

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерної інженерії та управління
(повна назва)

Кафедра Автоматизації проектування обчислювальної техніки
(повна назва)

АТЕСТАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти другий (магістерський)
(рівень вищої освіти)

Моделі та методи розпізнавання кіберфізичних об'єктів на основі використання метрики подібності – відмінності
(тема)

Виконав: студент II курсу, групи СКСм-19-1

Чугай В.А.
(прізвище, ініціали)

Спеціальність
123 – Комп'ютерна інженерія
(код і повна назва спеціальності)

Тип програми
освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма
Спеціалізовані комп'ютерні системи
(повна назва освітньої програми)

Керівник проф. Свірь І.Б.

Допускається до захисту

Зав. каф. АПОТ

(підпис)

Чумаченко С.В.
(прізвище, ініціали)

2020 р.
Харківський національний університет радіоелектроніки

Факультет Комп'ютерної інженерії та управління
Кафедра Автоматизації проектування обчислювальної техніки
Рівень вищої освіти другий (магістерський)
Спеціальність 123 – Комп'ютерна інженерія
Тип програми Освітньо-професійна
Освітня програма Спеціалізовані комп'ютерні системи

ЗАТВЕРДЖУЮ:

Зав. кафедри АПОТ 

Чумаченко С.В.

(підпис)

« » 2020 р.

ЗАВДАННЯ НА АТЕСТАЦІЙНУ РОБОТУ

студентові Чугаю Владиславу Андрійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Моделі та методи розпізнавання кіберфізичних об'єктів на основі використання метрики подібності – відмінності

затверджена наказом по університету від 30 жовтня 2020 р. № 1489 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 14 грудня 2020 р.

3. Вихідні дані до роботи комп'ютинг розпізнавання, моделі, цифрова схема, архітектура, моделювання

4. Перелік питань, що потрібно опрацювати в роботі Аналіз предметної області,

Аналіз джерел предметної галузі,

Метрика подібності-відмінності для кіберфізичного простору;

Методи визначення подібності двох об'єктів

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів)
слайди презентації


6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання	01.09.2020-05.09.2020	
2	Аналіз предметної області	07.09.2020-21.09.2020	
3	Аналіз джерел з проблемної галузі	22.09.2020-05.10.2020	
4	Метрика подібності-відмінності для кіберфізичного простору	06.10.2020-19.10.2020	
5	Методи визначення подібності двох об'єктів	20.10.2020-27.11.2020	
6	Оформлення пояснювальної записки	28.11.2020-30.11.2020	
7	Оформлення графічного матеріалу	01.12.2020-05.12.2020	
8	Перевірка виконаного проекту керівником	06.12.2020-09.12.2020	

Дата видачі завдання 01 вересня 2020 р.

Студент 
(підпис)

Керівник роботи _____

(підпис) 

проф. каф. АПОТ Свірь І.Б.,
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 60 с., 23 рис., 17 джерел.

МОДЕЛЬ, МЕТОД, АРХІТЕКТУРА, КІБЕРФІЗИЧНИЙ КОМП'ЮТИНГ, СОЦІАЛЬНИЙ КОМП'ЮТИНГ, КОМП'ЮТИНГ РОЗПІЗНАВАННЯ, МЕТРИКА ПОДІБНОСТІ – ВІДМІННОСТІ

У магістерській роботі розглядаються питання, пов'язані зі створенням моделей, методів, елементів архітектури комп'ютингу розпізнавання кіберфізичних об'єктів на основі використання метрики подібності – відмінності.

Мета – суттєве зменшення обчислювальної складності алгоритмів комп'ютингу розпізнавання кіберфізичних об'єктів шляхом розробки ефективної інфраструктури матричних даних (моделей, методів, архітектур) з використанням метрики подібності-відмінності.

Задачі дослідження орієнтовані на створення моделей, методів і архітектур комп'ютингу розпізнавання кіберфізичних об'єктів. Для досягнення поставленої мети необхідно вирішити такі задачі:

- проаналізувати сучасні технологічні тенденції та особливості побудови програмно-апаратних рішень;
- розширити моделі, методи, архітектури комп'ютингу розпізнавання кіберфізичних об'єктів.

Об'єкт дослідження – кіберсоціальні/кіберфізичні процеси і явища, автоматично керовані комп'ютинговими сервісами на основі використання метрики подібності – відмінності.

Предмет дослідження – матричні структури, які використовуються для створення методів і алгоритмів паралельного розподілу та об'єднання векторів

рядків і стовпців по матриці подібності-відмінності для синтезу архітектури кіберфізичного комп'ютингу розпізнавання.

ABSTRACT

Master's thesis: 60 pages, 23 figures, 17 sources under the list of references.

MODEL, METHOD, ARCHITECTURE, CYBERPHYSICAL COMPUTING, SOCIAL COMPUTING, RECOGNITION COMPUTING, SIMILARITY METRIC - SIMILARITY

The master's thesis deals with issues related to the creation of models, methods, elements of the architecture for computer recognition of cyberphysical objects based on the similarity - difference metrics.

The goal is to significantly reduce the computational complexity of cyberphysical object recognition computing algorithms by developing an efficient matrix data infrastructure (models, methods, architectures) using similarity-difference metrics.

The objectives of the study are focused on creating models, methods and architectures for computer recognition of cyberphysical objects.

To achieve this goal it is necessary to solve the following tasks:

- analyze modern technological trends and features of software and hardware solutions;

- expand models, methods, architectures of computer recognition of cyberphysical objects.

The object of study – cybersocial / cyberphysical processes and phenomena, automatically controlled by computer services based on the use of similarity metrics - differences.

The subject of research - matrix structures that are used to create methods and algorithms for parallel distribution and combining vectors of rows and columns on

the metrics of similarity-difference for the synthesis of the architecture of cyberphysical recognition computer.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	7
ВСТУП	8
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	9
1.1 Технологічні тенденції 2020-2021 від компанії Gartner	9
1.2 Програмно-апаратні рішення	13
1.3 Висновки до розділу 1	18
2 МЕТРИКА ПОДІБНОСТІ - ВІДМІННОСТІ ДЛЯ КІБЕРФІЗИЧНОГО ПРОСТОРУ	20
2.1 Комп'ютинг критичних систем	20
2.2 Єдність подібності-відмінності об'єктів або компонентів	25
2.3 Метрика і операції подібності-відмінності для пошуку дефектів при аналізі стовпців	29
2.4 Висновки до розділу 2	32
3 МЕТОДИ ВИЗНАЧЕННЯ ПОДІБНОСТІ ДВОХ ОБ'ЄКТІВ	33
3.1 Частотно-множинний метод визначення подібності двох об'єктів	33
3.2 Частотно-векторна модель і метод для обчислення подібності.	39
3.3 Секвенсор апаратного обчислення подібності-відмінності-включення об'єктів	45
3.4. Синтез цифрового SD-автомата для визначення подібності- відмінності об'єктів	50
3.5 Висновки до розділу 3	53
ВИСНОВКИ	54

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	55
ДОДАТОК А Тези доповіді, сертифікат	57
ДОДАТОК Б Слайди презентації	61

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ
І ТЕРМІНІВ

IoT – Internet of Things;

IoB – інтернет поведінки;

Amazon Web Services – AWS;

AI – Artificial Intelligent;

FDSOI – fully depleted silicon-on-insulator;

FinFET – Fin Field-Effect Transistor

ВСТУП

У роботі розглядаються питання, пов'язані зі створенням моделей, методів, елементів архітектури комп'ютингу розпізнавання кіберфізичних об'єктів на основі використання метрики подібності – відмінності.

Мета дослідження – суттєве зменшення обчислювальної складності алгоритмів комп'ютингу розпізнавання кіберфізичних об'єктів шляхом розробки ефективної інфраструктури матричних даних (моделей, методів, архітектур) з використанням метрики подібності-відмінності.

Сутність дослідження – створення комп'ютингу точного пошуку даних на основі аналізу регулярних матричних структур, які використовуються для створення методів і алгоритмів паралельного розподілу та об'єднання векторів рядків і стовпців по метриці подібності-відмінності.

Задачі дослідження орієнтовані на удосконалення та створення моделей, методів і архітектур комп'ютингу розпізнавання кіберфізичних об'єктів. Для досягнення поставленої мети необхідно вирішити такі задачі:

- проаналізувати сучасні технологічні тенденції та особливості побудови програмно-апаратних рішень;
- розширити моделі, методи, архітектури комп'ютингу розпізнавання кіберфізичних об'єктів.

Об'єкт дослідження – кіберсоціальні/кіберфізичні процеси і явища, автоматично керовані комп'ютинговими сервісами на основі використання метрики подібності – відмінності.

Предмет дослідження – регулярні матричні структури, які використовуються для створення методів і алгоритмів паралельного розподілу та об'єднання векторів рядків і стовпців по метриці подібності-відмінності для синтезу архітектури кіберфізичного комп'ютингу розпізнавання.

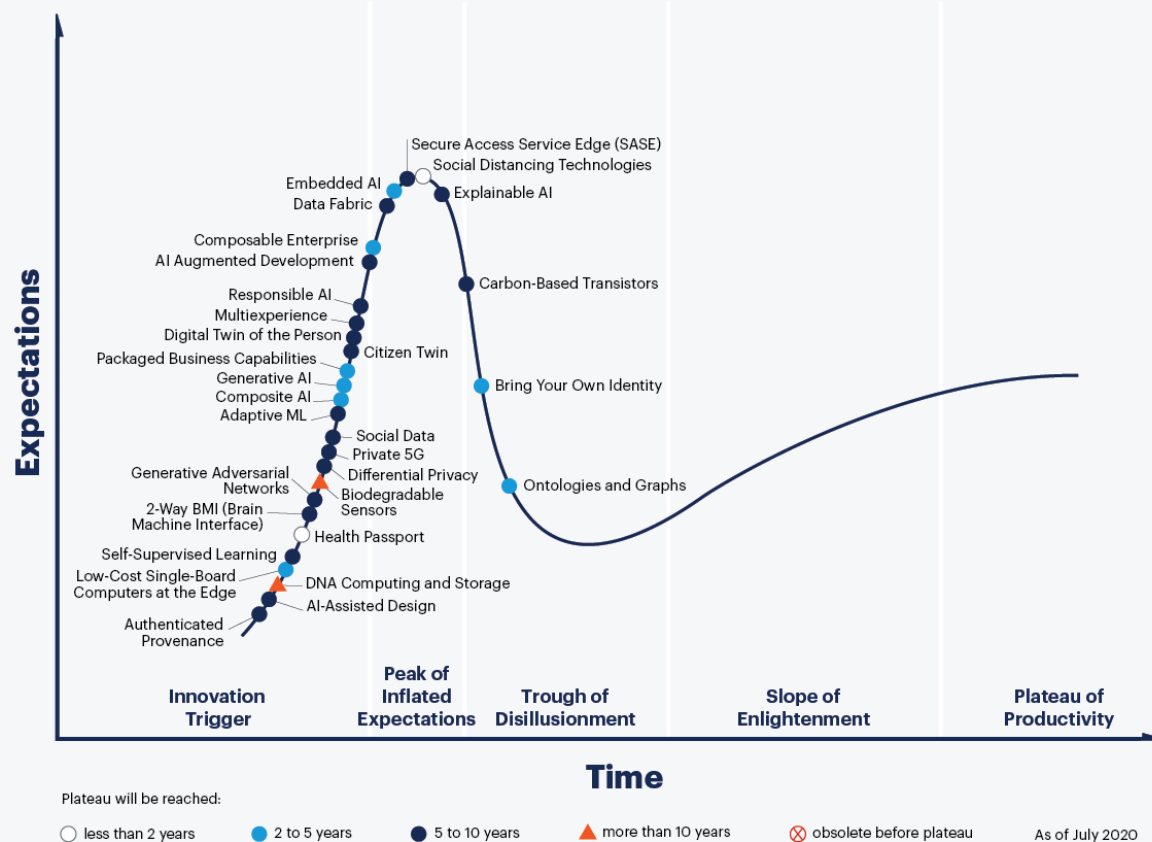
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

Дається огляд новітніх програмно-апаратних рішень, орієнтованих на створення ефективних засобів, які дають можливість здійснювати пошук даних в кіберпросторі за прийнятний час.

1.1 Технологічні тенденції 2020-2021 від компанії Gartner

В середині жовтня 2020 року компанія Gartner оприлюднила список основних технологічних тенденцій на 2021 рік (рис. 1.1).

Hype Cycle for Emerging Technologies, 2020



gartner.com/SmarterWithGartner

Source: Gartner
© 2020 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner and Hype Cycle are registered trademarks of Gartner, Inc. and its affiliates in the U.S.

Gartner

Рисунок 1.1 – Прогноз Gartner по розвитку нових технологій [1]

Термін «інтернет поведінки» (IoB) аналітики від Gartner вперше запропонували в прогнозі ще на 2020 рік як технології для моніторингу поведінкових явищ і управління даними у кіберпросторі, які на них впливають. До них відносяться: розпізнавання осіб, відстеження місця розташування і великі дані. Gartner прогнозує, що до кінця 2025 року понад половини населення світу буде задіяно хоча б в одній програмі IoB.

Технології комірчастой мережі дозволяють отримати доступ до будь-яких цифрових активів незалежно від місця положення активу або самої людини. Оскільки цифрові активи виходять за межі брандмауера, особливо при використанні хмарних технологій та дистанційної роботи, стають актуальними

нові підходи до кібербезпеки, зокрема, технології комірчастої мережі. Gartner прогнозує, що до 2025 року мережу кібербезпеки буде підтримувати більш половини запитів на управління цифровими правами.

Технології узагальненого досвіду об'єднують досвід клієнтів, співробітників і користувачів. З огляду на перехід діяльності в віртуальну сферу, що актуалізовано пандемією, подібні інструменти особливо затребувані. Gartner прогнозує, що компанії, які надають узагальнений досвід, зможуть обігнати конкурентів за ключовими показниками протягом найближчих трьох років.

Інтелектуальний складовий бізнес має на увазі, що додатки і пакетні бізнес-можливості розглядаються як окремі функціональні блоки, доступні через API. Вони можуть розроблятися всередині компанії або надаватися постачальниками, а нова структура дозволить об'єднати бізнес-можливості окремих пакетів і баз даних. Такий підхід дозволить компаніям проводити швидку реконфігурацію бізнесу.

Аналітики відзначають підвищення попиту на автоматизацію повторюваних ручних процесів і завдань (*гіперавтоматизація*). При цьому спостерігається перехід від автоматизації окремих завдань до автоматизації процесів з декількома завданнями, а також до функціональної автоматизації декількох процесів і навіть до автоматизації на рівні бізнес-екосистеми.

Вседоступна операційна модель ІТ стане ще одним трендом у 2021 році, оскільки передбачає інваріантність ведення бізнесу по відношенню до робочого місця, що в період пандемії стало особливо затребуваним. Gartner прогнозує, що до кінця 2023 року 40% компаній будуть застосовувати нову модель для віртуальної і фізичної взаємодії з клієнтами і співробітниками.

При *проектванні систем штучного інтелекту* тільки 53% проектів проходять шлях від прототипів до виробництва. Тому згідно з дослідженням Gartner, ще одним трендом стане забезпечення надійної структури, яка створить основу для проектування, масштабування і переходу нових систем штучного інтелекту у виробництво.

Тренд «розподілена хмара» передбачає розподіл загальнодоступних хмарних сервісів по різних фізичних локалізаціям, при якому за експлуатацію, управління і розвиток послуг відповідає постачальник даної хмари. Gartner прогнозує, що до 2025 року більшість платформ хмарних сервісів надаватимуть принаймні кілька розподілених хмарних послуг.

Конфіденційність даних стає все більш серйозною проблемою і до 2025 року половина великих компаній буде впроваджувати обчислювальні системи, які забезпечують конфіденційність даних в ненадійних середовищах.

Наприкінці серпня 2020 року компанія Gartner також визначила 5 головних ІТ-трендів, які суттєво вплинуть на бізнес, суспільство і кожну людину в найближчі 5-10 років.

Композитна бізнес-архітектура. Така модель дозволяє компаніям перейти від жорсткого традиційного планування до гнучкого реагування на потреби бізнесу. Модульна бізнес-модель ґрунтується на чотирьох принципах: модульність, ефективність, постійне вдосконалення і адаптивні інновації. Вона створює можливості для впровадження інноваційних підходів, знижує витрати і покращує партнерські відносини. Інші технології в рамках нової бізнес-моделі включають "пакетні" бізнес-послуги, фабрики даних, приватні 5G мережі, вбудований штучний інтелект.

Алгоритмічні моделі довіри забезпечують конфіденційність і безпеку даних, відстежують їх походження, а також підтверджують ідентичність людей і речей. Gartner вважає, що підвищений інтерес до blockchain призведе до розширення можливостей цифрової аутентифікації і перевірки. Серед технологій, пов'язаних з алгоритмічною довірою, аналітики відзначають безпечний доступ до послуг (SASE, Secure Access Service Edge), а також відповідальний і зрозумілий штучний інтелект, тобто алгоритм з простежуваними етапами.

Технології без кремнію. Згідно до закону Мура, кількість транзисторів в щільній інтегральній схемі буде подвоюватися кожні два роки, але технологія швидко наближається до фізичних меж кремнію. В результаті стали з'являтися

нові матеріали з розширеними можливостями, які дозволяють робити технології компактніше і швидше. Наприклад, ДНК-обчислення використовують ДНК і біохімічні реакції замість кремнію або квантових архітектур для виконання обчислень або зберігання даних. Інші технології в цій галузі включають біорозкладні датчики і транзистори на основі вуглецю.

Штучний інтелект, що формує, здатний динамічно змінюватися, щоб реагувати на ситуацію, що має кілька підтипів: здатний динамічно адаптуватися з часом, здатний створювати нові моделі для вирішення конкретних проблем, генеративний штучний інтелект, який може створювати новий контент (зображення, відео) або змінювати вже існуючий, а також композитний штучний інтелект, диференціальна конфіденційність, "малі дані" і навчання з самоконтролем.

Цифровізація особистості пов'язана з технологіями створення цифрових версій людини - цифрових двійників, які можуть існувати як у фізичному, так і у віртуальному просторі. Наприклад, двосторонній нейрокомп'ютерний інтерфейс як система для обміну інформацією між мозком і електронним пристроєм, яка може бути носієм або імплантом для реєстрації електроенцефалограми (ЕЕГ). Їх можна використовувати для ідентифікації, отримання доступу, оплати та аналітики. Такий інтерфейс також представляють собою додаткову уразливість, яку можуть використовувати зловмисники.

1.2 Програмно-апаратні рішення

Пропонується короткий огляд окремих новітніх програмно-апаратних рішень, орієнтованих на створення ефективних засобів і систем штучного інтелекту, які дають можливість здійснювати пошук даних в кіберпросторі за прийнятний час.

Історично виходить, що апаратні і програмні засоби комп'ютера на кожній стадії його розвитку складають між собою гармонійний альянс. Це означає, що

новим чіпам повинні відповідати нові програмні системи, так само як і під нові алгоритми і технології слід розробляти більш досконалі кремнієві кристали і структури. Не секрет, що провідна компанія планети Apple використовує стратегію вмирання старого заліза шляхом його примусової невідтримки новими програмними системами. Так виключається мезальянс між старим залізом і новими програмними додатками. Аналогічна ситуація має місце бути по відношенню до нового заліза і старих додатків. Зростання обчислювального потенціалу нових чіпів надає вченим можливість проектувати більш досконалі моделі, методи та алгоритми розв'язання актуальних ринкових і наукових завдань. Тому кожен фахівець з програмування, який бажає бути в тренді, повинен мати уявлення про проривні технології щодо нового покоління чіпів, комп'ютерів, мереж, дата центрів та хмарних сервісів.

1) Замість материнських плат [2] компанія SEA-Leti пропонує чіплет (рис. 1.2), який інтегрує цифрові системи на кристалах різних компаній за функціональністю і електроживленню. Система чіпів може пропускати 3 терабайта в секунду на одному квадратному міліметрі кремнію (150 мілліватт / 1 кв.мм потужність розсіювання) із затримкою 0,6 наносекунди на міліметр. Чіп SEA-Leti укладає шість 16-ядерних чіпсетів поверх тонкої пластини (інтерпозера) кремнію. Інтерпозер містить як схеми регулювання напруги, так і функціональну мережу, яка зв'язує різні компоненти з вбудованою пам'яттю до 2 терабайт. Активні вбудовані пристрої є кращий шлях інтегрування в системи розрізнених технологій і мікросхем від різних постачальників.

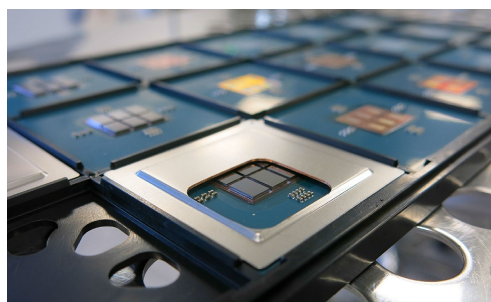


Рисунок 1.2 – Система чіплетов на силіконовій підкладці

2) Ізраїльський стартап Hailo [3] зібрав 60 мільйонів доларів для масового виробництва чіпа Hailo-8. Він призначений для глибокого навчання на автомобілях, роботах і інших «передових» машинах. Інноваційні мікросхеми призначені для зменшення витрат, розмірів і енергоспоживання при використанні штучного інтелекту для обробки з високою роздільною здатністю даних від датчиків, камер HD. Hailo ідентифікує свій проект структурно-визначеною архітектурою для потоку великих даних. Навчання називається «глибоким», тому що його нейронні мережі складаються з довгого стека нейронних шарів між входом і виходом. Програмні засоби Hailo використовують навчену нейронну мережу для перевірки можливостей пам'яті, обчислень і потоку даних, навантажених на кожен рівень (рис. 1.3). Додаток зіставляє інформацію з ресурсами процесора для формування структури, де кожен рівень є суміжним зі своїми сусідами. В результаті виходить мікросхема, яка може виконувати 26 трильйонів операцій в секунду і досягати 672 кадри на секунду за зразком класифікації зображень ResNet-50 з ефективністю 3,1 трильйона операцій в секунду на ват. Чіпи призначені для використання в периферійних системах, включаючи частково автономні транспортні засоби, інтелектуальні камери, смартфони, дрони і гаджети AR / VR, які вимагають складного глибокого навчання для точного виконання заданих функцій.

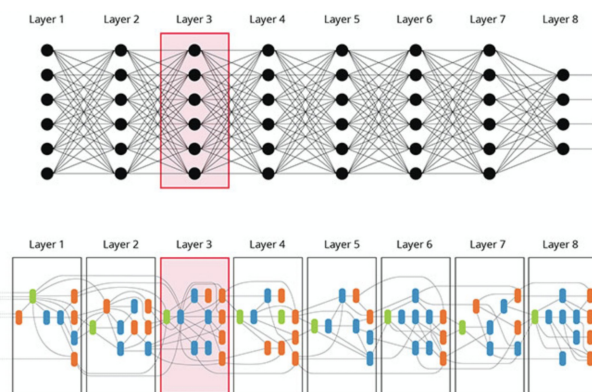


Рисунок 1.3 – Чіп Nailo і його структура для Machine Learning

3) Комп'ютер CS-1 [4] мав конструкцію, де близько трьох чвертей кожного шасі займали системи охолодження. У 2020 році сталася справжня революція: надзвичайно потужний комп'ютер, майже повністю складається з одного чіпа (рис. 1.4), розмір якого становить понад 46255 квадратних міліметрів, що більш ніж в 50 разів перевищує площу будь-якого іншого процесорного чіпа. Маючи 1,2 трильйона транзисторів, 400000 процесорних ядер, 18 гігабайт SRAM і міжкомпонентні з'єднання, здатні переміщати 100 мільйонів мільярдів біт в секунду, Wafer Scale Engine (WSE) Cerebras не піддається простому порівнянню з іншими системами. За даними компанії, 10-стійковий кластер TPU2 компанії Google AI споживає в п'ять разів більше енергії і займає в 30 разів більше місця, забезпечуючи лише одну третину продуктивності від комп'ютера WSE. Масивний чип є відповіддю, якої чекала спільнота штучного інтелекту: моделі нейронної мережі становляться все більш складними, здатність швидко навчатися або перенавчатися може бути вирішена комп'ютером WSE Cerebras.

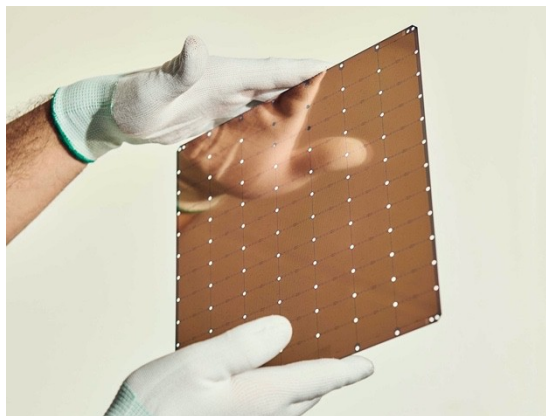


Рисунок 1.4 – Нейро-комп'ютер Wafer Scale Engine Cerebras

4) Наступний найпримітивніший пристрій складає основу для створення

як потужних дата центрів і суперкомп'ютерів, так і для кінцевих гаджетів і сенсорних пристроїв. Польовий (field-effect transistor) транзистор [5] складається з чотирьох основних частин: джерела (source), стоку (drain), каналу (conductive channel), який з'єднує їх, і затвора (gate) для управління потоком струму по каналу (рис. 1.5). Оскільки всі компоненти стали менше, то поведінка транзисторів істотно змінюється у бік порушення працездатності при тривалому їх використанні. Chenming Hu – батько FinFETs – вважав фундаментальну проблему абсолютно ясною для себе: зробити канал дуже тонким, щоб електрони не могли проходити крізь затвір. Рішення використовували витончення оксидного шару затвора для кращого контролю над каналом і зменшення струму витoku. Його дослідження показали, що підхід стоншування близький до межі, а саме: зробити шар оксиду досить тонким, щоб електрони могли перестрибнути через нього в кремнієву підкладку, створюючи ще одне джерело витoku. Натомість він запропонував дві інших ідеї. Перша – додати шар ізоляції, прихований в кремнії під транзистором, для утруднення електронам пробиратися крізь затвор зарядів. Цю конструкцію стали називати повністю збідненим кремній на ізоляторі або FDSOI (fully depleted silicon-on-insulator). Друга пов'язана з наданням затвору більшого контролю над потоком заряду за рахунок розширення тонкого каналу вертикально над підкладкою, схожою з акулячим плавником (shark's fin), щоб затвор міг оточувати або огортати (wrap) канал з трьох сторін, а не просто сидіти зверху. Ця структура називається FinFET (Fin Field-Effect Transistor), яка має додаткову перевагу, коли використання простору по вертикалі усуває деякі затори на площині 2D, відкриваючи еру тривимірних транзисторів. Особливості FinFET: Кожен транзистор має витік, стік, канал і затвор для управління струмом по каналу. Підняття каналу над поверхнею чіпа дозволяє затвору обертати його з трьох сторін, створюючи більший контроль.

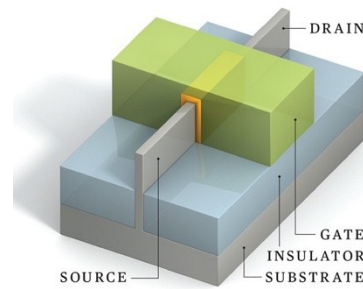


Рисунок 1.5 – Сьогодення та майбутнє мікроелектроніки:
FinFET або Fin Field-Effect Transistor

Новий швидкісний пристрій (рис. 1.6) машинного зору [6] може обробляти зображення в тисячі разів швидше існуючих аналогів за допомогою датчика зображення, що використовує штучну нейронну мережу. Технологія машинного зору часто відчуває затримки сканування пікселів рядок за рядком, перетворення відеокадрів в цифрові сигнали і передачі даних на комп'ютери для аналізу. Вчені з TU Wien прискорили машинний зір, прибравши посередника. Вони створили датчик зображення – штучну нейронну мережу, яка може одночасно отримувати і аналізувати дані. Потім вчені створили датчик – нейронну мережу, що навчається, на основі зв'язків між фотодіодами для класифікації зображень, включаючи літери. Датчик зображення не споживає електроенергію під час роботи. Самі виявлені фотони забезпечують енергію для роботи нейромережі. Звичайна технологія машинного зору зазвичай здатна обробляти до 100-1000 кадрів на секунду. Для порівняння: TU Wien система працює з 20 мільйонами кадрів на секунду. Швидкодію системи обмежено тільки швидкістю електронів в ланцюгах. Стратегія може працювати на рівні пикосекунд, або трильйонних доль секунди, що на три-чотири порядки швидше існуючих аналогів. Датчик орієнтований на пошук і аналіз даних в візуалізації динаміки рідини, процесів горіння або механічних поломок. Машинний зір в автономному водінні вимагає інших ефективних рішень.

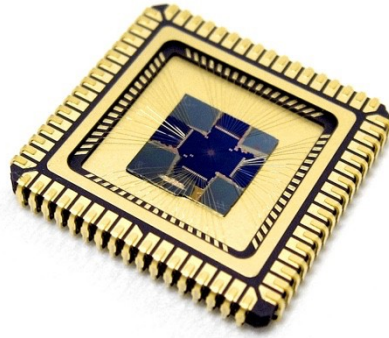


Рисунок 1.6 – Нейропроцесор для аналізу візуальних потоків даних

1.3 Висновки до розділу 1

Таким чином, навіть окремі кращі рішення в області комп'ютингу аналізу і пошуку даних свідчать про існування стійкого тренда на інтелектуалізацію «заліза» шляхом імплементації апаратних рішень в алгоритми штучного інтелекту і машинного навчання, що створюють засоби для високопродуктивних обчислень, необхідних сьогодні в кіберфізичному просторі для аналізу великих даних. Відомо, що моделі, методи, алгоритми та програми, пов'язані зі штучним інтелектом, прагнуть в своєму розвитку від імовірнісних характеристик процесів до детермінізму кінцевих автоматів, закладених спочатку в сучасний комп'ютинг. Ідеальним описом деякого процесу або явища завжди була детермінована таблиця істинності від n змінних, як деяку межу знань про об'єкт дослідження [7].

Тому шляхи вирішення актуальних завдань по розпізнаванню патернів, образів, станів будуть формуватися у зустрічних напрямках:

- 1) Від ймовірного незнання процесу до його автоматного детермінізму через час навчання.
- 2) Від комп'ютерного детермінізму, як ядра деякого знання, до

багатопараметричної деталізації аналізованого процесу або явища. Другий шлях передбачає розробку ефективної метрики подібності-відмінності, структур даних і алгоритмів для аналізу і навчання спочатку детермінованих систем.

2 МЕТРИКА ПОДІБНОСТІ - ВІДМІННОСТІ ДЛЯ КІБЕРФІЗИЧНОГО ПРОСТОРУ

Розглядається комп'ютинг критичних систем. Дається універсальна метрика пошуку даних в кіберпросторі на основі використання параметрів подібності-відмінності і матричної структури в двійковій формі. Описуються методи аналізу матричних структур даних по метриці подібності-відмінності для пошуку дефектів в цифрових системах.

2.1 Комп'ютинг критичних систем

Критична система являє собою сукупність взаємопов'язаних в

кіберфізичному просторі і часі відносин між компонентами для досягнення поставленої мети, відмови яких призводять до значних економічних, політичних, соціальних, екологічних і гуманітарних втрат. Прикладами критичних систем виступають технологічні і технічні об'єкти в галузях: енергетика, транспорт, промисловість, озброєння, кіберсоціальна сфера, банкінг, інтернет, державність, юриспруденція. Однозначно вчені і фахівці дійшли висновку, що 80% всіх відмов у критичних системах пов'язано з людським фактором. Отже, необхідно виключати людину з циклу моніторингу-управління шляхом передачі повноважень по прийняттю рішень детермінованому і практично безпомилковому комп'ютингу: мережевого, хмарного, термінального.

Комп'ютерна інженерія є галузь знань, яка займається теорією і практикою проектування, тестування, виробництва і експлуатації захищених програмно-апаратних масштабованих обчислювальних засобів для надійного метричного управління віртуальними, фізичними і соціальними процесами і явищами шляхом використання інтелектуальних хмарних і телекомунікаційних сервісів на основі цифрового моніторингу кіберфізичного простору за допомогою персональних гаджетів і вбудованих розумних сенсорів. Тут комп'ютинг, як глобальна методологія, яку має і комп'ютерна інженерія, є стратегія досягнення і візуалізації поставленої мети – створення продукції і / або сервісів при заданих ресурсах, яка системно представляється процесом моніторингу та актуації метричних відносин в замкнутій інфраструктурі управління і виконання. Взаємні відносини між комп'ютерною інженерією та комп'ютингом подано на рис. 2.1.

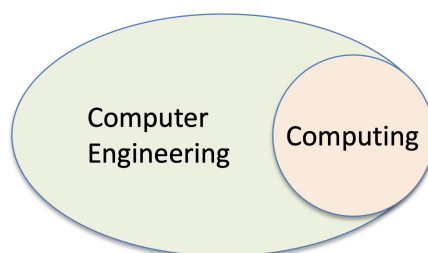


Рисунок 2.1 – Відносини між інженерією та комп'ютигом

Масштабований комп'ютинг як процес ціледосягнення має наступні найбільш характерні області застосування (рис. 2.2): космологічний, біологічний, природний, технологічний, кіберсоціальний, кіберфізичний, глобальний, транспортний, індустриальний, науково-освітній, Cloud-Edge, медичний, Cyber Security, Quantum, Design and Test.

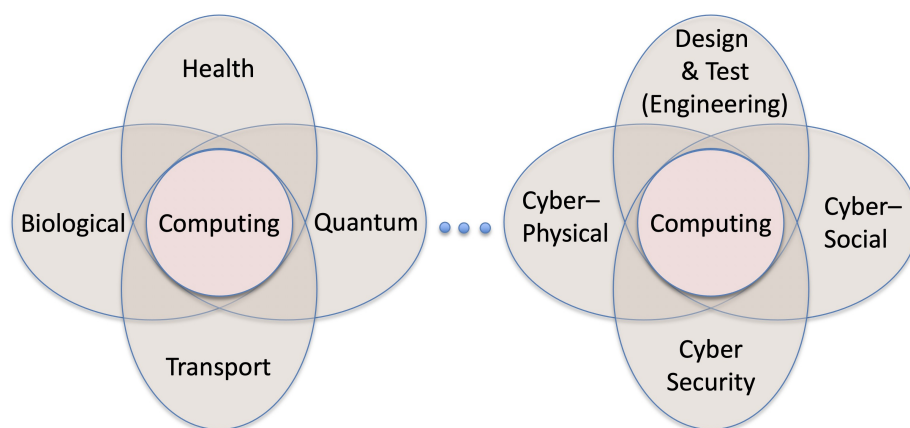


Рисунок 2.2 – Відносини між комп'ютигом і природою

Це означає, що методологія комп'ютингу однакова для дослідження всіх процесів і явищ природи і людської діяльності, а відмінність в представлених відносинах полягає лише в утриманні об'єктів досліджень, які своїми примітивами створюють постійно розширюваний універсум нових знань про комп'ютинговий розвиток природи.

Комп'ютинг системно може бути представлений (рис. 2.3) процесом моніторингу (5) і актуації (6) метричних відносин (2) в інфраструктурі управління (3) і виконання (4) для досягнення і візуалізації (8) поставленої мети – продукції і / або сервісів (1) при заданих ресурсах (7).

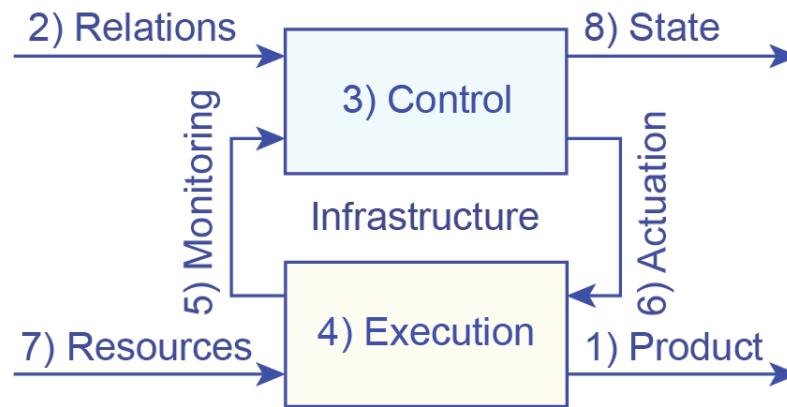


Рисунок 2.3 – Комп'ютинг критичної системи управління персоналом

Метричне і структурне визначення комп'ютингу за допомогою восьми взаємопов'язаних компонентів надає теоретичну фундаментальну основу для формального і фактичного створення системи цифрового управління будь-яким процесом в заданій сфері діяльності. Види комп'ютингу за введеною метрикою охоплюють всі сфери людської діяльності: космологічний, біологічний, флористичний, фізичний, віртуальний, квантовий, соціальний, державний, медичний, транспортний, інфраструктурний, науковий, освітній, виробничий, спортивний, відпочинку, подорожей, розваг. Природно, що в першу чергу комп'ютинг орієнтований на моніторинг і управління критичними об'єктами, процесами і явищами без- або з мінімальною участю людини.

Оскільки сьогодні не можна поки обійтися без людини, як одного з компонентів управління, то необхідно мінімізувати його можливі помилки при прийнятті оперативних і стратегічних рішень у критичній системі. Виконання завдання пов'язане з цифровізацією історії, знань, умінь і навичок кожного співробітника на основі детермінованої метрики, попередньо сформованої експертами.

Відношення розглядається як структура взаємопов'язаних компонентів, що визначає матеріально-енергетичні властивості процесу або явища в часі і просторі. Структура взаємозв'язків елементів визначає властивості надійності і продуктивності системи, процесу або явища, але ніяк не навпаки. Інакше,

заміщення одних елементів на інші в заданій структурі не змінює суті системи.

Таким чином, щоб змінити метричні параметри будь-якої системи, необхідно створити нову структуру відносин між її компонентами. Відомим прототипом тут виступає blockchain & smart contract computing, який реалізує онлайн моніторинг і управління всіма процесами в оцифрованому суспільстві на основі 5G-технологій, хмарних сервісів і edge computing.

Метрика якості критичної системи визначається параметрами: надійність, безвідмовність, довговічність, ремонтпридатність, збереженість, тестопригодність, керованість, спостережуваність, діагностованість, обслуговування, контролепригодність, безпека і живучість. Відмова є подія, що полягає в порушенні працездатного стану об'єкта [9]. Критичність відмови – сукупність ознак, що характеризують наслідки відмови, визначається прямими і непрямими втратами, витратами праці і часу на усунення наслідків відмов, можливістю і доцільністю ремонту, тривалістю простоїв, ступенем зниження продуктивності [10].

Що стосується критичних ситуацій і відмов, то в даний час в кіберфізичному просторі є вичерпна інформація про будь-який негативний процес або явище, яке можна запобігти засобами інтелектуального хмарного і edge computing моніторингу-управління, що становить сутність critical system computing (рис. 2.4).

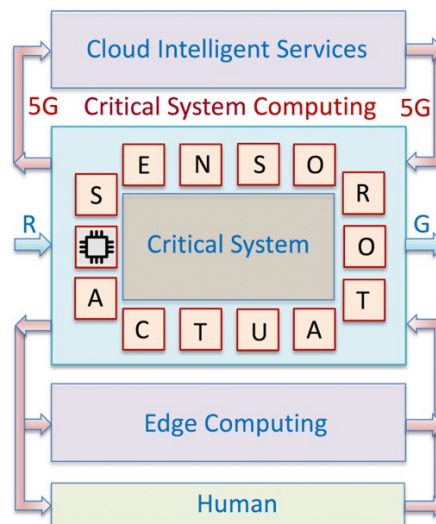


Рисунок 2.4 – Комп'ютинг критичної системи

Тут два обчислювача (хмарний і термінальний) обслуговують критичну систему за допомогою сенсорних датчиків і актюаторів. Природно, що хмарний комп'ютинг інваріантний до часу і геопозиції критичного об'єкта, наприклад, автомобіля (Synopsys, GMC, Tesla). Якість і надійність тут забезпечуються стандартами: JTAG IEEE 11.49, SECT IEEE 1500, JTAG IEEE 1687, ISO 9001. Технології граничного сканування згаданих стандартів створюють додаткові лінії і spare-компоненти, що дозволяють досягти високих рівнів якості і надійності за рахунок online тестування і відновлення працездатності критичних систем з використанням вбудованих засобів BIST і хмарних test services.

Природно, тільки компетентний оператор має можливість втручатися в роботу критичної системи за допомогою термінального комп'ютера. Для цього вичерпний моніторинг компетенцій кожної людини досить просто здійснити за допомогою пошукових та спеціальних додатків з метою подальшого прийняття актюаторних рішень про призначення працівника на функціональну позицію, що становить сутність personnel computing або HR-Management. Для вимірювання компетенцій працівників, представлених векторами змінних, використовується метрика Левенштайна, яка дає можливість визначити схожість-відмінність між претендентами і зразковим патерном, а також квазі-оптимальні маршрути трансформування однієї метрики-моделі в іншу.

2.2 Єдність подібності-відмінності об'єктів або компонентів

Метрика подібності-відмінності об'єктів, процесів або явищ (рис. 2.5), представлена рівнянням $SD=ab=U=1$, є базисом для вирішення найпоширеніших на технологічному ринку завдань, пов'язаних з вимірюваннями в кіберфізичному просторі і інтернеті [8].

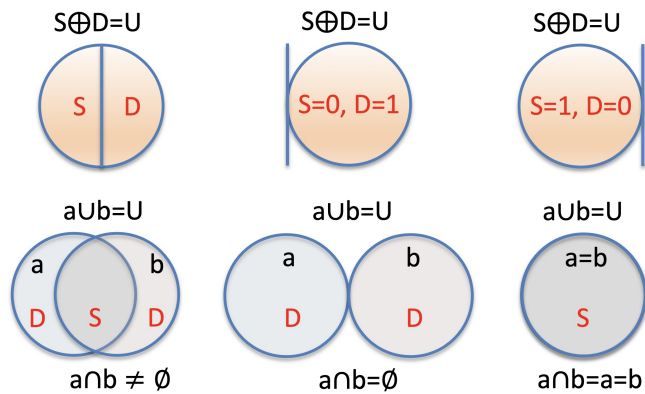


Рисунок 2.5 – Метрика відносини між схожістю-відмінністю двох об'єктів

Такими науково-практичними завданнями (рис. 2.6) є: 1) Розпізнавання образів і патернів. 2) Машинне навчання. 3) Прийняття рішень. 4) Управління процесами і об'єктами. 5) Аналітика великих даних. 6) Цифровизація процесів і явищ. 7) Тестування і діагностика систем. 8) Створення і використання регулярних баз даних.

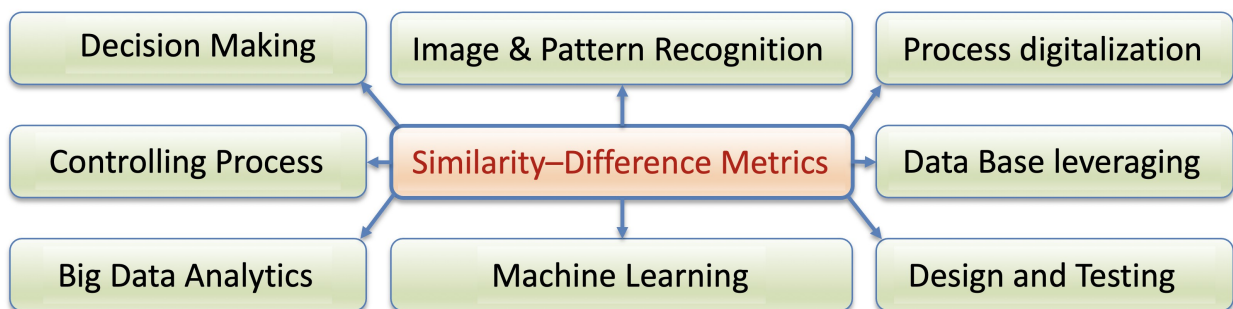


Рисунок 2.7 – Технологічний ринок завдань

Модель відносин між процесами і/або явищами оптимально представляється у вигляді двійковій (бінарної) матриці M , яка формує в загальному випадку декартів добуток двох множин, наприклад, множину тестів на множину функціональностей: $\langle TF \rangle$.

Подібність розглядається як відношення кількості однакових компонентів процесу або явища до їх загальної кількості в матриці заданих істотних параметрів. Різниця – як відношення кількості відмінних один від одного компонентів процесу або явища до їх загальної кількості в матриці заданих істотних параметрів. Подібність і відмінність є оцінка відношення між процесами і явищами.

Для визначення скалярної оцінки подібності між усіма парами компонентів необхідно з матриці M побудувати квадратичну матрицю їх подібності (TT or FF) шляхом використання формули, яка визначає відносини подібності двох компонентів:

Подібність визначається непустим перетином вектора-рядка матриці M з диз'юнкцією всіх інших рядків. Формула для обчислення скалярної оцінки подібності між компонентами зводиться до визначення подібності між поточним рядком і диз'юнкцією-об'єднанням всіх інших рядків, знайдених в матриці M (один з усіма, але в форматі вектора):

Результатом виконання даної формули є скалярна оцінка, яка формує нормоване відношення кількості одиничних координат, спільних для даного компонента і вектора логічного об'єднання всіх інших компонентів, до числа одиничних координат аналізованого рядка. Сукупність таких скалярних оцінок для всіх рядків матриці M : дає можливість ранжувати їх за ступенем убудання подібності. Природно, що рядки з максимальними оцінками подібності слід виключати з сукупності валідних компонентів, перетин яких між собою повинно в ідеалі бути рівним порожній множині. Доповнення: якщо порахувати

різницю для компонента , а потім всіх їх підсумувати, то вийде оцінка сукупностей непересічних частин компонентів, які формують інтегральну функціональність матриці. Розподіл даної оцінки на довжину матриці буде визначати відсоток корисної функціональності:

Після виключення невалідних компонентів необхідно виконувати процедуру верифікації шляхом визначення їх спільної загальної подібності, сенс якої полягає у підрахунку одиничних координат після одночасного перетину всіх векторів , віднесених до суми одиничних координат, отриманих після об'єднання векторів, що в сукупності дає нуль, як позитивний результат:

Дана формула визначає рівень спільності, якщо значення отриманого результату відмінно від нуля. В цьому випадку необхідно виключати з множини вектори, які мають високий рівень рейтингу за метрикою подібності з іншими рядками. Для масштабного сприйняття оцінок подібності-відмінності необхідно також виводити як супровідний результат базу-метрику, яка представляє собою в даному випадку кількість одиничних координат розглянутого вектора. Аналогічно слід чинити і при верифікації компонентів шляхом їх спільного перетину, де слід виводити інформацію про сукупній кількості одиничних координат, отриманих в результаті об'єднання всіх векторів, що фігурують в знаменнику. Тоді вихідні дані будуть мати формат «схожість - число одиниць вектора»: .

Інтегральний рівень подібності в матриці є більш структурованою скалярною оцінкою в порівнянні з , яка визначається шляхом підсумовування 1-координат після об'єднання подібності всіх векторів щодо диз'юнкції інших,

поділеній на суму 1-координат після об'єднання всіх рядків матриці M:

Якщо результат , то кожен вектор не має подібності з іншими, що означає коректне розпізнавання компонентів-рядків в матриці M. В іншому випадку необхідно видаляти ті рядки, які мають максимальні оцінки подібності з іншими векторами. Висновки до отриманих формул можна зобразити у вигляді рис. 2.8.

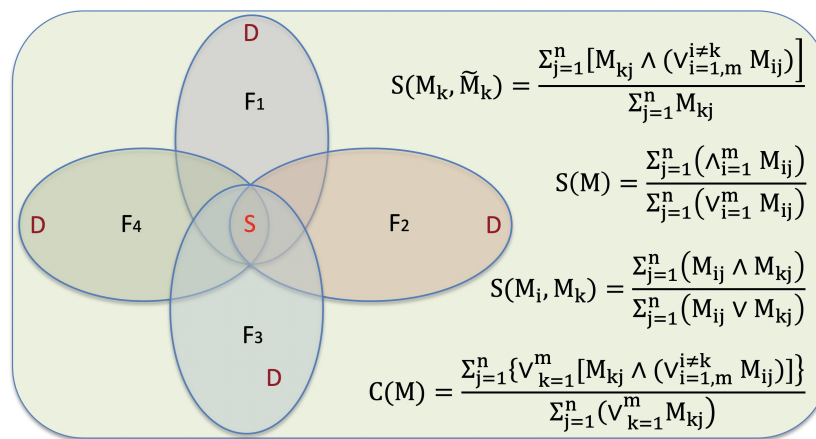


Рисунок 2.8 – Визначення подібності між парою векторів матриці M, один з рештою, а також всіма векторами між собою

Формули, що представлені спільно з діаграмами, створюють інструментарій для векторного і паралельного виконання процедур, пов'язаних з пошуком даних в матричному просторі з координатами, заданими в двійковому, чисельному або багатозначному вигляді.

2.3. Метрика і операції подібності-відмінності для пошуку дефектів при аналізі стовпців

Теоретико-множинна і векторно-логічна (кубітна) взаємодія константних дефектів і їх похідних (подібності-відмінності) визначається операціями (), які представлені нижче в таблицях [8,11]:

	0	1	X					10	01	11		
0	0		0					10	10	00	10	00
1		1	1					01	00	01	01	00
X	0	1	X					11	10	01	11	00
								00	00	00	00	

	0	1	X					10	01	11		
0	0	X	X					10	10	11	11	
1	X	1	X	1				01	00	01	01	11
X	X	X	X	X				11	11	11	11	11
	0	1	X					10	01	11	00	

a	0	1	X					10	01	11	
	1	0						01	10	00	

	0	1	X					10	01	11		
0			1					10	00	11	01	10
1			0					01	11	00	10	01
X	1	0						11	01	10	00	11
	0	1	X					10	01	11	00	

Слід зауважити, що виконання логічних операцій над векторами-кубітами є більш технологічним для сприйняття людиною і комп'ютером: перетинання-кон'юнкції, об'єднання-диз'юнкції, доповнення-заперечення, симетричної різниці - що виключає або. По суті, дані логічні операції становлять сутність подібності-відмінності процесів або явищ, які фігурують в метричному рівнянні пошуку даних [12]: $SD=ab=$. Розділяй і об'єднуй, виключаючи протиріччя – основна ідея пошуку. У разі діагностування згадана ідея може бути представлена розподілом сукупності векторів матриці на М два підмножини по

параметру $R = \{0,1\}$, як показано на рис. 2.9. Отримані підмножини векторів множини M в загальному випадку взаємодіють між собою за подібністю, що визначається відповідно до формули: $=$.

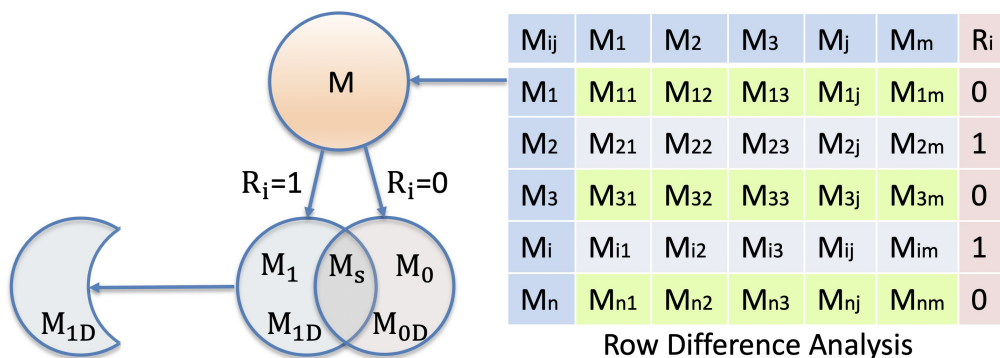


Рисунок 2.9 – Аналіз відмінності по рядках матриці

З огляду на той факт, що множина M і похідні підмножини створюються за допомогою елементів, які утворюють вектори, то актуальними є процедури, що формують подібності та відмінності між підмножинами :

$$= \setminus(or);$$

$$= \setminus ;$$

$$= ;$$

$$= .$$

Модифікація даних формул у бік більш точного вирішення завдань діагностування на основі додаткового аналізу подібності окремих векторів множини призводить до результату:

$$= \setminus or);$$

$$= \setminus ;$$

$$= ;$$

$$= .$$

Тут додається операція пошуку загальної частини або подібності між усіма векторами, які відносяться до підмножини S . Стосовно до класу константних несправностей, схожість асоціюється до пошуку спільних одиночних дефектів, які присутні у всіх векторах і відсутні у всіх векторах множини S . Що стосується пошуку кратних дефектів, то остання система рівнянь трансформується в сукупність виразів:

$$\begin{aligned}
 &= \text{or }); \\
 &= \setminus ; \\
 &= ; \\
 &= .
 \end{aligned}$$

Подані вирази орієнтовані на підвищення глибини діагностування, що означає зменшення потужності підмножини

2.4 Висновки до розділу 2

Описано комп'ютинг критичних систем. Модифіковано універсальну метрику пошуку даних в кіберпросторі на основі використання параметрів подібності-відмінності і матричної структури в двійковій формі. Описано методи аналізу матричних структур даних по метриці подібності-відмінності для пошуку дефектів в цифрових системах.

3. МЕТОДИ ВИЗНАЧЕННЯ ПОДІБНОСТІ ДВОХ ОБ'ЄКТІВ

Розглядаються частотно-множинний метод визначення подібності двох об'єктів, а також частотно-векторна модель і метод для обчислення подібності. Розробляється прототип секвенсора апаратного обчислення подібності-відмінності-включення об'єктів. Пропонується синтез цифрового SD-автомата для визначення подібності-відмінності об'єктів.

3.1 Частотно-множинний метод визначення подібності двох об'єктів

Технологічним ядром для вирішення практичних завдань управління персоналом в критичних системах є кіберфізичний комп'ютинг, оформлений в структури Machine Learning і SCADA - Supervisory Control And Data Acquisition. Такий комп'ютинг передбачає аналітику великих даних, яка використовує примітивні теоретико-множинні операції, процедури та алгоритми паралельної дії з метою підвищення продуктивності при пошуку квазіоптимальних рішень. У середовищі комп'ютингу пошуку на анвансцену виходить метрика подібності-відмінності. Тому важливо мати ефективний спеціалізований процесор, як найпростіше ядро, для паралельного і високопродуктивного рішення задач синтезу і аналізу. Структурно метрика подібності-відмінності двох процесів, явищ, об'єктів, компонентів використовує дві формули, які оперують у бінарній алгебрі логіки двома паралельними операціями `and`, хог для отримання результуючих векторів:

;

.

Але такі формули мало що дають для знання відносин між процесами (явищами), коли необхідно і дуже важливо визначити загальні структури даних, щоб зрозуміти, як трансформуються окремі компоненти (координати векторів) один в другий при синтезі і аналізі. Більш того, тут весь процес синтезу знаходиться в обчислювальній залежності від технологічно досконалих структур даних. Нормована метрика подібності-відмінності використовує дві формули, що також оперують в алгебрі логіки двома паралельними операціями, але доповнені арифметикою підрахунку одиничних координат, отриманих в результаті виконання логічних операцій. Крім того, з'являється спільний знаменник у вигляді диз'юнкції однойменних координат векторів, який служить інтегратором розрізнених структур даних процесів в загальний вектор тільки істотних координат, відносно яких виконується нормування подібності-відмінності:

;

.

Наприклад, два вектора $a = i$ і $b = j$, що мають несуттєві нульові однойменні координати, автоматично виключаються з нормованого оцінювання, завдяки врахуванню і підрахунку тільки одиничних значень в результуючих векторах:

;

;

Природно, немає потреби обчислювати обидві оцінки за даними формулами. Досить визначити одну з них, а другу можна отримати за формулою доповнення:

.

Тут відмінністю формованої оцінки від відстані Хеммінга є виключення з метрики і структур даних умови існування двох нулів на координатах з однаковими адресами-індексами, що істотно підвищує адекватність вимірювання двох процесів. Що стосується багатозначної алгебри (теорії множин), де замість алфавіту $\{0,1\}$ виступають символи, літери, цифри, слова, тексти, об'єкти, процеси, то схожість-відмінність, як правило, розглядається в рамках метрики або відстані Левенштайна. Тут фігурують три елементарних операції: заміна символів, вставка і видалення, які трансформують одне слово в інше.

Пропонується інший розв'язок визначення подібності між словами, який характеризується синтезом уніфікованої структури даних, вирівнює пари слів будь-якої довжини до однієї розмірності за рахунок виконання єдиної операції - вставки порожнього символу. Як наслідок, обчислювальна складність алгоритму для синтезу уніфікованої структури єдиної розмірності зводиться до пошуку місць для вставки кінцевого числа $n = 0,1,2,3, \dots$ порожніх символів з

метою вирівнювання довжини двох слів (об'єктів, процесів).

Як приклад, розглядається перетворення одного слова в інше шляхом вставки порожніх символів:

CONDUCTION
BOND IANA

Виконання алгоритму вставки порожніх символів з метою отримання мінімальної відмінності і максимальної схожості за умови трансформування одного слова в інше дає результат:

CONDUCTION -
BOND - - - IANA

Кількість порожніх символів для вирівнювання двох слів дорівнює чотирьом. Після цього здійснюється тривіальний підрахунок відстані Левенштайна, яка дорівнює числу координат, мають різні символи в матриці трансформації слів, що означає , .

Таким чином, будь-яку пару процесів або явищ можна привести до структурної матриці однакової довжини з метою подальшого підрахунку нормованих оцінок подібності-відмінності шляхом арифметичного додавання виконання логічних умов в чисельнику і знаменнику:

;

Для заданого прикладу трансформованої взаємодії пари слів нормовані оцінки подібності-відмінності мають вигляд:

;

Більш складна конструкція нормованої подібності-відмінності визначається не за рівністю, а за приналежністю однієї координати вектора-слова іншій координаті другого вектора, якщо координати представлені деякими множинами. В цьому випадку формули для обчислення оцінок матимуть вигляд:

;

Тут корисними будуть теоретико-множинні операції для алфавіту Кантора, наприклад, які визначені такими квадратичними таблицями істинності:

	0	1	X	
0	1		0,5	
1		1	0,5	
X	0,5	0,5	1	

	0	1	X	
0	1	1	1	0,5
1	1	1	1	0,5
X	1	1	1	1
	0,5	0,5	1	

	0	1	X	
0			0,5	0,5
1			0,5	0,5
X	0,5	0,5		
	0,5	0,5	1	

Елементарні таблиці дають можливість привести теоретико-множинні операції до їх норм, складання яких формує точні оцінки подібності-відмінності. Наприклад, для наступних двох багатозначних векторів $a = 1XXX10X1$, $b = 01X00XX1$, оцінки подібності-відмінності, отримані за чисельними таблицями істинності, матимуть вигляд:

;

.

Загальна структура визначення подібності-відмінності пари векторів за три векторних паралельних чисельних операції () представлена у наступній таблиці:

a	1	X	X	X	1	0	X	1
b	0	1	X	0	0	X	X	1
	0	0,5	1	0,5	0	0,5	1	1
	1	0,5	0	0,5	1	0,5	0	0
	1	1	1	1	1	1	1	1

Таким чином, отримані дві оцінки подібності, які взаємно-доповнюють один одного до 1:

;

Координати наступних векторів проміжних обчислень також є взаємно-

доповнюючими до 1, що є умовою валідації процесу визначення подібності-відмінності:

	0	0,5	1	0,5	0	0,5	1	1
	1	0,5	0	0,5	1	0,5	0	0

Обчислювальна складність алгоритму синтезу структурної уніфікованої метрики трансформування одного слова в інше дорівнює

3.2 Частотно-векторна модель і метод для обчислення подібності

Багатозначна структура пари векторів, що відповідають множині слів-примітивів (\cdot), може бути використана для ефективного визначення подібності текстових фрагментів, а також обчислення рівня плагіаризма. Спрощена діаграма вирішення даного завдання за допомогою трьох векторних логічних операцій може бути представлена у вигляді рис. 3.1.

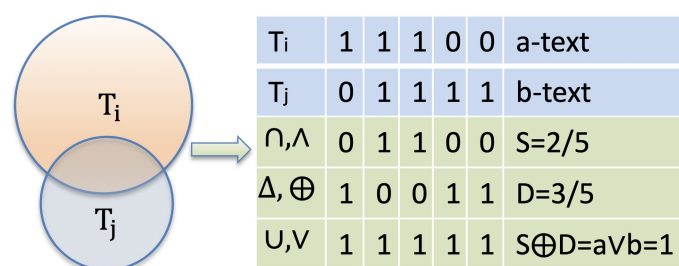


Рисунок 3.1 – Діаграма пошуку подібності текстових фрагментів

У координатах замість одиничних значень можуть бути представлені слова або будь-які інші дані, які фігурують у векторах-множинах (,). Однак попереднє унітарне кодування слів або пропозицій істотно полегшує виконання алгоритму визначення подібності-відмінності. Слід зауважити, що замість двійкового коду координати вектора можуть бути відзначені частотністю слів або даних у формі дійсних або цілих чисел, а також часовими або іншими параметрами компонентів, що не змінює суті алгоритму для метричного нормованого оцінювання подібності текстових фрагментів.

Характеристика метрики:

1) Визначається унікально істотними змінними для вектор-множин, взаємодіючих між собою.

2) Не має змінних, які не істотні для взаємодіючих множин.

3) Враховує частоту зустрічальності кожного компонента для обчислення подібності-відмінності.

4) Є універсальною моделлю для визначення подібності-відмінності будь-яких дискретних процесів і явищ.

Частотно-векторна модель двох взаємодіючих підмножин для визначення подібності-відмінності обробляється за такими формулами:

;

.

Тут в логічних операціях беруть участь потужності підмножин - частотності - кожного компонента, які істотно підвищують адекватність подібності-відмінності текстів, процесів і явищ. Така модифікація формул, де фігурують три операції: вибір мінімального значення з двох однойменних координат (, вибір максимального значення на парі координат (, обчислення модуля різниці

двох значень координат, обробляє не тільки частотність, а й чисто двійкове подання вектор-множин. Однак арифметичні операції тут набувають недолік – не можна використовувати паралельні обчислення над координатами векторів.

Переваги запропонованої метрики подібності процесів і явищ:

1) Інваріантність частотно-множинного уявлення примітивів-даних, в порівнянні з кортеж-орієнтованою відстанню Левенштайна, дає можливість зменшити обчислювальну складність алгоритму визначення подібності-відмінності від експоненційної до квадратичної. Дана перевага дає можливість адекватно оцінювати плагіаризм для текстів на слов'янських мовах, де допускається зміна порядку слів у реченнях.

2) Векторна, апаратно орієнтована, модель унітарного кодування множин примітивів-слів дає можливість обчислювати схожість-відмінність за один автоматний такт.

3) Синтезована унікальна метрика також показує шлях перетворення одного тексту в інший, а також обчислювальну складність такої трансформації, яка вимірюється відмінностями в вектор-множинах.

4) Частотно-векторна структура є універсальною моделлю для визначення подібності-відмінності будь-яких дискретних процесів і явищ для вирішення завдань перетворення одного об'єкта в інший, прийняття рішень, пошуку дефектів, класифікації та кластеризації даних.

У загальному випадку для вимірювання відносин між двома об'єктами або процесами формуються чотири оцінки (рис. 3.2):

1) Подібність – загальний критерій в og -метриці двох текстів.

2) Різниця, як інтегральна оцінка в og -метриці, яка доповнює подібність до одиниці – 1.

3) Норма відмінності для першого тексту, що приведена до загальної og -метриці текстів.

4) Норма відмінності для другого тексту, приведена до загальної og -метриці текстів. Норми відмінності для двох текстів складають інтегральну або

загальну відмінність в og-метриці.

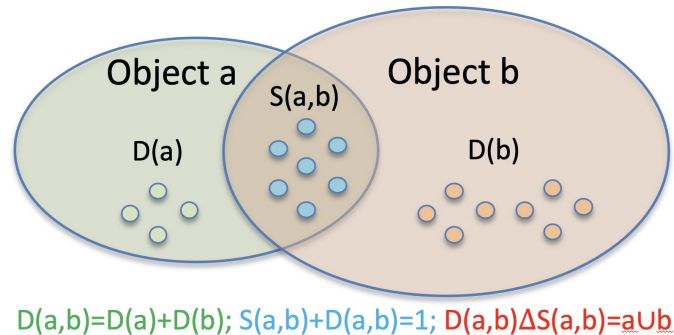


Рисунок 3.2 – Метричні оцінки подібності-відмінності

Слід зауважити, що норма подібності, що приведена до загальної og-метриці об'єктів, є однаковою. Іноді існує необхідність в обчисленні норм подібності-відмінності, наведених до кожного об'єкту, який бере участь у взаємодії. З урахуванням сказаного, кількість оцінок подібності-відмінності для визначення ставлення дорівнюватиме восьми.

Істотність og-метрики для визначення чотирьох оцінок подібності-відмінності проявляється також при формуванні маршруту трансформування одного явища, об'єкта чи процесу в інший шляхом усунення відмінностей між ними. Завданнями, які підпадають під згадану істотність, є:

1) Маршрут навчання, що використовує існуючий статус студента і його кінцеву мету - соціальну позицію.

2) Маршрут соціального визнання, що використовує існуючий статус громадянина і його кінцеву мету - соціальну позицію, посаду, нагороду, гроші.

3) Маршрут наукового визнання, що використовує існуючий статус вченого і задану кінцеву мету - нобелівську або державну премію.

4) Маршрут перемоги у виборах, використовує існуючий статус громадянина і метрику-патерн бажаної посади.

5) Маршрут трансформації ХНУРЕ у Стенфорд на основі порівняння їх метрик.

6) Маршрут трансформування відсталого підприємства (університету, держави, конституції) в успішне, цифрове на основі визначення відмінностей двох метрик. Актуальність успішності компанії шляхом цифровізації суб'єкта ринку представлена в статистиці використання cloud-edge computing:

Digital Services - Unique companies, as users	
Cloud: 5,111,272	Microsoft Office 365: 148,013
ERP: 243,065	Microsoft Azure: 24,978
CRM: 203,035	Microsoft Dynamics: 80,333
Accounting: 401,597	AWS: 1,369,944
HCM: 265,049	Salesforce: 114,870
BI: 175,850	SAP: 60,018

Тут ERP - Enterprise Resource Planning, CRM - Customer Relationship Management, HCM - Human Capital Management, BI - business intelligence, AWS - Amazon Web Services, SAP ERP - enterprise resource planning software, створений німецькою компанією SAP SE.

Цифровізація – створення глобальної інфраструктури для масштабованого тотального комп'ютингу на основі ідентифікації всіх процесів і явищ в кіберфізичному просторі з метою інтелектуального метричного управління об'єктами на основі вичерпного моніторингу для досягнення якості життя і збереження екології планети (рис. 3.3).

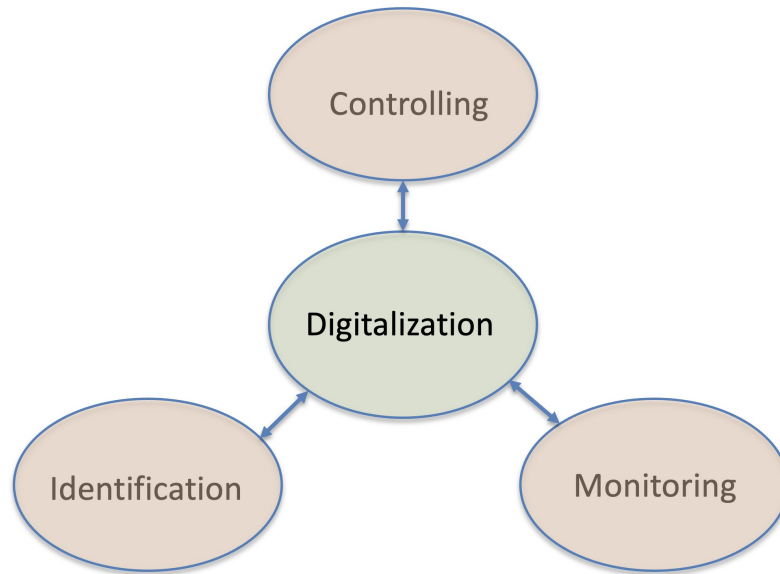


Рисунок 3.3 – Компоненти цифровізації кіберфізичного простору

При цьому складність виникає тільки під час розв'язання задачі цифрової ідентифікації громадян. Всі інші питання вирішуються за допомогою розумних сенсорів і актюаторів, які виконують cloud-edge комп'ютинг в повній мірі над будь-якими об'єктами, включаючи: безпілотний транспорт, медицину, побут, відпочинок, харчування, торгівлю, кадри, освіту, державні послуги, індустрію, включаючи критичні системи.

7) Маршрут трансформування (корекції) несправного продукту, цифрової системи, програмного додатка в справний на основі визначення відмінностей двох метрик.

8) Маршрут трансформування деструктивного генома вірусу в корисний білок на основі визначення відмінностей двох метрик, або вироблення антитіл, що нейтралізують деструктивні геноми вірусів.

3.3 Секвенсор апаратного обчислення подібності-відмінності-включення об'єктів

Мета – визначення якісної і кількісної метричної взаємодії між процесами і явищами шляхом використання паралельних регістрових логічних операцій.

Завдання полягає у синтезі логічної структури для обчислення метричної взаємодії між процесами і явищами, заданими в двійковому коді.

Абсолютні оцінки відмінності і подібності не мають великого сенсу при обчисленні взаємодії між множинами. Більш інформативними є нормовані оцінки подібності-відмінності, наведені до знаменника у вигляді суми істотних координат двох векторів, що виключають тільки нульові за однойменними координатами значення. Знаменник для отримання нормованих оцінок виходить шляхом підсумовування всіх (одиничних) координат після виконання операції диз'юнкції $N=$. Чисельником може бути відстань по Хеммінгу, яка формує абсолютну оцінку відмінності між двійковими векторами. Але далі включається модифікація, пов'язана з розподілом подібності або відмінності на кількість одиничних координат, отриманих після логічного додавання двійкових векторів. Дане доповнення робить нормовану оцінку відмінності (подібності) більш значущою за рахунок зменшення знаменника.

Таким чином, мають місце дві нормованих оцінки:

=

Відмінність даного підходу від Хеммінга і Левенштайна та інших метрик [13-17] полягає у видаленні чисто **нульових однойменних координат** зі структури даних для оцінювання подібності-відмінності.

Наприклад, нехай є два вектора або рядки, що відповідають об'єктам:

money	1	0	1	1	0	0	0	1	0	1
things	0	1	1	1	0	0	0	0	1	1

Рядки повинні бути трансформовані до наступної мінімальної структури, які мають тільки значущі стовпці, які обов'язково містять одиницю при їх диз'юнкції:

money	1	0	1	1	1	0	1
things	0	1	1	1	0	1	1

Потім до рядків застосовується процедура обчислення подібності-відмінності за наведеними вище формулами:

Природно, що .

Крім того, можна підрахувати похідну між нормами відмінності-подібності

Тут, шляхом віднімання або без нього видно, яка оцінка більше - схожість або відмінність. Однак в даній метриці краще уникнути появи негативних чисел, які потребують додаткової інтерпретації.

Таким чином, процедура підрахунку наведених оцінок подібності-

відмінності зводиться до виконання трьох векторних паралельних операцій ($\&$) з подальшим підрахунком одиниць в отриманих результуючих векторах. Якісний аналіз для прийняття рішення про включення ключового даного в клас еквівалентності заснований на оцінках максимальної схожості або мінімальної відмінності двох рядків-об'єктів.

Цифрова логічна схема підрахунку подібності-відмінності представлена на рис. 3.4.

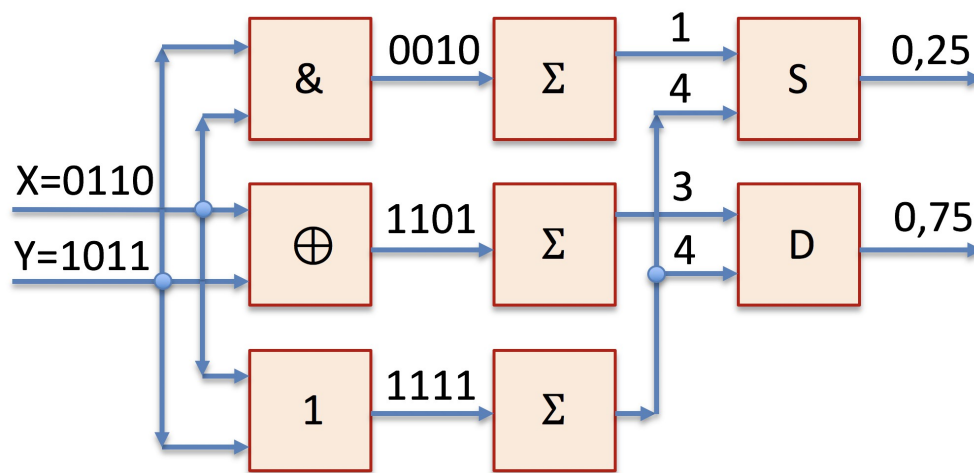


Рисунок 3.4 – Логічна схема обчислення оцінок подібності об'єктів

Недоліком наведеної LS-метрики визначення подібності є нерозрізнення відношень включення одного об'єкту в інший. Усунути дану недосконалість можна тільки шляхом хог-порівняння вихідних векторів з результатами їх перетину:

Відношення включення виконується, якщо значення дробу в виразі дорівнює нулю:

Одночасна наявність нулів на обох виходах свідчить про еквівалентність векторів. Модифікована логічна схема обчислення оцінок подібності об'єктів, що дозволяє визначати відносини включення представлена на рис. 3.5.

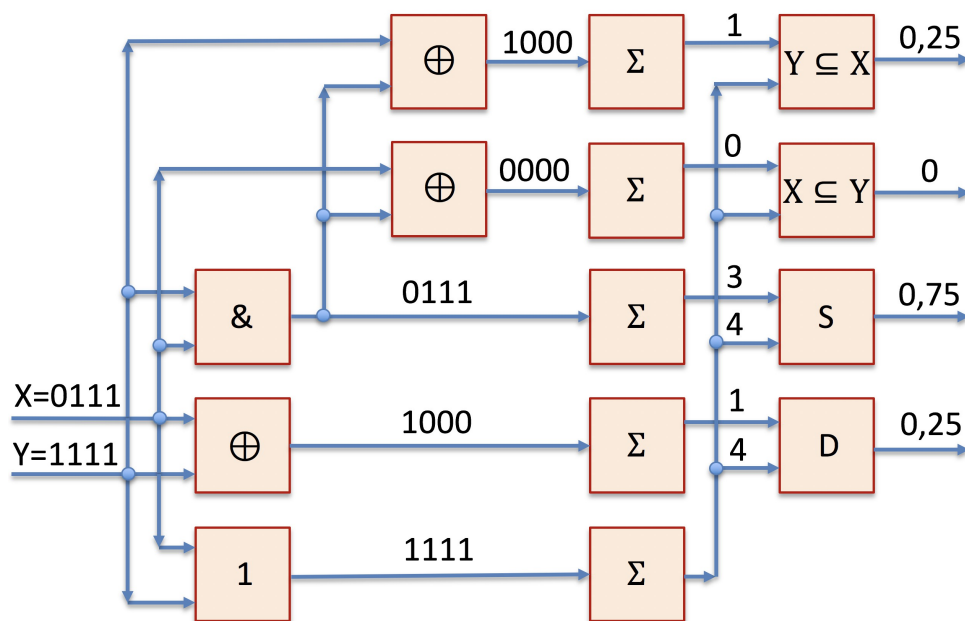


Рисунок 3.5 – Схема визначення норм подібності і відношення включення 0-сигналом

Важливим також представляються не тільки загальні симетричні оцінки

подібності-відмінності в метриці вирівняних по довжині векторів, але і який відсоток від кожного вектора-ключа займають дані оцінки, що формують асиметрію відносин. Асиметричні відносини дають можливість визначити пріоритети двох векторів або який з них вибрати як головний.

Наприклад, для двох векторів $A = 1011111$, $B = 0111110$ ідентифікатором спільності буде вектор A , оскільки вага його істотних одиничних координат відмінності більше, ніж у векторі B : $D1(A)=2 > D1(B)=1$ при наявності домінуючої ненульової подібності (непорожнього перетину) над інтегральної відмінності двох векторів $S(A) = 4 > D(B) = 3$. В результаті пишуться два даних через конкатенацію, а вектор-рядки об'єднуються. Перевага даного модифікованого методу вимірювання подібності-відмінності полягає в еквівалентності довжин ключів-векторів, що йде від метрики Левенштайна, де також немає і не може бути нульових однойменних координат (вектор-стовпців), що дорівнюють нулю. Наслідком такого факту слід вважати повну симетрію абсолютних і відносних оцінок подібності-відмінності, але не асиметричної, в загальному випадку, кількості одиниць в векторах, наведених до загальної довжини векторів:

$$A = \text{Image} = 0111111;$$

$$B = \text{Mirror} = 1011110.$$

Симетрія оцінок подібності-відмінності (рис. 3.6) зовсім не означає симетрію рівності числа одиничних значень координат в векторах, що дає можливість метрично визначати верховенство одного вектора над іншим для подальшої ідентифікації загального об'єднання двох векторів.

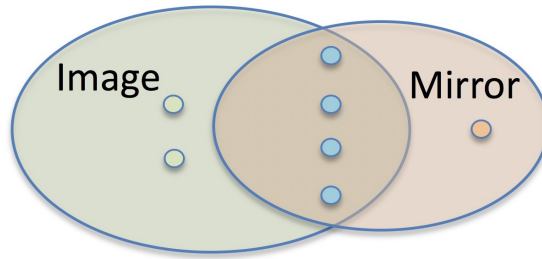


Рисунок 3.6 – Асиметрія подібності-відмінності для чистого Левенштайна

3.4. Синтез цифрового SD-автомата для визначення подібності-відмінності об'єктів

Структура секвенсора для визначення структурних оцінок подібності-відмінності-включення представлена на рис. 3.7.

Тут є 4 шари перетворення даних:

- 1) Формування вхідних даних про об'єкти.
- 2) Синтез загальної метрики для вимірювання об'єктів.
- 3) Унітарне кодування об'єктів в синтезованої метриці параметрів.
- 4) Подача векторів, що відповідають об'єктам, на входи SD-автомата.
- 5) Обчислення-формування чотирьох виходів-значень, що визначають відносини між об'єктами.

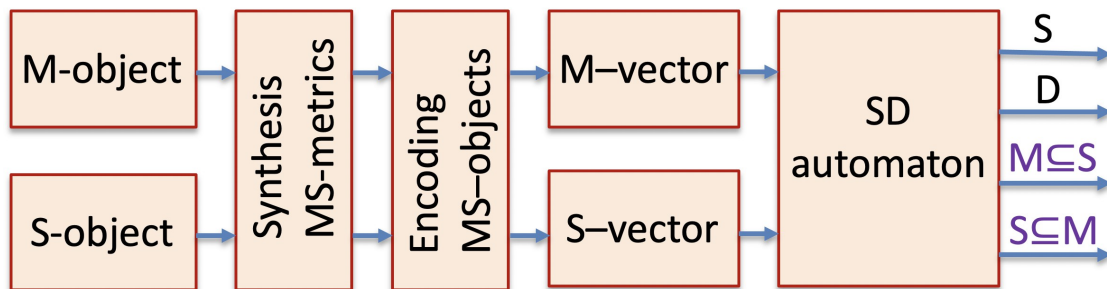


Рисунок 3.7 – Секвенсор визначення відношень між об'єктами

Структурна схема пристрою, реалізованого у вигляді SD-автомата (Similarity-Difference) на мові опису апаратури Verilog для визначення подібності-відмінності між двома об'єктами, зображена на рис. 3.8.

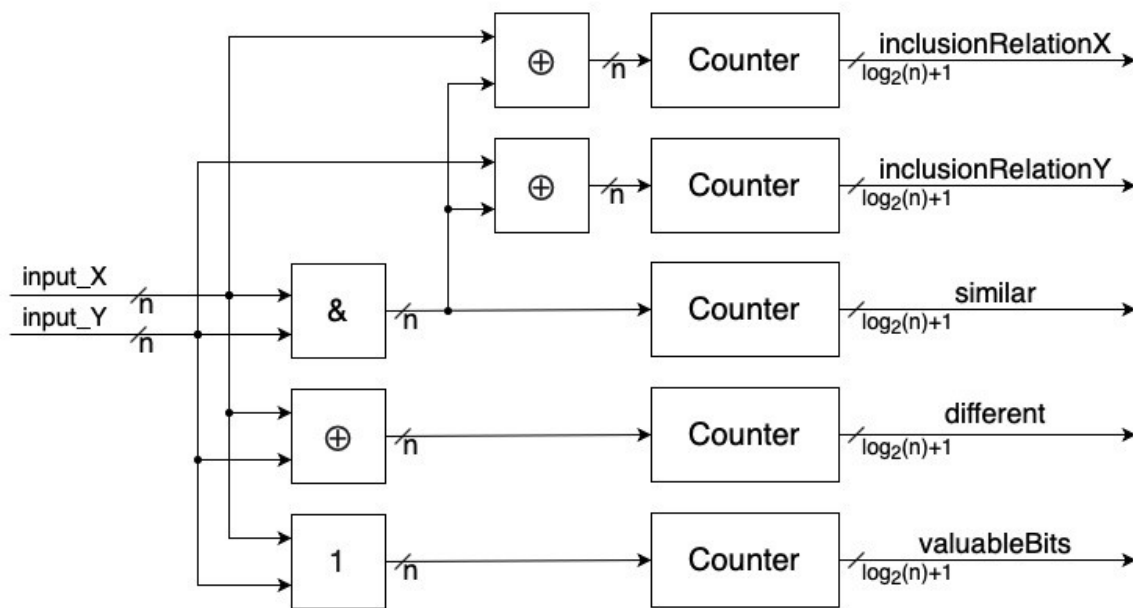


Рисунок 3.8 – Структурна схема SD-автомата

Пристрій виконує підрахунок кількості співпадаючих ненульових бітів (similar), що розрізняються (different), значущих координат (valuableBits), а також ступеня входження X і Y (inclusionRelationX і inclusionRelationY відповідно). Координати вважаються значущими, якщо хоча б в одному з вхідних векторів у біті по даній координаті зустрічається одиниця. Для більш точного підрахунку схожості / відмінності відбувається видалення чисто нульових однойменних координат X і Y - враховуються тільки ненульові. Підрахунок кількості співпадаючих ненульових бітів відбувається в два етапи. На першому етапі виконується операція «&». Після цього отриманий вектор співпадаючих ненульових бітів подається на лічильник, в якому виконується

підрахунок кількості одиниць в даному векторі.

Лічильник може бути реалізований декількома способами. Перший - комбінаційна схема. В даному випадку процес підрахунку кількості одиниць відбувається на логічній схемі, де затримка між подачею вхідних впливів і отриманням результату на виході формується найдовшим ланцюгом логічних елементів комбінаційної схеми. Варіант дозволяє максимально скоротити час підрахунку результату, але він є витратним по апаратурі логіки.

Другий спосіб - використання лічильника, який тактується зовнішнім пристроєм і використовує тригери. Варіант використовує мінімум апаратури, але має великі часові витрати для отримання результату.

Третій спосіб заснований на комбінації двох згаданих вище варіантів. Реалізувати підрахунок одиниць у векторі можна шляхом використання лічильника на тригерах, який містить комбінаційну схему. Це дозволяє за один автоматний такт проводити перевірку декількох координат, зменшуючи тим самим час, необхідний для отримання вихідного результату.

Використовуючи поєднання варіантів, можна домогтися паритету між прийнятною розмірністю апаратури і часом підрахунку подібності-відмінності. На практиці можуть бути реалізовані всі варіанти. Це залежить від обмежень і потреб ринку (замовника), для якого проектується даний модуль.

Для підрахунку метрик подібності та відмінності в діапазоні від нуля до одиниці необхідно зробити розподіл чисел *similar* і *different* на число *valuableBits* відповідно. Залежно від специфікації проекту дана операція може виконуватися як безпосередньо в даному модулі, так і на процесорному блоці, а при відсутності необхідності може взагалі не виконуватися, оскільки часто значення зберігаються в цілих числах. Те ж саме стосується і метрики приналежності.

3.5 Висновки до розділу 3

Розглянуто частотно-множинний метод визначення подібності двох об'єктів, а також частотно-векторна модель і метод для обчислення подібності.

Запропоновано метод і апаратна реалізація секвенсора для пошуку подібності-відмінності-включення, який характеризується отриманням структурованої оцінки взаємодії двох об'єктів, яка дає можливість більш точно визначати шляхи трансформування одного об'єкта в інший, а також вибирати більш суттєві об'єкти з пари при прийнятті рішення.

Запропоновані методи можуть бути застосовані для пошуку одиночних і кратних дефектів в цифрових програмних і апаратних системах. Показано, що методи аналізу можна використовувати для матриць, заданих чисельними параметрами в довільній формі, а також для пошуку даних, пов'язаних з ідентифікацією процесів і явищ в кіберпросторі.

Розроблені методи орієнтовані на апаратну реалізацію виконання паралельних логічних реєстрових операцій, які забезпечують істотне (на порядок) підвищення швидкодії в порівнянні з існуючими аналогами, виконаними в програмному коді.

Обмеження запропонованих методів пошуку даних пов'язані з необхідністю виконання процедури синтезу матричних структур даних, яка має квадратичну обчислювальну складність.

ВИСНОВКИ

Виконано огляд новітніх програмно-апаратних рішень, орієнтованих на створення ефективних засобів і систем штучного інтелекту, машинного навчання, які дають можливість здійснювати пошук даних в кіберпросторі за прийнятний час.

Запропоновано метрика пошуку даних в кіберпросторі на основі використання параметрів подібності-відмінності по матричній структурі в двійковій формі, що носить універсальний характер.

Розроблено методи аналізу матричних структур даних по метриці подібності-відмінності, які можна застосовувати для пошуку даних в довільних матричних структурах.

Аналіз літературних джерел та інтернет-ресурсів за тематикою роботи дає можливість зробити наступні висновки:

1) існуючі рішення в області комп'ютингу аналізу і пошуку даних свідчать про тенденції на інтелектуалізацію «заліза» шляхом імплементації апаратних рішень в алгоритми штучного інтелекту і машинного навчання, створюють засоби для високопродуктивних обчислень, необхідних сьогодні в кіберфізичному просторі для аналізу великих даних;

2) моделі, методи, алгоритми, пов'язані зі штучним інтелектом, наближаються до детермінізму кінцевих автоматів, закладених спочатку в сучасний комп'ютинг. Ідеальним описом тут виступає таблиця істинності від n змінних як база даних щодо об'єкту дослідження.

Таким чином, вирішення актуальних завдань з розпізнавання патернів, образів, станів повинні включати розробку ефективної метрики подібності-відмінності, структур даних і алгоритмів для аналізу і навчання від початку детермінованих систем.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. <https://www.gartner.com/smarterwithgartner/5-trends-drive-the-gartner-hype-cycle-for-emerging-technologies-2020/>
2. <https://spectrum.ieee.org/tech-talk/semiconductors/processors/core-processor-chiplets-isscc-news>
3. <https://spectrum.ieee.org/tech-talk/semiconductors/processors/israeli-edgeai-startup-gets-60million-preps-for-mass-production>
4. <https://spectrum.ieee.org/semiconductors/processors/cerebrass-giant-chip-will-smash-deep-learnings-speed-barrier>
5. <https://spectrum.ieee.org/semiconductors/devices/how-the-father-of-finfets-helped-save-moores-law>
6. <https://spectrum.ieee.org/tech-talk/computing/hardware/image-neural>
7. Abramovici M. Digital System Testing and Testable Design / M. Abramovici, M.A. Breuer and A.D. Friedman.- Comp. Sc. Press. – 1998. – 652 p.
8. Hahanov V. Cyber Physical Computing for IoT-driven Services. New York. Springer. 2018.
9. ГОСТ 20911–89. Техническая диагностика. Термины и определения.
10. ГОСТ 27.002-89. Надежность в технике. Основные понятия. Термины и определения.
11. [Vladimir Hahanov](#); [Svetlana Chumachenko](#); [Eugenia Litvinova](#); [Abdullayev Vugar Hacimahmud](#); [Anastasia Hahanova](#); [Tetiana Soklakova](#). [Cyber Social Computing. 2018 IEEE East-West Design & Test Symposium \(EWDTS\)](#). 2018.
12. M. Karavay, V. Hahanov, E. Litvinova, H. Khakhanova and I. Hahanova, "Qubit Fault Detection in SoC Logic," 2019 IEEE East-West Design & Test Symposium (EWDTS), Batumi, Georgia, 2019, pp. 1-7.
13. Tanimoto T.T., An elementary mathematical theory of classification and

prediction, IBM Report. (November 1958), cited in: G. Salton, Automatic Information Organization and Retrieval (McGraw-Hill, 1968) p. 238.

14. Temma S., Sugii M. and Matsuno H., "The Document Similarity Index based on the Jaccard Distance for Mail Filtering," 2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), JeJu, Korea (South), 2019, pp. 1-4.

15. Verma N. K., Dutta E. and Yan Cui, "Hausdorff distance and global silhouette index as novel measures for estimating quality of biclusters," 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Washington, DC, 2015, pp. 267-272.

16. [Бляшке. Круг и шар](#). Метрика Хаусдорфа. – М.: Наука, 1967.

17. Jaccard P. Distribution de la flore alpine dans le Bassin des Dranses et dans quelques regions voisines // Bull. Soc. Vaudoise sci. Natur. — 1901. — V. 37, Bd. 140. — S. 241—272.

1. ДОДАТОК А

Тези доповіді, сертифікат

УДК 658:512.011: 681.326: 519.713

1. Технічні науки

МОДЕЛІ ТА МЕТОДИ РОЗПІЗНАВАННЯ КІБЕРФІЗИЧНИХ ОБ'ЄКТІВ НА ОСНОВІ ВИКОРИСТАННЯ МЕТРИКИ ПОДІБНОСТІ – ВІДМІННОСТІ

В.А. Чугай,

*магістрант кафедри автоматизації
проектування обчислювальної техніки*

С.В. Чумаченко

*д-р техн. наук, професор,
завідувач кафедри автоматизації
проектування обчислювальної техніки, науковий керівник
Харківський національний університет радіоелектроніки
м. Харків, Україна*

Питання, пов'язані зі створенням моделей, методів, елементів архітектури комп'ютингу розпізнавання кіберфізичних об'єктів на основі використання метрики подібності-відмінності, є актуальними у зв'язку з необхідністю створення ефективних засобів, які дають можливість здійснювати пошук даних в кіберпросторі за прийнятний час [1].

Мета дослідження – суттєве зменшення обчислювальної складності алгоритмів комп'ютингу розпізнавання кіберфізичних об'єктів шляхом розробки ефективної інфраструктури матричних даних (моделей, методів, архітектур) з використанням метрики подібності-відмінності. Задачі дослідження орієнтовані на удосконалення та створення моделей, методів і архітектур комп'ютингу розпізнавання кіберфізичних об'єктів.

Об'єкт дослідження – кіберсоціальні/кіберфізичні процеси і явища, автоматично керовані комп'ютинговими сервісами на основі використання метрики подібності – відмінності. Предмет дослідження – регулярні матричні

структури, які використовуються для створення методів і алгоритмів паралельного розподілу та об'єднання векторів рядків і стовпців по метриці подібності-відмінності для синтезу архітектури кіберфізичного комп'ютингу розпізнавання.

Для досягнення поставленої мети необхідно вирішити такі задачі: проаналізувати сучасні технологічні тенденції; розширити моделі, методи, архітектури комп'ютингу розпізнавання кіберфізичних об'єктів. У процесі реалізації завдань дослідження виконано огляд новітніх програмно-апаратних рішень, орієнтованих на створення ефективних засобів, які дають можливість здійснювати пошук даних в кіберпросторі за прийнятний час.

Термін «інтернет поведінки» (IoB) аналітики від Gartner вперше запропонували в прогнозі ще на 2020 рік як технології для моніторингу поведінкових явищ і управління даними, які на них впливають. До них відносяться: розпізнавання осіб, відстеження місця розташування і великі дані. Gartner прогнозує, що до кінця 2025 року понад половини населення світу буде задіяно хоча б в одній програмі IoB [1]. Технологічним ядром для вирішення практичних завдань є кіберфізичний комп'ютинг, оформлений в структури Machine Learning. Такий комп'ютинг передбачає аналітику великих даних, яка використовує примітивні теоретико-множинні операції, процедури та алгоритми паралельної дії з метою підвищення продуктивності при пошуку квазіоптимальних рішень [2,3]. У середовищі комп'ютингу пошуку на анвансцену виходить метрика подібності-відмінності. Тому важливо мати ефективний спеціалізований процесор, як найпростіше ядро, для паралельного і високопродуктивного рішення задач синтезу і аналізу [2]. Структурно метрика подібності-відмінності використовує дві формули, які оперують двома паралельними операціями and, хог для отримання результуючих векторів: $S(a,b) = a_i \wedge b_i$; $D(a,b) = a_i \oplus b_i$. Важливо визначити загальні структури даних, щоб зрозуміти, як трансформуються окремі компоненти (координати векторів) один в другий при синтезі і аналізі. Процес синтезу знаходиться в об- числювальній

залежності від технологічно досконалих структур даних. Нормована метрика подібності-відмінності використовує дві формули, що також використовують в алгебрі логіки дві паралельні операції, але доповнені арифметикою підрахунку одиничних координат, отриманих в результаті виконання логічних операцій. З'являється також спільний знаменник у вигляді диз'юнкції однойменних координат векторів, який служить інтегратором розрізнених структур даних процесів в загальний вектор тільки істотних координат, відносно яких виконується нормування подібності-відмінності:

$$S(a, b) = \frac{\sum_{i=1}^n (a_i \wedge b_i)}{\sum_{i=1}^n (a_i \vee b_i)}; D(a, b) = \frac{\sum_{i=1}^n (a_i \oplus b_i)}{\sum_{i=1}^n (a_i \vee b_i)}.$$

Запропоновано метод і апаратна реалізація секвенсора для пошуку подібності-відмінності-включення, який характеризується отриманням структурованої оцінки взаємодії двох об'єктів, яка дає можливість більш точно визначати шляхи трансформування одного об'єкта в інший, а також вибирати більш суттєві об'єкти з пари при прийнятті рішення. Запропоновані методи можуть бути застосовані для пошуку одиночних і кратних дефектів в цифрових програмних і апаратних системах. Показано, що методи аналізу можна використовувати для матриць, заданих чисельними параметрами в довільній формі, а також для пошуку даних, пов'язаних з ідентифікацією процесів і явищ в кіберпросторі. Розроблені методи орієнтовані на апаратну реалізацію виконання паралельних логічних реєстрових операцій, які забезпечують істотне (на порядок) підвищення швидкодії в порівнянні з існуючими аналогами, виконаними в програмному коді. Обмеження запропонованих методів пошуку даних пов'язані з необхідністю виконання процедури синтезу матричних структур даних, яка має квадратичну обчислювальну складність.

Література:

1. <https://www.gartner.com/smarterwithgartner/5-trends-drive-the-gartner-hype-cycle-for-emerging-technologies-2020/>
2. Hahanov V. Cyber Physical Computing for IoT-driven Services. New

York. Springer. 2018. 279 p. [<https://www.springer.com/gp/book/9783319548241>]

3. V. Hahanov, S. Chumachenko, E. Litvinova, A. V. Hacimahmud, A. Hahanova and T. Soklakova, "Cyber Social Computing," 2018 IEEE East-West Design & Test Symposium (EWDTS), Kazan, 2018, pp. 1-8, doi: 10.1109/EWDTS.2018.8524663.



ДОДАТОК Б

Харківський національний університет радіоелектроніки
Факультет Комп'ютерної інженерії та управління
Кафедра Автоматизації проектування обчислювальної техніки

Магістерська атестаційна робота

Чугай Владислав

Моделі та методи розпізнавання кіберфізичних об'єктів на основі використання метрики подібності-відмінності

Науковий керівник – д.т.н., проф. Свірв І.Б.

Харків – 2020

ЗМІСТ

1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

(Технологічні тенденції 2021 від компанії Gartner, програмно-апаратні рішення)

2. МЕТРИКА ПОДІБНОСТІ - ВІДМІННОСТІ ДЛЯ КІБЕРФІЗИЧНОГО ПРОСТОРУ

3. МЕТОДИ ВИЗНАЧЕННЯ ПОДІБНОСТІ ДВОХ ОБ'ЄКТІВ

4. АРХИТЕКТУРА ТА СЕКВЕНСОП АПАРАТНОГО ОБЧИСЛЕННЯ ПОДІБНОСТІ- ВІДМІННОСТІ-ВКЛЮЧЕННЯ ОБ'ЄКТІВ ДЛЯ КОМП'ЮТИНГУ РОЗПІЗНАВАННЯ



МЕТА РОБОТИ

У магістерській роботі розглядаються питання, пов'язані зі створенням моделей, методів, елементів архітектури комп'ютерного розпізнавання кіберфізичних об'єктів на основі використання метрики подібності – відмінності.

Мета дослідження (глобально) – суттєве зменшення обчислювальної складності алгоритмів комп'ютерного розпізнавання кіберфізичних об'єктів шляхом розробки ефективної інфраструктури матричних даних (моделей, методів, архітектур) з використанням метрики подібності-відмінності.

Розробка архітектур кіберфізичного комп'ютерного розпізнавання на основі синтезу моделей і методів, а також метрики подібності-відмінності.

Об'єкт дослідження – кіберсоціальні/кіберфізичні процеси і явища, автоматично керовані комп'ютерними сервісами на основі використання метрики подібності – відмінності.

Предмет дослідження – матричні структури, які використовуються для створення методів і алгоритмів паралельного розподілу та об'єднання векторів рядків і стовпців по матриці подібності-відмінності для синтезу архітектури кіберфізичного комп'ютерного розпізнавання.



21.12.2020

3



ЗАДАЧІ РОБОТИ

Для досягнення поставленої мети необхідно вирішити такі **задачі**:

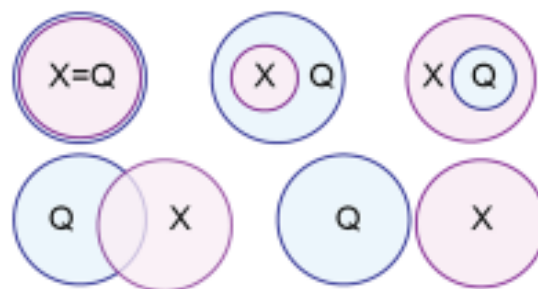
- проаналізувати сучасні технологічні тенденції та особливості побудови програмно-апаратних рішень;
- розширити моделі, методи, архітектури комп'ютерного розпізнавання кіберфізичних об'єктів.

Задачі дослідження орієнтовані на розвиток та удосконалення моделей, методів, технологій і архітектур комп'ютерного розпізнавання.

4

Комп'ютинг розпізнавання

- Пошук можна розглядати як вид комп'ютингу, призначений для визначення чи розпізнавання подібності/відмінності між процесами або явищами. Метод пошуку при цьому використовує бінарні властивості суспільного та приватного, подібності та відмінності для застосування бінарних апаратів – інтегрування і диференціювання, тестування і діагностування, множення і розподілу, додавання і віднімання, кон'юнкції і диз'юнкції. Метрика розглядається як спосіб визначення відстані між схожістю і відмінностями.
- Комп'ютинг розпізнавання визначається як пошук відстані (подібності та відмінності) між процесами або явищами. Подібність розглядається як ступінь спільності або перетину.
- Види теоретико-множинного перетину між двома множинами елементів:



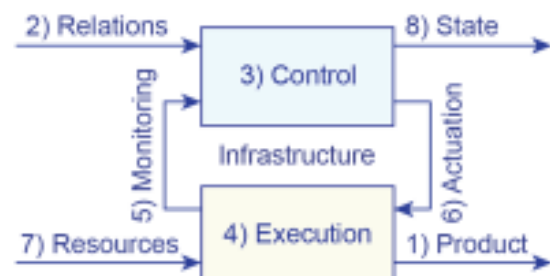
21.12.2020

7

Комп'ютинг критичних систем

- Комп'ютинг орієнтований на моніторинг і управління критичними об'єктами, процесами і явищами без- або з мінімальною участю людини.
- Оскільки сьогодні не можна поки обійтися без людини, як одного з компонентів управління, то необхідно мінімізувати його можливі помилки при прийнятті оперативних і стратегічних рішень у критичній системі. Виконання завдання пов'язане з цифровізацією історії, знань, умінь і навичок кожного співробітника на основі детермінованої метрики, попередньо сформованої експертами.

Метричне і структурне визначення комп'ютингу за допомогою восьми взаємопов'язаних компонентів надає теоретичну фундаментальну основу для формального і фактичного створення системи цифрового управління будь-яким процесом в заданій сфері діяльності.

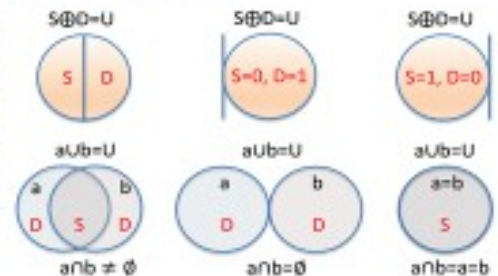


21.12.2020

8

Метрика подібності-відмінності

Розглядається комп'ютинг критичних систем. Дається універсальна метрика пошуку даних в кіберпросторі на основі використання параметрів подібності-відмінності і матричної структури в двійковій формі. Описуються методи аналізу матричних структур даних по метриці подібності-відмінності для пошуку дефектів в цифрових системах.



Метрика подібності-відмінності об'єктів процесів або явищ (рисунок вгорі), представлена рівнянням $S \oplus D = a \cup b = U$, є базисом для вирішення найпоширеніших на технологічному ринку завдань (рисунок зліва), пов'язаних з вимірюваннями в кіберфізичному просторі і інтернеті.

21.12.2020

9

Метрика і операції подібності-відмінності для пошуку дефектів при аналізі стовпців

- Теоретико-множинна і векторно-логічна (кубітна) взаємодія константних дефектів і їх похідних (подібності-відмінності) визначається операціями ($\cap, \cup, \bar{\Delta}, \Delta$) \approx ($\wedge, \vee, \bar{\oplus}, \oplus$), які представлені в таблицях

\cap	0	1	X	\emptyset	\wedge	10	01	11	00	
0	0	\emptyset	0	\emptyset	10	10	00	10	00	
1	\emptyset	1	1	\emptyset	\rightarrow	01	00	01	01	00
X	0	1	X	\emptyset	11	10	01	11	00	
\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	00	00	00	00	00	

a	0	1	X	\emptyset	a	10	01	11	00	
\bar{a}	1	0	\emptyset	X	\rightarrow	\bar{a}	01	10	00	11

Δ	0	1	X	\emptyset	\oplus	10	01	11	00	
0	\emptyset	X	1	0	10	00	11	01	10	
1	X	\emptyset	0	1	\rightarrow	01	11	00	10	01
X	1	0	\emptyset	X	11	01	10	00	11	
\emptyset	0	1	X	\emptyset	00	10	01	11	00	

\cup	0	1	X	\emptyset	\vee	10	01	11	00	
0	0	X	X	0	10	10	11	11	0	
1	X	1	X	1	\rightarrow	01	00	01	01	11
X	X	X	XX		11	11	11	11	11	
\emptyset	0	1	X	\emptyset	00	10	01	11	00	

21.12.2020

10

МЕТОДИ ВИЗНАЧЕННЯ ПОДІБНОСТІ ДВОХ ОБ'ЄКТІВ

- Розглядаються частотно-множинний метод визначення подібності двох об'єктів, а також частотно-векторна модель і метод для обчислення подібності.



ЧАСТОТНО-МНОЖИННИЙ МЕТОД

- Розглядаються визначення подібності двох об'єктів, а також частотно-векторна модель і метод для обчислення подібності.
- Кількість порожніх символів для вирівнювання двох слів дорівнює чотирьом. Після цього здійснюється тривіальний підрахунок відстані Левенштайна, яка дорівнює числу координат, мають різні символи в матриці трансформації слів, що означає $D(a, b) = 6$, $S(a, b) = 5$.

Приклад

CONDUCTION
BONDIANA

- Виконання алгоритму вставки порожніх символів з метою отримання мінімальної відмінності і максимальної схожості за умови трансформування одного слова в інше дає результат:

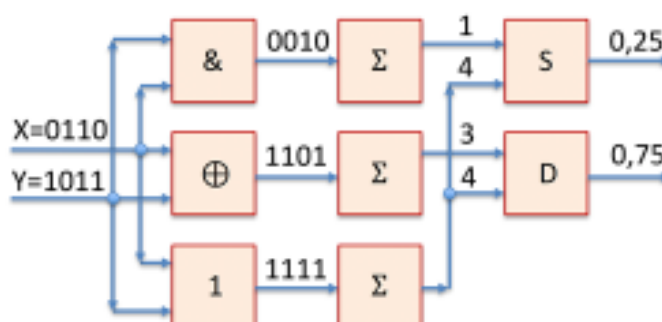
CONDUCTION -
BOND - - - IANA

СЕКВЕНСОР АПАРАТНОГО ОБЧИСЛЕННЯ ПОДІБНОСТІ-ВІДМІННОСТІ-ВКЛЮЧЕННЯ ОБ'ЄКТІВ

- Завдання полягає у синтезі логічної структури для обчислення метричної взаємодії між процесами і явищами, заданими в двійковому коді.
- Мають місце дві нормованих оцінки:

$$D^n = \frac{D}{N} = \frac{\sum_{i=1}^n (X_i \oplus Y_i)}{\sum_{i=1}^n (X_i \vee Y_i)}$$

$$S^n = \frac{S}{N} = \frac{\sum_{i=1}^n (X_i \wedge Y_i)}{\sum_{i=1}^n (X_i \vee Y_i)}$$



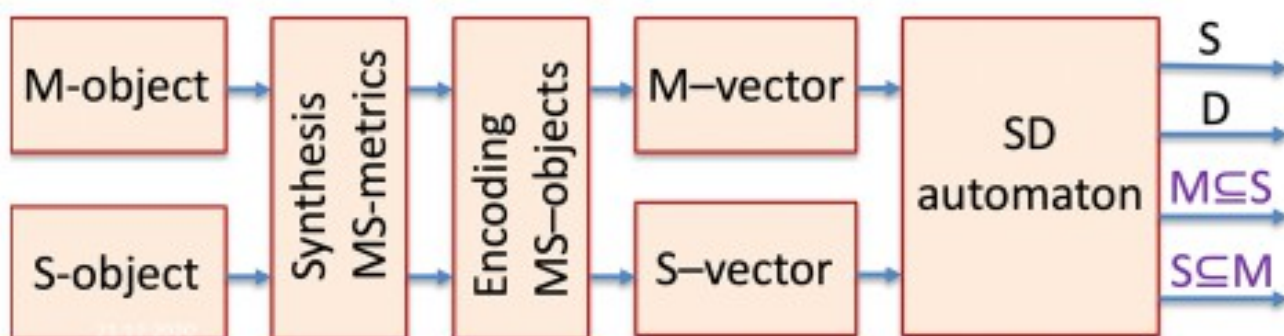
21.12.2020

13

СИНТЕЗ ЦИФРОВОГО SD-АВТОМАТА ДЛЯ ВИЗНАЧЕННЯ ПОДІБНОСТІ- ВІДМІННОСТІ ОБ'ЄКТІВ

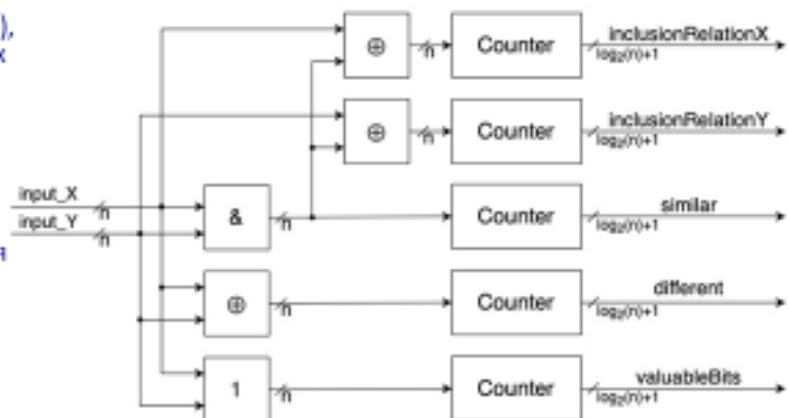
Структура секвенсора для визначення структурних оцінок подібності-відмінності-включення представлена на рисунку. Тут є 4 шари перетворення даних:

- 1) Формування вхідних даних про об'єкти.
- 2) Синтез загальної метрики для вимірювання об'єктів.
- 3) Унітарне кодування об'єктів в синтезованій метриці параметрів.
- 4) Подача векторів, що відповідають об'єктам, на входи SD-автомата.
- 5) Обчислення-формування чотирьох виходів-значень, що визначають відносини між об'єктами.



СТРУКТУРНА СХЕМА SD-АВТОМАТА

Пристрій виконує підрахунок кількості співпадаючих ненульових бітів (*similar*), що розрізняються (*different*), значущих координат (*valuableBits*), а також ступеня входження X і Y (*inclusionRelationX* і *inclusionRelationY* відповідно). Координати вважаються значущими, якщо хоча б в одному з вхідних векторів у біті по даній координаті зустрічається одиниця. Для більш точного підрахунку схожості / відмінності відбувається видалення чисто нульових однойменних координат X і Y - враховуються тільки ненульові.

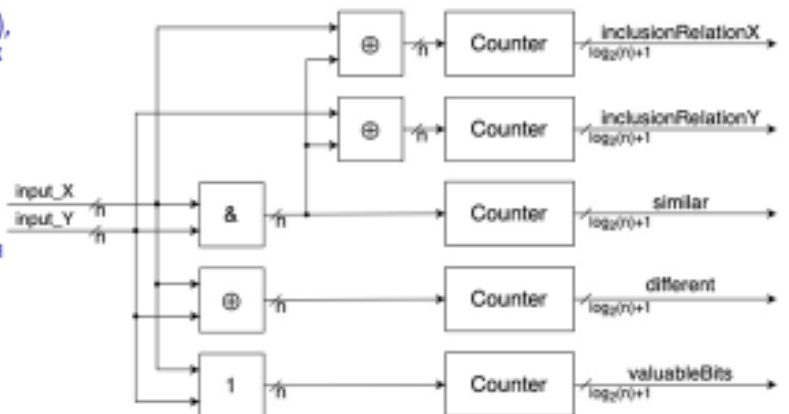


21.12.2020

15

СТРУКТУРНА СХЕМА SD-АВТОМАТА

Пристрій виконує підрахунок кількості співпадаючих ненульових бітів (*similar*), що розрізняються (*different*), значущих координат (*valuableBits*), а також ступеня входження X і Y (*inclusionRelationX* і *inclusionRelationY* відповідно). Координати вважаються значущими, якщо хоча б в одному з вхідних векторів у біті по даній координаті зустрічається одиниця. Для більш точного підрахунку схожості / відмінності відбувається видалення чисто нульових однойменних координат X і Y - враховуються тільки ненульові.



21.12.2020

15

ДЯКУЮ ЗА УВАГУ



