

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет комп'ютерної інженерії та управління
(повна назва)

Кафедра електронних обчислювальних машин
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

Рівень вищої освіти другий (магістерський)

Алгоритми обробки даних для підвищення
енергозбереження бездротових сенсорних мереж

(тема)

Виконав:

студент II курсу, групи СПМ-21-2
Гаптельманов А.В.
(прізвище, ініціали)

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

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

Освітня програма Системне програмування
(повна назва освітньої програми)

Керівник: проф. Міхаль О.П.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ЕОМ

(підпис)

Коваленко А.А.

(прізвище, ініціали)

2023 р.

Харківський національний університет радіоелектроніки

Факультет навчально-науковий центр заочної форми навчання

Кафедра електронних обчислювальних машин

Рівень вищої освіти другий (магістерський)

Спеціальність 123 «Комп'ютерна інженерія»
(код і повна назва)

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

Освітня програма Системне програмування
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

“ _____ ” _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студенту Гаптельманову Артему Володимировичу
(прізвище, ім'я, по батькові)

1. Тема роботи Алгоритми обробки даних для підвищення енергозбереження бездротових сенсорних мереж

затверджена наказом по університету від “ 03 ” квітня 2023 р. № 318 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 17 травня 2023 р.

3. Вхідні дані до роботи _____

машинне навчання

енергозбереження

класифікація

вузол

сенсорна мережа

адаптивний алгоритм

4. Перелік питань, що потрібно опрацювати у роботі _____

Методи та засоби обробки інформації в бездротових сенсорних мережах

Методи і моделі класифікації станів в бездротових сенсорних мережах

Програмна реалізація алгоритму та тестування

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) 16 слайдів

6. Консультанти розділів роботи (заповнюється 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	Термін виконання етапів роботи	Примітка
1	Отримання завдання. Аналіз предметної області	03.04.2023–10.04.2023	
2	Аналіз існуючих алгоритмів	11.04.2023–26.04.2023	
3	Розробка адаптивного алгоритму	27.04.2023–29.04.2023	
4	Визначення критеріїв для машинного навчання	30.04.2023–02.05.2023	
5	Отримання та аналіз результатів	03.05.2023–06.05.2023	
6	Оформлення пояснювальної записки	07.05.2023–13.05.2023	

Дата видачі завдання 03 квітня 2023 р.

Студент _____
(підпис)

Керівник роботи _____
(підпис)

проф. Міхаль О.П.
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 61 с., 18 рис., 1 дод., 17 джерел.

МАШИННЕ НАВЧАННЯ, СЕНСОРНА МЕРЕЖА, ДЕРЕВО РІШЕНЬ, АДАПТИВНИЙ АЛГОРИТМ, ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ.

Метою кваліфікаційної роботи аналіз алгоритмів обробки даних для підвищення енергозбереження бездротових сенсорних мереж з використанням методів машинного навчання.

У ході виконання кваліфікаційної роботи розроблено алгоритми роботи сенсора бездротової сенсорної мережі та побудовано бінарне дерево рішень, які можуть бути використані для підвищення енергозбереження функціонування вузлів в бездротових сенсорних мережах. Розглянуто методи машинного навчання. Для реалізації запропонованого алгоритму обрано метод дерева рішень. Вибір пов'язаний з тим, що даний метод має високу інтерпретацію та схожий із процесом прийняття рішення оператором. Грунтуючись на ідемпотентності предикатів дерева рішень, пропонується скоротити кількість сеансів зв'язку за рахунок динамічного визначення частоти передачі. Таким чином, цей підхід дозволить позбутися постійної передачі даних бездротовим каналом зв'язку, тим самим буде зекономлено енергоресурси вузлів мережі.

ABSTRACT

Master's thesis: 61 pages, 18 figures, 1 appendices, 17 sources.

MACHINE LEARNING, SENSOR NETWORK, DECISION TREE, ADAPTIVE ALGORITHM, SOFTWARE.

The major goal of this thesis is the analysis of data processing algorithms to increase the energy saving of wireless sensor networks using machine learning methods.

In order to algorithms for the operation of the sensor of a wireless sensor network were developed and a binary decision tree was built, which can be used to increase the energy saving of the functioning of nodes in wireless sensor networks. Methods of machine learning are considered. The decision tree method was chosen to implement the proposed algorithm. The choice is due to the fact that this method has a high interpretation and is similar to the process of decision-making by the operator. Based on the idempotency of the predicates of the decision tree, it is proposed to reduce the number of communication sessions due to the dynamic determination of the transmission frequency. Thus, this approach will make it possible to get rid of the constant transmission of data through a wireless communication channel, thereby saving energy resources of network nodes.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	7
ВСТУП	8
1 МЕТОДИ ТА ЗАСОБИ ОБРОБКИ ІНФОРМАЦІЇ В БСМ.....	12
1.1 Енергоживлення вузла БСМ	12
1.2 Аналіз енергоживлення вузла БСМ	14
1.3 Енергетичні втрати вузла БСМ. Класифікація.....	18
1.4 Аналіз існуючих підходів енергоефективної передачі інформації у бездротових сенсорних мережах	18
1.5 Класифікація функцій БСМ	21
1.6 Класифікація функцій БСМ	23
1.7 Машинне навчання	27
2 МЕТОДИ І МОДЕЛІ КЛАСИФІКАЦІЇ СТАНІВ В БСМ	33
2.1 Характеристики ознакового простору в БСМ.....	34
2.2 Модель завадання класифікації станів в БСМ	36
2.3 Алгоритм побудови бінарного дерева рішень	37
3 ПРОГРАМНА РЕАЛІЗАЦІЯ АЛГОРИТМУ ТА ТЕСТУВАННЯ	40
3.1 Вимоги до алгоритму.....	40
3.2 Структурно-функціональна модель алгоритму	42
3.3 Проведення експериментів	45
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	51
ДОДАТОК А Г Графічний матеріал кваліфікаційної роботи.....	53

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ
І ТЕРМІНІВ

БСМ – бездротова сенсорна мережа

МН – машинне навчання

ПЗ – програмне забезпечення

СТЗ – система технічного зору

ШНМ – штучна нейронна мережа

ВСТУП

Структура бездротової сенсорної мережі (БСМ) часто залежить від її цільового призначення та навколишнього середовища інфраструктури. У класичному варіанті БСМ є безліч взаємопов'язаних по бездротовій мережі вузлів (сенсорів), розташованих на певній території та виконуючих збір характеристик довкілля. Як правило, вузол мережі – це малогабаритний пристрій, до складу якого входять датчик, блок обробки даних (мікропроцесор і зовнішня пам'ять), приймач і автономне джерело живлення. Сенсор реєструє зміни конкретних параметрів довкілля (освітлення, температура, вологість, тиск, рух тощо), потім обробляє їх та передає на базову станцію. Вузол мережі також може виступати в ролі ретранслятора повідомлень, за рахунок чого досягається значна зона покриття мережею за низьких фінансових витрат.

БСМ використовують для виконання таких функцій, як автоматичний збір даних, стеження параметрами середовища [1]. Серед основних типів БСМ виділяють: наземні, підземні, підводні, мультимедійні та мобільні мережі. Їх відмінні риси полягають у різних топологіях, технічній складовій сенсора, періодичності та швидкості передачі даних, особливості каналу зв'язку. За рахунок низької вартості, швидкості розгортання та високої ефективності БСМ широко застосовуються практично у всіх сферах життєдіяльності, особливо популярні в оборонній галузі, промисловості, сільському господарстві та охороні здоров'я.

За великим потенціалом використання БСМ стоїть чимало наукових та технічних проблем, що починаються ще з проектування мережі [2]. Ці проблеми мають концептуально новий характер. Важливість їх рішення полягає в тому, що невраховані деталі можуть суттєво вплинути на функціонування всієї мережі у майбутньому. На відміну від інформаційних мереж, БСМ містять набір відмінних характеристик, які сприяють створенню

нових способів та методик проектування. В силу особливостей використання БСМ та обмеженого функціоналу сенсорів, що формують мережу, виникають нові параметри системи, такі як розгортання та конфігурація маршрутизації, зона покриття, енергоспоживання та термін автономної роботи вузла мережі [3]. Оскільки все частіше БСМ є важливим елементом у процесах зовнішніх систем, то необхідно приділяти велику увагу надійності та відмовостійкості мережі.

Проведений аналіз публікацій [4-7] показав актуальність питання енергозбереження в БСМ. Значна частина досліджень спрямована на рішення вузькоспеціалізованих проблем, що належать до конкретних областей застосування БСМ, і лише невелика частина досліджень сфокусована на загальних задачах, властивих різним типам БСМ [8, 9]. При проектуванні БСМ за заданими вимогами часто стикаються з безліччю обмежень, зумовлених зовнішнім середовищем, рівнем розвитку техніки, доступними ресурсами [10,11]. Серед розв'язуваних завдань виділяють наступні значні напрямки: проектування апаратного забезпечення, способи і алгоритми управління та обробки інформації, питання експлуатації.

Серед перерахованих вище завдань приділимо увагу розробці методів і алгоритмів управління та обробки інформації у мережі. Ключовими параметрами при оцінці запропонованих методів та алгоритмів комунікації в мережі є швидкість передачі даних та термін експлуатації БСМ.

Термін експлуатації мережі безпосередньо залежить від тривалості функціонування вузлів мережі. У свою чергу, вузол мережі може припинити функціонувати з причин внутрішньої несправності, агресивного зовнішнього середовища, виходу із зони дії бездротової мережі або вичерпання енергії джерела живлення. Вузол мережі споживає електроенергію для збору, обробки та передачі інформації, також при цьому повинна бути вирішена задача маршрутизації при топології, відмінної від топології «зірка». Як правило, вузол мережі містить автономне джерело живлення з лімітованим енергоресурсом та відсутні способи поповнення енергії із зовнішнього

середовища. Час автономної роботи вузла мережі пов'язано безпосередньо з часом служби автономного джерела живлення.

Незважаючи на те, що для вузлів БСМ можна передбачити механізми отримання енергоресурсів із навколишнього середовища (сонячні батареї, механічна, електромагнітна енергія), не завжди існує прямий і постійний доступ до перерахованих ресурсів [12]. Також варто враховувати, що елементи перетворення енергії з довкілля мають високу вартість, яка доходить до собівартості самого датчика; в такому випадку використання перетворювачів стає безглуздом, оскільки часто однією з базових вимог до БСМ є низька вартість датчиків. У дослідженні [13] було підкреслено, що питома енергетична щільність сучасних хімічних джерел живлення зростає повільніше, ніж обсяг споживаної енергії. Ще важливим фактом є те, що джерела енергії з великою енергоємністю становлять підвищену небезпеку і вимагають особливих умов експлуатації, тому що в іншому випадку можуть статися загоряння та вибухи.

Отже, вибір енергоефективних методів і алгоритмів збору та передачі інформації між вузлами мережі є однією з основних актуальних наукових проблем під час проектування БСМ, що підтверджує актуальність даної роботи.

Метою кваліфікаційної роботи є аналіз алгоритмів обробки даних для підвищення енергозбереження бездротових сенсорних мереж з використанням методів машинного навчання.

Завдання:

- аналіз енергоспоживання в бездротових сенсорних мережах з метою визначення залежності енергоспоживання від виконуваних сенсором функцій;
- аналіз методів машинного навчання;
- розробка алгоритму енергозбереження в бездротових сенсорних мережах;
- побудова структурно-функціональної моделі адаптивного

енергозберігаючого алгоритму;

- реалізація та тестування.

Ключова ідея БСМ – автоматизація збору інформації про навколишнє середовище та керованих об'єктах. БСМ особливо корисні там, де присутність людини в контрольованій зоні неможлива або збирання інформації необхідно проводити тривалий час. БСМ отримали широке поширення після активного розвитку сучасної мікроелектроніки, технологій бездротової взаємодії та відповідного апаратного забезпечення. Енергозбереження даного типу мереж є важливим та актуальним.

1 МЕТОДИ ТА ЗАСОБИ ОБРОБКИ ІНФОРМАЦІЇ В БСМ

Сенсорні мережі дозволяють автоматизувати збирання даних про навколишнє середовищі чи контрольованому об'єкті. Бездротова взаємодія між вузлами мережі значно розширює сферу застосування, але несе безліч нових задач, якими представлено ряд досліджень.

Одним з важливих завдань при проектування БСМ є збільшення терміну експлуатації мережі, який, своєю чергою, пов'язані з терміном експлуатації вузлів мережі. Функціонування сенсора може бути порушено з кількох причин, і в цій роботі розглядаються проблеми обмеженого енергоресурсу біля вузла мережі та неенергоефективних алгоритмів обробки даних.

1.1 Енергоживлення вузла БСМ

Аналіз споживання електроенергії вузлом БСМ ґрунтується на експериментальних даних, оскільки дослідження теоретичної моделі енергоспоживання вузла БСМ досить трудомістке і має низьку точність через наявність безлічі факторів, таких як: тип сенсора, дальність передачі даних, що виконуються обчислення. Результатом аналізу енергоспоживання є залежність витраченої енергії від виконуваних сенсором функцій. Перед аналізом енергоспоживання вузла мережі розглянемо його узагальнену структуру у складі БСМ. До складу вузла мережі (сенсора) входять (рисунок 1.1):

- а) датчики, що вимірюють параметри навколишнього середовища;
- б) блок обробки даних, що складається з мікроконтролера та зовнішньої пам'яті;
- в) приймач;
- г) автономне джерело живлення.

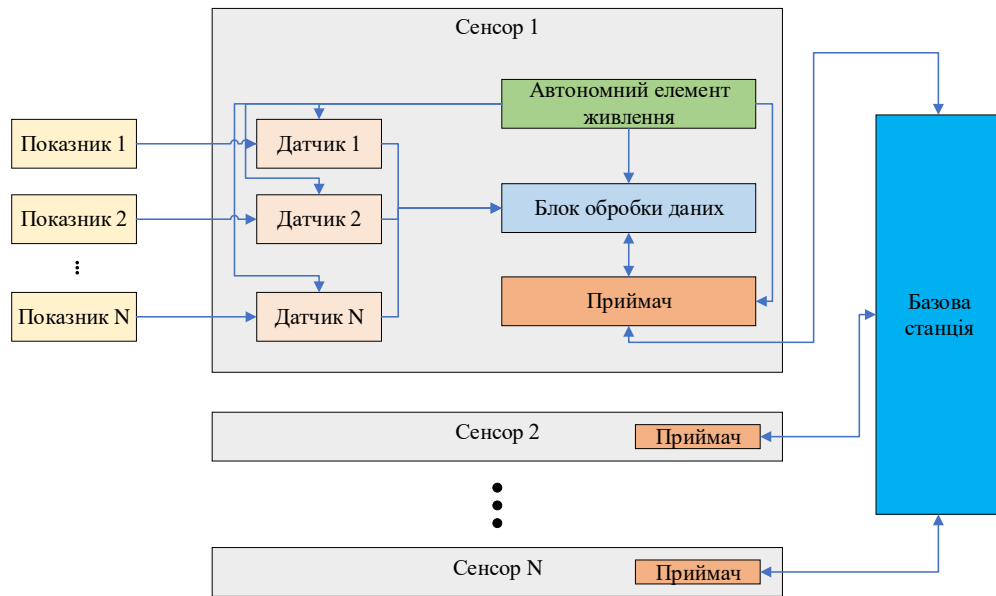


Рисунок 1.1 – Узагальнена структура вузла у складі БСМ

Вузол мережі вимірює параметри навколишнього середовища, обробляє їх, а потім посилає базову станцію, куди надходять дані з усіх сенсорів.

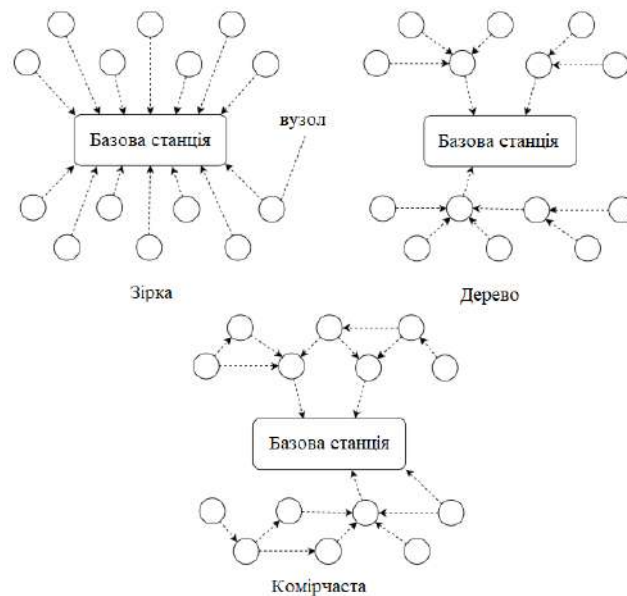


Рисунок 1.2 – Топології БСМ

З представленої структури видно, що енергія автономного елемента живлення витрачається на енергопостачання датчиків, мікроконтролера з пам'яттю та на приймач.

Ця структура відповідає топології мережі «зірка», коли кожен сенсор взаємодіє з базовою станцією безпосередньо. Якщо розглядати топологію «дерево» або «комірчасту», то пропадає частина зв'язків «Сенсор» - "Базова станція" і з'являються нові зв'язки "Сенсор" - "Сенсор". В БСМ топології «дерево» та «коміркова» використовуються для того, щоб збільшити зону покриття мережі. Це досягається за рахунок методів маршрутизації даних від віддалених сенсорів через проміжні вузли мережі.

У БСМ топології «дерево» та «коміркова» використовуються для того, щоб збільшити зону покриття мережі. Це досягається за рахунок методів маршрутизації даних від віддалених сенсорів через проміжні вузли мережі. Відмінність у енергоспоживанні залежно від топології БСМ полягає в тому, що при топологіях «дерево» та «коміркова» на частину вузлів мережі, які ближче до базової станції, буде надходити додаткове навантаження через передачу інформації в якості маршрутизації. В результаті при однаковій кількості даних у БСМ при топологіях «дерево» та «коміркова» енергоспоживання буде більше, ніж при топології «зірка» через додаткові сеанси зв'язку.

1.2 Аналіз енергоживлення вузла БСМ

На рисунку 1.3 представлений графік енергоспоживання вузла БСМ на одному робочому циклі, коли сенсор, що знаходиться в сплячому режимі, по команді від таймер прокидається, щоб виконати передачу даних по мережі, і знову повертається у сплячий режим. Етапи функціонування датчика пронумеровано від 0 до 10. Інтервалами роботи мікроконтролера є 1-3, 4-5, 6-7, 8-9, інтервалами передачі даних по мережі є: 3-4, 5-6, 7-8.

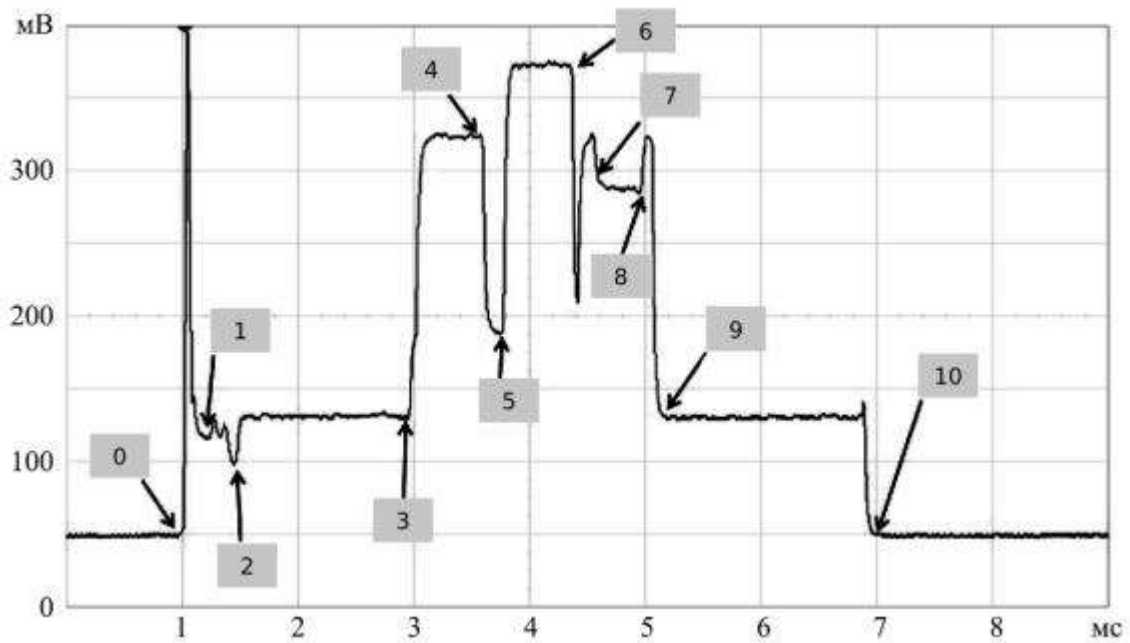


Рисунок 1.3 – Топології БСМ

Проаналізуємо графік енергоспоживання вузла БСМ у розрізі виконуваної сенсором функції та відповідних показань сили струму та напруги:

- проміжок 0-1: ініціалізація роботи приймача; сила струму 12 mA у межах 0.2 ms (спостерігається різкий казок енергоспоживання);
- проміжок 1-2: ініціалізація мікроконтролера з частотою 16 MHz за допомогою периферійного осцилятора; 6 mA у межах 0.25 ms;
- проміжок 2-3: перехід мікроконтролера на частоту 32 MHz периферійного осцилятора; 7.5 mA не більше 1.7 ms;
- проміжок 3-4: приймач працює в режимі читання каналу зв'язки для виявлення команди від базової станції; 27 mA межах 1.2 ms; - проміжок 4-5: перемикання приймача в режим відправлення даних для зв'язку з базовою станцією; 14 mA у межах 0.2 ms;
- проміжок 5-6: відправлення даних по мережі, сила струму – 32 mA в межах 0.5 ms;
- проміжок 6-7: повернення приймача в режим читання; 25 mA у межах

0.2 ms;

- проміжок 7-8: приймач працює в режимі читання каналу зв'язку; 23 mA у межах 0.35 ms;

- проміжок 8-9: завершення роботи приймача;

- проміжок 9-10: перехід сенсора в режим очікування; сила струму – 1 mA не більше 1 s.

Аналіз значень енергоспоживання показав, що близько 70% енергії автономного джерела живлення витрачається на комунікацію бездротовою мережі, іншими словами, комунікація вимагає вдвічі більше електроенергії, ніж при зборі та обробці інформації. На рисунку 1.4 представлений приклад осцилограми енергоспоживання вузла мережі у разі безперервного функціонування, коли сенсору необхідно обробити ланцюжок пакетів даних, що надходить.

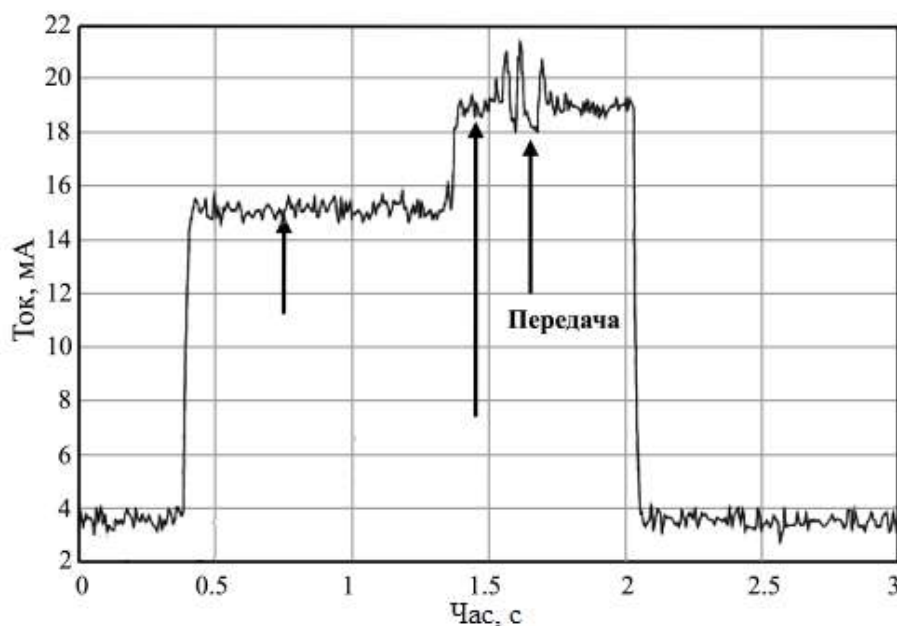


Рисунок 1.4 – Осцилограма роботи сенсора в залежності від виконуваних ним функцій

З рисунку 1.4 можна відзначити, що енергоспоживання сенсора по відношенню до виконуваних ним функцій є змінним. При зборі та обробці

даних витрата досягає близько 15 мА, а при читанні бездротового каналу зв'язку витрата збільшується до 20 мА, при надсиланні даних – до 22 мА. Зазначені величини можуть відрізнятись при різних технічних реалізаціях сенсор. Сенсор мережі складається з п'яти структурних частин, три з яких (мікроконтролер, пам'ять та приймач) можуть бути виконані на одному кристалі, що позначається як на зменшенні розмірів самого сенсора, так і на підвищенні його енергоефективності.

Запропонована методика розрахунку енергоспоживання є приблизною, так як досить складно обчислити дійсні абсолютні значення параметрів розрахунку через технічні відмінності конкретній реалізації вузлів БСМ. Дана методика дозволяє оцінити базові фактори, які впливають на енергоспоживання вузла БСМ. На основі аналізу представлених циклів роботи сенсора було розроблено Приклад алгоритму роботи сенсора (рисунок 1.5).

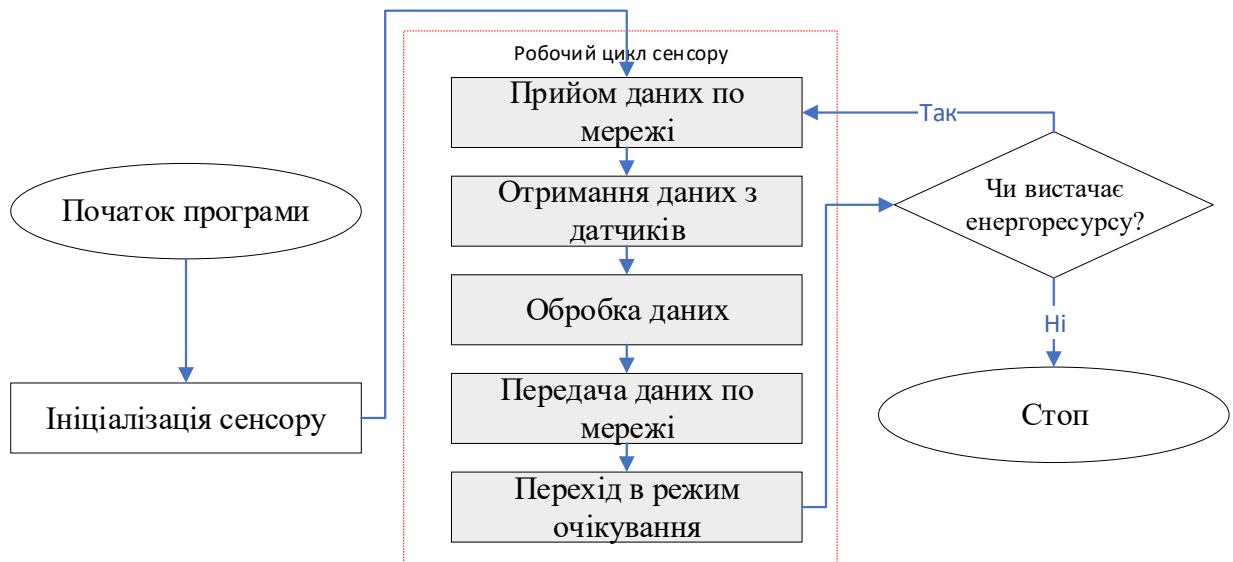


Рисунок 1.5 – Алгоритм роботи сенсора БСМ

Як правило, сенсор мережі функціонує циклічно, доки не закінчиться енергоресурс, тому представлений алгоритм не має завершення.

1.3 Енергетичні втрати вузла БСМ. Класифікація

Енергетичними втратами називаються енерговитрати, які не виконують безпосередню корисну роботу та їх прагнуть ліквідувати. Основними причинами енергетичних втрат у БСМ є:

- колізії, - надмірне прослуховування;
- прослуховування в очікуванні;
- службова інформація в пакетах, що передаються.

При одночасному передачі даних різними вузлами мережі відбувається колізія пакетів: сигнали змішуються фізично і надалі стороні, що приймає, неможливо розрізнити пакети. У такому разі вузли мережі намагаються повторно надіслати дані, тим самим витрачають додаткову електроенергію. У випадку, коли сенсор при відправленні пакета з даними одночасно прослуховує канал зв'язку - називається надлишковим прослуховуємо, тому що в цей момент часу в каналі зв'язку буде лише власний пакет та прослуховування даного пакета марно.

Прослуховування в очікуванні відбувається, коли вузол мережі безперервно прослуховує канал бездротового зв'язку в очікуванні можливого повідомлення, хоча воно може бути відсутнім тривалий проміжок часу. Пакет з даними, що передається, містить не тільки показання параметрів контрольованого об'єкта, але ще й набір додаткової інформації, яка необхідна для того, щоб ідентифікувати відправника та одержувача, перевірити коректність даних. В результаті фінальний розмір переданого пакета збільшується, що також негативно позначається на часі передачі по мережі.

1.4 Аналіз існуючих підходів енергоефективної передачі інформації у бездротових сенсорних мережах

Проведемо аналіз методів та засобів, що забезпечують збільшення часу автономної роботи вузла мережі за рахунок використання ефективних

методів та алгоритмів передачі. J

ohn Heidemann для зниження кількості передач даних у мережі пропонує метод GAF (Geographical Adaptive Fidelity), який всю область БСМ ділить на квадратні зони, у межах яких одночасно дані передає лише один вузол, інші ж перебувають у сплячому режимі. Даний метод частково вирішує проблему колізій та надмірного прослуховування. Результати дослідження показали, що запропонований метод дозволяє заощаджувати від 40% до 60% енергоресурсів усієї мережі. Недоліком цього методу є те, що кожен сенсор повинен мати механізм уточнення власної географічної позиції для визначення своєї приналежності до однієї із зон, навіщо, зазвичай, використовують модуль GPS (Global Positioning System), що істотно позначається на ціні сенсора та всієї мережі в цілому.

Енергетичний баланс при побудові маршруту передачі розглянуто у роботі Yunfeng Chen. Автор пропонує новий підхід EBMR при маршрутизації, в основі якого лежить принцип передачі даних різними маршрутами для того, щоб максимально розподілити енергоспоживання між вузлами мережі. Таким чином, автор вирішує проблему, коли вузли мережі, найбільш близькі до базової станції, суттєву частину електроенергії витрачають на ретрансляцію повідомлень. Дослідження методу EBMR показало, що з його використанні можна економити до 50% електроенергії. Дискретний режим роботи вузла мережі запропонований Терентьевим М. Н.

Суть дискретного режиму полягає в тому, щоб сенсор більшу частину часу проводив у режимі «сну», коли споживання електроенергії мінімально. За розкладом сенсор переходить в активний режим, щоб виконати вимірювання та надіслати їх бездротовою мережею. При цьому велика увага була приділена синхронізації сенсорів, щоб вузли мережі перебували в активному режимі одночасно, інакше можлива ситуація, коли дані передаються, але їх ніхто не приймає.

У роботі Бакіна Є. А. досліджується проблема колізій під час передачі даних, коли кілька сенсорів у каналі зв'язку створюють один для одного

перешкоди, що призводять до втрати повідомлень, внаслідок чого сенсори витрачають більше енергії на передачу даних та з'являються тимчасові затримки при доставці повідомлень. Цю проблему автор пропонує вирішувати за рахунок розробки розкладу передачі даних на основі інформації про періоди посилок кожного сенсор. Це рішення дозволяє боротися з колізіями в мережі.

У своїй роботі Бакін Є.А. проводить дослідження тривалості періоду збору інформації в різних топологіях та на основі аналізу отриманих даних пропонує алгоритм зі складання безконфліктного розкладу передачі даних, що забезпечує мінімальний період збирання інформації. Проблема низької тривалості роботи сенсора від автономного джерела живлення у випадку, коли він знаходиться на великій відстані від приймача сигналу (базової станції), досліджена у роботі Фоміна А. Д. У даній ситуації найбільший ресурс батареї витрачається на передачу даних, так чим більше відстань до приймача даних, тим більше енергії необхідно передачі. Фомін А.Д. пропонує реалізувати агрегування інформації в БСМ, що в результаті має сприяти зростанню енергоефективності системи.

Ця технологія корисна в тому випадку, коли для вирішуваного завдання достатньо визначити агреговану характеристику параметрів довкілля для певної ділянки мережі. В такому випадку один із вузлів ділянки вибирається «агрегатором». Агрегатор отримує дані обчислюваної характеристики з найближчих сенсорів, розраховує агрегатну функцію (min, max і т.д.) та надсилає результат на головний модуль. В такому у випадку сумарне енергоспоживання при надсиланні даних значно менше, ніж у разі відсутності агрегатора.

Лінським Є. М. запропоновано методику оцінки якості передачі даних за параметрів: надійність (ймовірність успішної передачі даних), швидкодія (затримка доставки); енергоефективність (енерговитрати). У роботі представлений алгоритм допоміжної відправки даних та пошук оптимальної конфігурації запропонованого алгоритму для різних типів повідомлень.

Способи конфігурування потоків інформації в БСМ галузевого застосування, що вирішують проблему, що виникає при динамічній зміні складу мережі та стану її сенсорів, представлені у дослідженні Іванової І. А.

Для вирішення зазначених проблем при конфігуруванні та експлуатації мережі представлений спосіб, який утворює інформаційні канали зв'язку між елементами системи для підвищення енергозбереження та актуальності поданої інформації про об'єкти мережі. Для цього була спроектована схема передачі інформації в рамках БСМ з великим набором сенсорів за умов динамічної модифікації складу мережі. Завдання підвищення якості функціонування БСМ та питання енергетичного балансу мережі представлено у науковому дослідженні Комарова М. М.

Під енергетичним балансом мережі розуміється забезпечення рівномірного споживання енергоресурсів усієї мережі для вирівнювання терміну автономної роботи її вузлів. Для досягнення поставленої мети побудовано модель та алгоритм функціонування додатково вузла мережі «диспетчера якості», який займається безпосередньо контролем енергетичного балансу мережі та за необхідності перенаправляє потоки даних, враховуючи вплив зовнішніх факторів та енергоспоживання, яке необхідне підтвердження доставки інформації від детектора.

1.5 Класифікація функцій БСМ

На кожен параметр БСМ істотно впливають методи і алгоритми управління та обробки інформації. У свою чергу, алгоритми залежать від необхідного функціоналу до БСМ. Розуміння завдань, які вирішуються за допомогою БСМ, дозволить формалізувати функціональні вимоги до алгоритму, що розробляється. Виділимо основні функції БСМ з урахуванням аналізу існуючих застосувань. Як було зазначено раніше у вступі, БСМ використовуються у безлічі областей: оборонної галузі, промисловості, сільському господарстві та медицині. Розглянемо по кілька прикладів

впровадження БСМ із кожної області, приділяючи особливу увагу виконуваних функцій. Одним із перших впроваджень БСМ можна вважати звукову систему спостереження SOSUS, яка з 1950-х років використовується ВМФ США для виявлення та ідентифікації підводних човнів.

Як датчики застосовуються гідрофони, які можуть виявити підводні човни 1-го та 2-го покоління на відстані до кількох тисяч кілометрів. З 2007 року усі берегові пости системи SOSUS працюють в автоматичному режимі, тобто не вимагають несення цілодобової вахти операторами. З цього факту можна зробити два висновки: по-перше, спочатку система з якихось причин контролювалася в ручному режимі досить тривалий час, по-друге, перехід на автоматичний режим означає, що система використовує алгоритми детектування та ідентифікації підводних човнів за даними з сенсорів. На жаль, більш докладні дані про алгоритми та в цілому про систему закриті та є секретною інформацією ВМФ США.

В оборонній галузі також згадується про тактичні БСМ, які допомагають приймати рішення при воєнних діях. Основним завданням БСМ є моніторинг ворожої території та стеження за пересуванням військових сил противника. Для цього за допомогою безпілотного літального апарату на цільову область скидають безліч сенсорів, які починають передавати дані до командного пункту. Важливими завданнями для тактичних БСМ є самоорганізація, надійність, швидкість передачі даних та захист від несанкціонованого доступу. Серед БСМ виділяють категорію акустичних мереж, що у складі сенсори містять чутливі елементи до звукових хвиль. Одним з застосування подібних мереж є система виявлення снайперів. Контролюючи гучність пострілу, за даними з кількох сенсорів встановлюється орієнтовне розташування пострілу. Для даної БСМ розробляються спеціальні алгоритми локалізації контрольованого об'єкту.

Також сенсорні мережі активно застосовуються при дослідженнях природних явищ. Наприклад, співробітники Гарвардського університету впровадили БСМ для моніторингу вулканічних активностей: вивержень,

землетрусів та поштовхів. У рамках цієї роботи було поставлено завдання детектування досліджуваних подій у режимі реального часу, щоб передавати дані по мережі тільки при настанні цікавих подій. За рахунок цього було збільшено час автономної роботи сенсорів, що було важливим завданням, так як заміна джерел живлення чи сенсорів загалом була дорогою операцією. Також була створена БСМ ZebraNet, яка допомагає зоологам контролювати міграцію зебр, вивчати їхню поведінку та реакцію на зовнішні впливи.

1.6 Класифікація функцій БСМ

Більшість досліджень у галузі підвищення енергоефективності вузла БСМ мають характерну рису: запропоноване рішення є результатом пошуку компромісу між підвищенням енергоефективності мережі та зниженням якості інших параметрів мережі, таких як: своєчасність доставки повідомлень, простота реалізації комунікації та обробки інформації. Таким чином, досі залишається відкритим питання побудови алгоритмів передачі даних та обробки інформації в БСМ, що забезпечують одночасне підвищення енергозбереження, продуктивності та адаптивності мережі до довкілля.

Розглянемо п'ять основних напрямів функціоналу БСМ (рисунок 1.6).

Збір даних – на зовнішній запит від базової станції сенсори зчитують показання з датчиків про параметри зовнішнього середовища або контрольованого об'єкта та відправляють на базову станцію. Моніторинг являє собою збір даних із певної періодичністю. Чергова ітерація збору даних може бути ініційована не тільки на запит від базової станції, а й за заданим розкладом, який зберігається локально у пам'яті сенсора.

Як правило, на основі моніторингу в наступним приймається рішення про стан мережі або події подіях. Під функцією контролю мають на увазі моніторинг параметрів навколишнього середовища та подальшу їх перевірку на спрацювання заздалегідь певних умов. Спостереження – передача показань на головний модуль, якщо сенсор виявив конкретний об'єкт чи

поведінку за заданим шаблоном. Формування керуючих команд – на основі отриманих із сенсорів показань приймається рішення про те, яку керуючу команду необхідно сформувавши для керованої системи. Для функцій моніторингу, контролю, стеження та формування керуючих команд потрібно на основі даних з вузлів БСМ приймати рішення про стан мережі або необхідну керуючу команду.

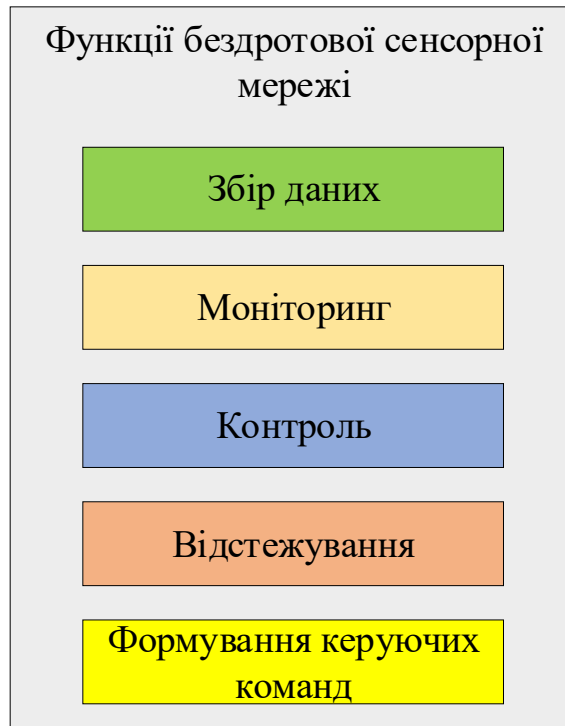


Рисунок 1.6 – Функції бездротової сенсорної мережі

Зазначені методи прийняття рішення мають свої недоліки: під час прийняття рішення оператором потрібні людські ресурси, а також з'являється фактор помилкового ухвалення рішення; при використанні алгоритму ми стикаємося з ситуацією, коли алгоритми фіксуються на конкретних параметрах і при зміні структури мережі необхідна їхня модифікація.

База знань БСМ – інформація, отримана раніше під час експлуатації мережі, містить структуровані дані з вузлів мережі та результати класифікації з допомогою експертної оцінки. Користь бази знань полягає в тому, що вона

дозволяє приймати рішення значно швидше через можливу наявність історичних даних, які містять схожі вхідні дані та результат їх класифікації. З представленого алгоритму видно, що у кожному циклі прийняття рішення є кроком збору даних із сенсорів. З точки зору енергоспоживання ця операція ресурсозатратна, оскільки містить велику кількість сеансів зв'язку бездротової мережі: як мінімум n передач с n сенсорів при топології мережі «зірка», при топології «дерево» та «комірчаста» кількість передач збільшується з допомогою маршрутизації.

У разі, якщо набір вхідних даних на поточному циклі ухвалення рішення немає подібних значень з бази знань, необхідно виконати етапи дослідження системи, проведення експертної оцінки та перевірки рішення. Після прийняття рішення вноситься до бази знань, щоб за наступних подібних вхідних даних не повторювати попередні операції. Щоб оцінити енергоспоживання мережі, необхідно позначити параметр – частота ухвалення рішення, яка є причиною частоти збору даних із сенсорів БСМ. Зміна даного параметра покращує одні властивості системи, а інші – погіршує, у результаті можна сформулювати технічне протиріччя.

Для зниження рівня споживаних енергоресурсів дослідники вдаються до різних методів, таких як запровадження розкладу передачі даних, зменшення обсягу переданих даних та ін. Таким чином, тенденція нових розробок іде до того, щоб зберегти енергоресурси за рахунок зменшення кількості сеансів зв'язку між вузлами мережі. З іншого боку, завдання БСМ – своєчасно збирати дані про параметри навколишнього середовища та на їх основі приймати наступні рішення.

У результаті, рішення, спрямоване зменшення кількості сеансів зв'язку, створює технічну суперечність: «Зі збільшенням числа передач даних збільшується енергоспоживання вузла мережі, що призводить до зменшення часу автономного термін служби сенсора, зі зменшенням кількості передач даних – збільшується ймовірність пропуску важливих подій, що може призвести до невірному ухваленню рішення»

Якщо відмовитися від періодичної передачі даних, потрібно забезпечити іншу методику взаємодії. Для цього скористаємося «принципом динамічності», тобто змінюватимемо характеристику системи так, щоб вона була оптимальною кожному етапі роботи системи.

В даному випадку принцип динамічності можна характеризувати як здійснення передачі даних лише тоді, коли у цьому є реальна потреба. Таким чином, для вирішення технічної суперечності пропонується розробити алгоритм обробки інформації у бездротових сенсорних мережах, який з одного боку скоротить кількість сеансів зв'язку, тим самим підвищить енергоефективність вузла БСМ, а з іншого боку гарантуватиме, щоб головний модуль завжди володів необхідною інформацією про параметри довкілля для коректного прийняття рішення.

Для динамічного визначення частоти передачі даних пропонується використовувати інформацію з бази знань, щоб сенсори самостійно ухвалювали рішення, коли необхідно передавати дані на базову станцію для загальної класифікації системи. Для реалізації цього алгоритму Придатними є методи машинного навчання. Дані методи спрямовані на створення математичної моделі процесу прийняття рішення на основі знань, отриманих з навчальної вибірки.

Реалізація процесу ухвалення рішення про необхідність передачі даних на головний модуль має розташовуватися в логіці роботи вузла мережі, оскільки першоджерелом інформації є дані з датчиків, що знаходяться безпосередньо у складі вузла мережі. З іншого боку, для ухвалення рішення про передачі даних можливо знадобляться знання про процес класифікації, то є дані про математичну модель прийняття рішення про стан мережі

Тенденція нових розробок іде до того, щоб зберегти енергоресурси за рахунок зменшення кількості сеансів зв'язку між вузлами мережі. З іншого боку, завдання БСМ – своєчасно збирати дані про параметри навколишнього середовища та на їх основі приймати наступні рішення. У результаті, рішення, спрямоване на зменшення кількості сеансів зв'язку, створює

технічну суперечність: «Зі збільшенням числа передач даних збільшується енергоспоживання вузла мережі, що призводить до зменшення часу автономного термін служби сенсора, зі зменшенням кількості передач даних – збільшується ймовірність пропуску важливих подій, що може призвести до невірному ухвалення рішення».

Для вирішення даної технічної суперечності пропонується розробити алгоритм обробки інформації у БСМ, який з одного боку скоротить кількість сеансів зв'язку, тим самим підвищить енергоефективність вузла БСМ, а з іншого боку гарантуватиме, щоб головний модуль завжди володів необхідною інформацією про параметри довкілля для коректного прийняття рішення.

Для динамічного визначення частоти передачі даних пропонується використовувати інформацію з бази знань, щоб сенсори самостійно ухвалювали рішення, коли необхідно передавати дані на базову станцію для загальної класифікації системи. Для реалізації цього алгоритму придатними є методи машинного навчання [14]. Дані методи спрямовані на створення математичної моделі процесу прийняття рішення на основі знань, отриманих з навчальної вибірки. Реалізація процесу ухвалення рішення про необхідність передачі даних на головний модуль має розташовуватися в логіці роботи вузла мережі, оскільки першоджерелом інформації є дані з датчиків, що знаходяться безпосередньо у складі вузла мережі. З іншого боку, для ухвалення рішення про передачу даних, можливо, знадобляться знання про процес класифікації, тобто дані про математичну модель прийняття рішення про стан мережі.

1.7 Машинне навчання

Аналіз існуючих застосувань [15,16] з використанням методів машинного навчання у роботі БСМ виявив, що часто такий підхід пов'язаний з тим, що ручне створення моделі ухвалення рішення є складним процесом,

іноді навіть нездійсненним за рахунок різноманіття зовнішніх факторів, що впливають на результат, а також їх непередбачуваність у майбутньому. Отже, створення методів прийняття рішення в БСМ є трудомісткою операцією, вимагає адаптивної настройки параметрів алгоритму для виконуваного завдання. Можлива ситуація, коли модель має надто багато вхідних даних, в результаті процес прийняття рішення може бути занадто довгим і задовольняти поставленим вимогам. Тому важливо генерувати модель, яка з одного боку надаватиме коректні результати прийняття рішення, а з іншого – має бути не надто складним, щоб час прийняття рішення було у допустимому інтервалі. У машинному навчанні для спрощення моделі використовують ймовірні підходи, які дозволяють знайти компроміс між складністю моделі та точністю прийнятих рішень. Тому найчастіше під час використання методів машинного навчання виконують попередній етап аналізу вхідних даних.

Серед основних завдань машинного навчання (рисунок 1.7) виділяють завдання регресії, класифікації, кластеризації, пошуку аномалій. Для кожного завдання використовуються різні методи, ефективність яких залежить від типу вхідних даних та бажаного результату.

Представлені методи дозволяють генерувати модель прийняття рішення шляхом навчання на раніше відомій вибірці даних та результатах їх обробки. У випадку з БСМ вхідними даними будуть параметри зовнішнього середовища, отримані з сенсорів, та результати ухвалення рішення про стан мережі. Отже, у цій роботі для завдання ухвалення рішення пропонується використати методи машинного навчання.

Для цього на основі раніше відомих прецедентів (бази знань) буде отримано математичну модель, за допомогою якої надалі автоматично прийматиметься рішення, імітуючи роботу оператора. Завдання визначення стану в сенсорній мережі за даними із сенсорів відноситься до задачі класифікації.

Найбільш популярними методами класифікації на даний момент є:

а) k-найближчих сусідів;

- б) наївний байєсівський класифікатор;
- в) метод опорних векторів;
- г) лінійна регресія;
- г) індукція правил (дерево рішень, випадковий ліс).

Вибір методу класифікації залежить від показників вхідних даних (кількість параметрів, пропуски в даних), точності та повноти класифікації, часу навчання та перенавчання.

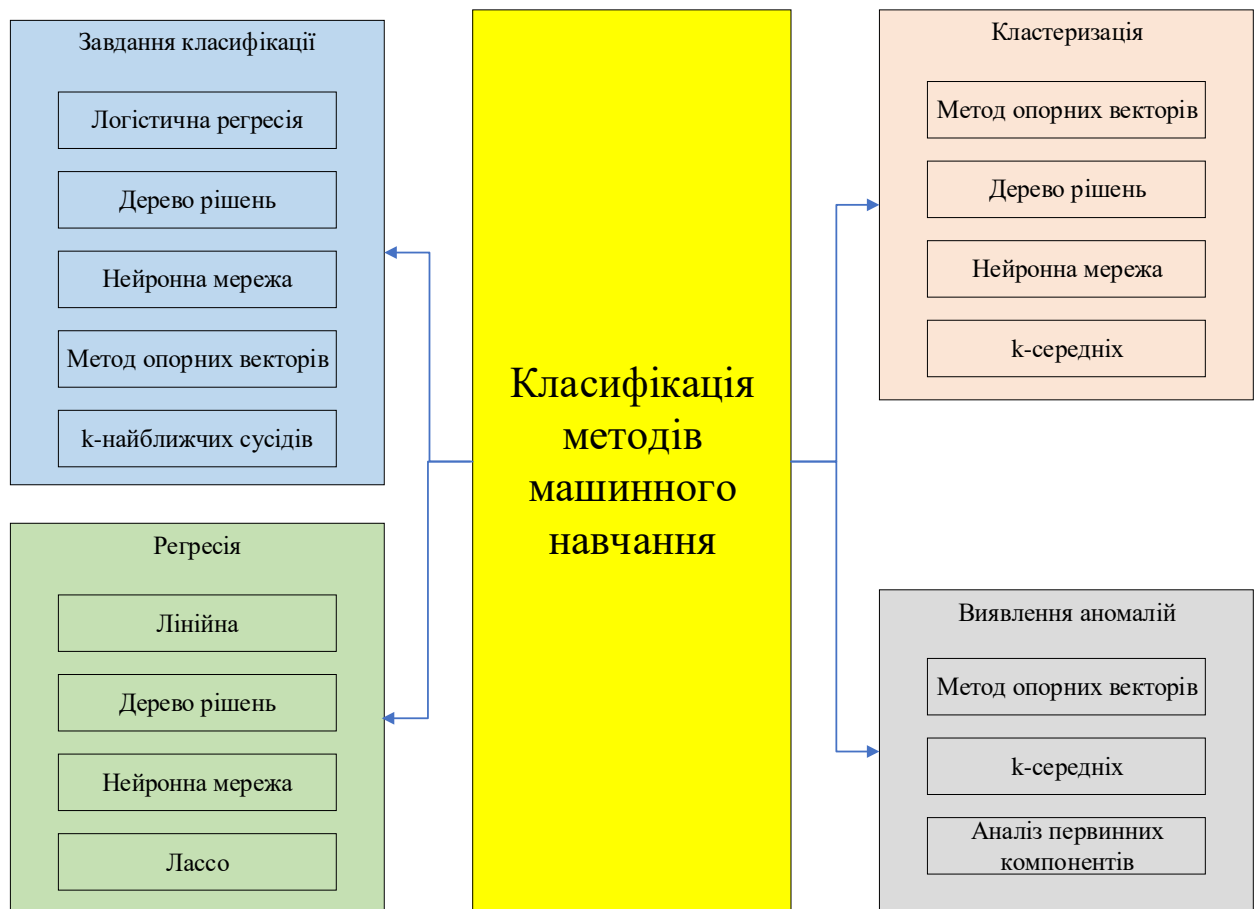


Рисунок 1.7 – Узагальнена класифікація методів машинного навчання

Велика кількість досліджень із завдання класифікації об'єктів представлено у галузі машинного навчання, штучного інтелекту. Проведено аналіз методів машинного навчання з метою пошуку найбільш відповідної реалізації запропонованого алгоритму. Для ухвалення рішення про стан об'єкта використовуються різні моделі класифікації об'єктів, які широко

представлені в галузі машинного навчання. Безліч розроблених методів класифікації об'єктів дозволяє обрати найбільш відповідний спосіб залежно від специфіки даних, їх кількості, необхідної точності та часу класифікації. Для вирішення поставленої задачі пропонується використовувати метод дерева рішень, оскільки він має ряд переваг: інтерпретованість моделі, робота з перепустками в даних.

У цьому алгоритмі процес класифікації вхідних об'єктів побудований на кон'юнктивних правилах, що складаються з безлічі предикатів за вхідними даними. Для цього на основі навчальної вибірки виконується побудова логічної схеми, що реалізує організовану систему ієрархічно послідовних питань. Процес класифікації полягає у почерговій відповіді на питання моделі, причому наступне питання залежить від попередньої відповіді. В результаті після відповіді на всі питання визначається єдиний клас, який і є відповіддю класифікації. Також є можливість для кожного предикату моделі зберігати ймовірність відповіді, тоді за неможливості відповісти питанням (наприклад, за відсутності даних), обирається найбільш популярна відповідь, тим самим метод дозволяє оперувати з перепустками у вхідних даних. До недоліків даного методу відноситься неможливість перенавчання моделі, так як навчання використовується вся навчальна вибірка.

На алгоритми обробки інформації в БСМ накладають низку характерних обмежень, таких як розмір програмного коду через обмежений ресурс внутрішньої пам'яті, кількість виконуваних операцій для економії енергоресурсу.

Далі зробимо опис моделі дерева рішень для завдання класифікації БСМ.

Структура даних «дерево» є звичайним пов'язаним графом (без наявності циклів), у якого виділено кореневу вершину. Листком дерева називають вершину, яка не має нащадків. Шлях у графі "дерево" проходить від кореня до листа без повторення вершин. Дерево рішень представлене кореневим деревом, у вершинах якого містяться логічні правила, а листі –

результати класифікації. Логічні правила містять умову у форматі «якщо... то... інакше...». Дерева рішень відмінно справляються із завданням багатокласової класифікації, тобто дозволяють визначати належність вхідного об'єкта до одного із заздалегідь відомих класів. Цільова змінна може мати категоріальні, дискретні та бінарні значення. Алгоритм класифікації стану мережі за допомогою дерева рішень являє собою процес перетворення вхідного вектора даних на результат класифікації. Для кожної вершини дерева використовуються дані вхідного вектору, щоб відповідати на поставлені запитання. У листі дерева розташовуються результати класифікації, і навіть розмір класу, щоб у разі пропуску даних обирати найбільш популярний клас.

Запитання, що знаходяться у вершинах дерева рішень, використовують виключно один параметр вхідного об'єкта та мають уявлення "значення параметра x_{nk} більше значення ...". При позитивній відповіді перехід здійснюється до правого нащадка, при негативному – до лівого. Якщо нащадок містить правило, знову виконується процес відповіді питання. Внаслідок відповідей на запитання досягається один із листів дерева, де міститься інформація про цільовий параметр. За рахунок прозорості інтерпретації роботи алгоритму цей метод має високу наочність.

Для формалізації завдання класифікації станів у БСМ представлено математична модель, яка в загальному вигляді описує набір параметрів, яким надалі слід оперувати запропонованим алгоритмом. Перевагою даної моделі є те, що відбувається абстрагування від технічної реалізації сенсорної мережі, тим самим спрощується застосування методів управління та обробки інформації із суміжних дисциплін. Велика кількість досліджень із завдання класифікації об'єктів представлено у галузі машинного навчання, штучного інтелекту.

Для ухвалення рішення про стан об'єкта використовуються різні моделі класифікації об'єктів, які широко представлені в галузі машинного навчання. Безліч розроблених методів класифікації об'єктів дозволяє вибрати найбільш

підходящий спосіб залежно від специфіки даних, їх кількості, необхідної точності та часу класифікації. Для вирішення поставленої задачі пропонується використовувати метод вирішального дерева, оскільки він має ряд переваг: інтерпретованість моделі, робота з перепустками в даних. На алгоритми обробки інформації в БСМ накладають низку характерних обмежень, таких як розмір програмного коду через обмежений ресурс внутрішньої пам'яті, кількість виконуваних операцій для економії енергоресурсу. У роботі представлена модель обробки інформації в БСМ основі вирішального дерева, що відрізняється від відомих аналогів можливістю локального прийняття рішення про необхідність передачі даних по бездротовій мережі та дозволяє підвищити енергоефективність функціонування бездротової сенсорної мережі.

2 МЕТОДИ І МОДЕЛІ КЛАСИФІКАЦІЇ СТАНІВ В БСМ

Для математичного опису задачі класифікації введемо параметри вхідних даних: N – кількість сенсорів у БСМ, K – кількість унікальних параметрів зовнішнього середовища, контрольованих датчиками (температура, вологість, тощо), X - безліч показань з усіх сенсорів мережі. Кожен сенсор у певний момент часу $t \in T$ представлений вектором $x \in X$, де $n = 1..N$. Цей вектор належить K -мірному простору значень, при цьому елемент x_{nk} може бути виражений бінарним ознакою (спрацьовування тригера) або кількісною ознакою (значення температури, вологості). Отже, значення x_{nk} характеризує показання датчика у складі n -го сенсора, який контролює k -й параметр навколишнього середовища середовища.

У свою чергу сенсорна мережа описується матрицею об'єктів-ознаків:

$$X = \begin{pmatrix} x_{11} & x_{21} & \dots & x_{N1} \\ x_{12} & x_{22} & \dots & x_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1K} & x_{2K} & \dots & x_{NK} \end{pmatrix} \quad (2.1)$$

Також введемо параметри для завдання класифікації: S – кількість станів, у яких може бути БСМ (наприклад, підвищена вологість, спалах, витік газу), Y – безліч станів бездротової сенсорної мережі, які можна записати у вигляді:

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_S \end{pmatrix}. \quad (2.2)$$

У певний момент часу БСМ перебуває у стані $y \in Y$, яка може бути

виражена бінарною або номінальною ознакою. Також можливе уявлення, при якому в один момент часу БСМ перебуває у кількох незалежних станах. Подібне уявлення може бути виражено як набір бінарних параметрів, де кожен параметр характеризує належність мережі до конкретного стану. В результаті для кожного параметра довкілля буде визначено стан $y \in Y$, що є окремим випадком $y \in Y$. Завдання класифікації станів бездротової сенсорної мережі полягає у встановленні залежності (відображення):

$$y^*: X \rightarrow Y. \quad (2.3)$$

Через специфіку сенсорних мереж безліч вхідних даних X має ряд особливостей, що накладають обмеження на методи розв'язання задачі.

2.1 Характеристики ознакового простору в БСМ

Кожен сенсор мережі передає показання від 1 до K параметрів навколишнього середовища середовища, і відсутність будь-якого параметра трактується як пропуск даних у вхідній матриці X . Виділимо два різновиди пропуску даних: зумовлені та випадкові. Найчастіше перепустки у вхідній матриці об'єктів-ознаків будуть визначено заздалегідь, тобто протягом тривалого часу сенсор не надаватиме дані за заздалегідь визначеними параметрами. Це може бути наслідком того, що збирання всіх параметрів навколишнього середовища конкретним сенсором неможливий чи надмірний. В результаті маємо безліч перепусток даних:

$$Q \subseteq X, \text{ где } \forall x_{nk} \in Q : x_{nk} \notin \mathbb{R}. \quad (2.4)$$

Таким чином, при вирішенні завдання необхідно оперувати ознаковим простором $X' = X \setminus Q$. Крім заздалегідь визначених перепусток у даних, у

практиці зустрічаються 2 типи випадкових пропусків даних: пропуск одного параметра сенсора або пропуск всіх даних із сенсора. Перепустка параметра сенсора x_{nk} може бути пов'язаний з несправністю датчика, який знаходиться в n -му сенсорі і зчитує k -й параметр довкілля.

У свою чергу пропуск сенсора x_n може статися або через вихід з ладу n -го сенсора мережі, або через проблем доставки пакетів від вказаного сенсора. Проблема пропуску даних розглядалася у роботі з дослідження вулканічних активностей. Автори роботи зазначили, що стикалися з перепустками в даних через проблеми зі зв'язком під час поганих погодних умов, внаслідок чого частина даних не доходила до базової станції.

Обов'язковою умовою алгоритму детектування подій була наявність даних з усіх сенсорів, і даний алгоритм не був пристосований до оперування з неповним набором вхідних даних. Тому в моменти часу з перепустками в даних система не детектувала події незважаючи на те, що з інших сенсорів дані були успішно доставлені.

Таким чином, алгоритм, що розробляється, повинен бути відмовостійким при перепустках у даних, особливо для випадкових перепусток. Якщо розроблений алгоритм буде некоректним або недостатньо достовірним видавати класифікацію, то ефективність даного алгоритму буде вкрай низька.

Датчики характеризуються похибкою вимірів. Похибка датчика обумовлена його типом, моделлю та складовими елементами. Також, на похибка впливають умови експлуатації (температура, вологість).

Якщо один параметр середовища вимірюється кількома типами датчиків, то число k можна прийняти кількістю різних типів датчиків, і в результаті завдання зводиться до вищезгаданому виразу.

Отже, додатковою умовою до завдання класифікації встановлюється забезпечення перешкодостійкості алгоритму за допустимих похибки.

Кількість датчиків у мережі може постійно змінюватися: несправні сенсори видаляються з мережі, нові – додаються. Розроблений алгоритм

повинен враховувати дані обставини та за будь-якої зміни складу мережі продовжувати виконувати класифікацію з достатньою точністю, а також мати можливість переконфігурації моделі.

Значення параметрів навколишнього середовища можуть містити тренд і циклічність (час доби, сезонність). Ця властивість вимагає, щоб алгоритм був адаптивний до трендів і циклічності, оскільки статично заданий алгоритм один період часу обчислюватиме коректний результат, а в інший час - хибний, що вкрай неприпустимо.

Налаштування алгоритму щоразу при зміні тренду є ресурсозатратним процесом, а у разі несвоєчасного оновлення алгоритму можливе некоректне функціонування системи. Безліч станів системи Y також може змінюватися в часі чинність нових технічних вимог. Деякі стани сенсорної мережі можуть втрачати актуальність, а з появою нових типів сенсорів можуть відстежуватися раніше невідомі стани.

2.2 Модель завдання класифікації станів в БСМ

Враховуючи раніше викладені характеристики параметрів вхідних даних, модернізуємо модель завдання класифікації. Позначимо X – безліч параметрів навколишнього середовища на всьому протягом функціонування мережі Y – безліч усіх станів БСМ.

$$X^t = \begin{pmatrix} x_{11} & x_{21} & \dots & x_{N1} \\ x_{12} & x_{22} & \dots & x_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1K} & x_{2K} & \dots & x_{NK} \end{pmatrix}, Y^t = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_S \end{pmatrix}, \quad (2.5)$$

де N – кількість сенсорів у мережі при $t = [T_x; T_{x+1}]$, K – кількість контрольованих параметрів довкілля при $t = [T_x; T_{x+1}]$, $s = 1..S$ (S – кількість станів бездротової сенсорної мережі при $t = [T_x; T_{x+1}]$).

2.3 Алгоритм побудови бінарного дерева рішень

Алгоритм побудови бінарного дерева рішень представлений на рис. 5. На першій ітерації алгоритму розглядається вся безліч вхідних даних, далі алгоритм повторюється рекурсивно для підмножин, отриманих при розбиття:

Крок 1. Перевірка умови виходу з алгоритму. При виконанні умови слід вибрати для поточного вузла (аркуша дерева) значення з множини Y (результат класифікації вхідного об'єкта), інакше переходимо до виконання кроку 2.

Крок 2. Безліч даних S розбивається на 2 непересічні підмножини S_1 і S_2 за раніше встановленим критерієм мінімізації ентропії. Цей критерій виражений як функція $Q_t(x)$, на вхід якої подаються різні розбиття множини S і виконується пошук максимальної якості розбиття (максимальне зменшення ентропії).

Крок 3. Для підмножин S_1 і S_2 алгоритм продовжує виконуватись рекурсивно, для цього кожна підмножина окремо передається до кроку 1.

Для функції розбиття зазвичай використовують формулу розрахунку ентропії, і на кожній ітерації алгоритму прагнуть її мінімізувати. Таким чином, досягається найменший розмір дерева рішень. Для зменшення розміру дерева застосовують спеціальні методики раннього завершення алгоритму, такі як обмеження по глибині дерева, мінімальну кількість об'єктів у множині S , мінімально-допустиму зміну значення ентропії.

В результаті використання даних методик зменшується ймовірність перенавчання моделі дерева рішень, що позитивно позначається на точності класифікації моделі, і навіть зменшується загальний розмір моделі.

Спосіб розрахунку абсолютного значення енергоефективності пропонованого алгоритму досить складно висловити, оскільки підсумковий результат в основному залежить від динаміки параметрів довкілля. Для більш точної оцінки енергоефективності запропонованого адаптивного енергозберігаючого алгоритму необхідно мати дані щодо поведінки

параметрів зовнішнього середовища.

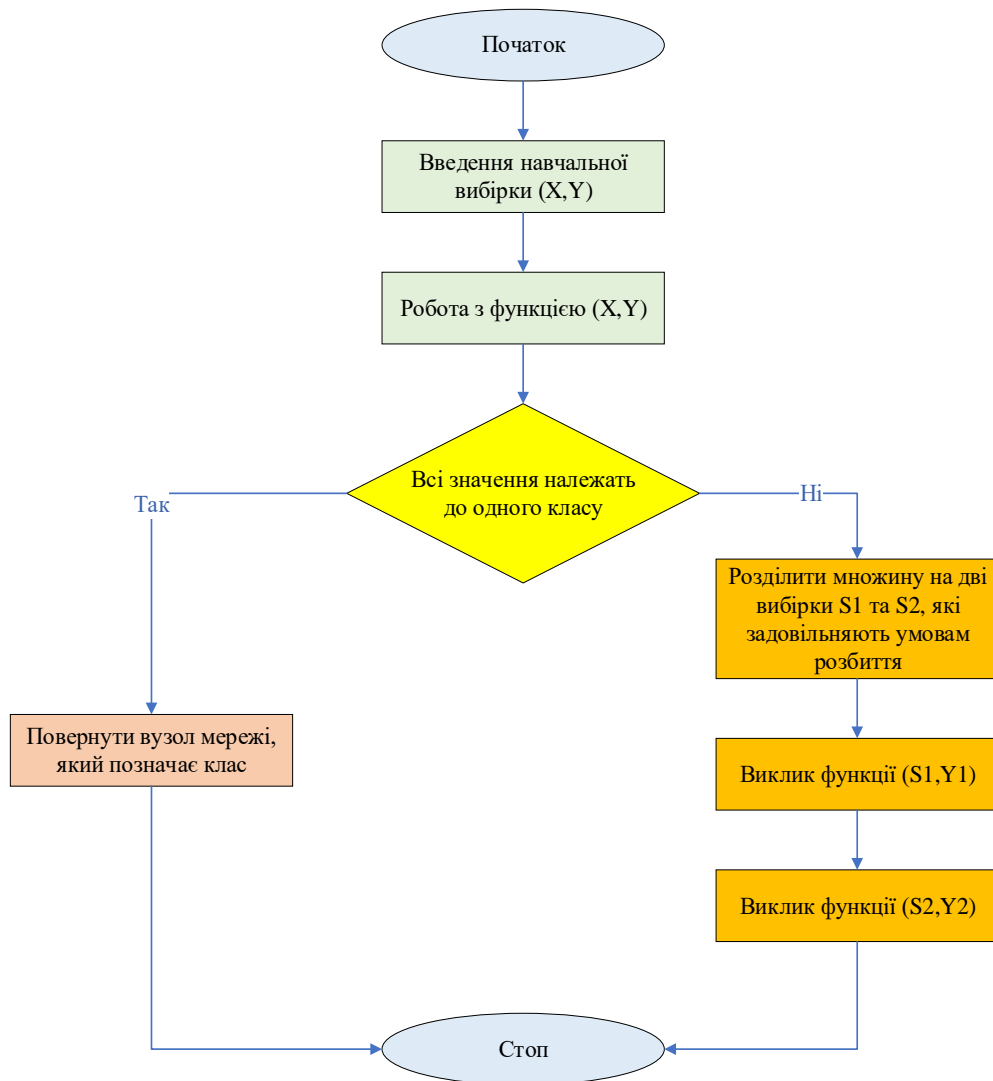


Рисунок 1.5 – Алгоритм побудови бінарного дерева рішень

Реалізація процесу ухвалення рішення про необхідність передачі даних на головний модуль має розташовуватися в логіці роботи вузла мережі, оскільки першоджерелом інформації є дані з датчиків, що знаходяться безпосередньо у складі вузла мережі.

З іншого боку, для ухвалення рішення про передачі даних можливо знадобляться знання про процес класифікації, то Існують дані про математичну модель прийняття рішення про стан мережі [7]. Для побудови моделі класифікації потрібно мати навчальну вибірку, яка представлена у

вигляді бази знань про раніше прийняті рішення. Даний факт обмежує застосування даного методу, тому що не завжди є подібна база, та її формування – трудомістка завдання. Також недоліком методу дерева рішень є необхідність навчання нової моделі дерева рішень за зміни параметрів моделі класифікації.

Розглянемо граничні випадки поведінки бездротової сенсорної мережі:

- постійний стан сенсорної мережі. У такому разі передачі даних не буде, оскільки передача даних означала б, що мережа змінює своє стан, що суперечить початковій умові. Таким чином, основна частина електроенергії використовуватиметься на вимір параметрів довкілля.

- постійна зміна станів сенсорної мережі. В цьому випадку на кожен період зчитування параметрів зовнішнього середовища виконуватиметься передача даних. У кращому разі, передавати дані буде лише один сенсор, у гіршому - всі сенсори.

Спосіб розрахунку абсолютного значення енергоефективності пропонованого алгоритму досить складно висловити, оскільки підсумковий результат в основному залежить від динаміки параметрів довкілля. Для більш точної оцінки енергоефективності запропонованого адаптивного енергозберігаючого алгоритму необхідно мати дані щодо поведінки параметрів зовнішнього середовища.

3 ПРОГРАМНА РЕАЛІЗАЦІЯ АЛГОРИТМУ ТА ТЕСТУВАННЯ

3.1 Вимоги до алгоритму

Функціональні вимоги до алгоритму, що розробляється, формуються виходячи з загальноприйнятих вимог до програмного забезпечення, а також технічних вимог у сфері застосування БСМ. До технічних обмежень алгоритму відносяться:

- розмір обсягу програми (через обмежений розмір внутрішньої пам'яті вузла мережі);
- глибина вкладеності стеку (через обмежений розмір оперативної пам'яті вузла мережі);
- набір інструкцій (специфікація мікроконтролера).

До вимог алгоритму для застосування в БСМ відносяться:

- швидкість виконання програми;
- формат вхідних/вихідних даних;
- автономність.



Рисунок 3.1 – Python3

В якості лінгвістичного забезпечення для реалізації алгоритму обрано мову високого рівня Python3 (рисунок 3.1). Вибір мови обумовлений тим, що

на його базі представлено популярну бібліотеку scikit-learn (рисунок 3.2), що містить готові реалізації моделей машинного навчання, у тому числі дерева рішень – DecisionTreeClassifier (рисунок 3.3), а також супутні методи перевірки якості класифікації, регуляризації моделі.

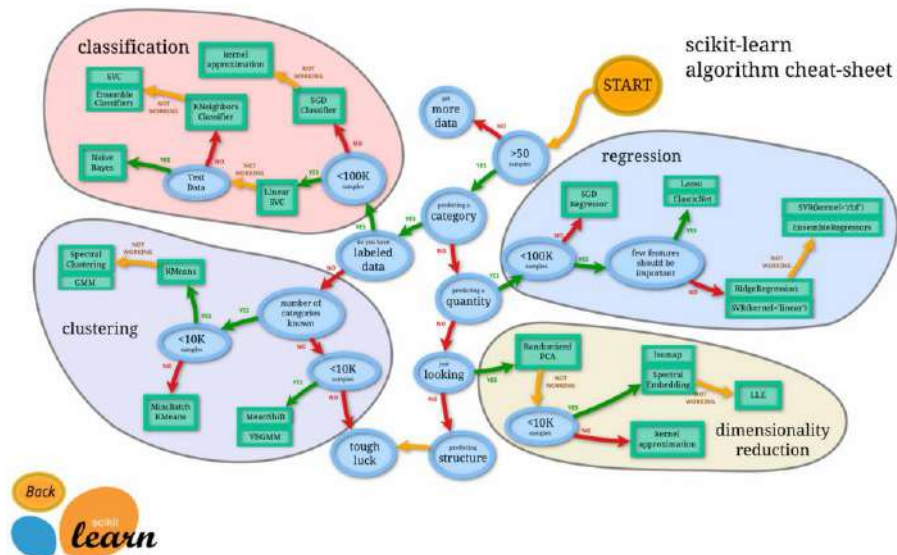


Рисунок 3.2 – Scikit-learn

Розробка виконана відповідно з об'єктно-орієнтованими парадигмами.

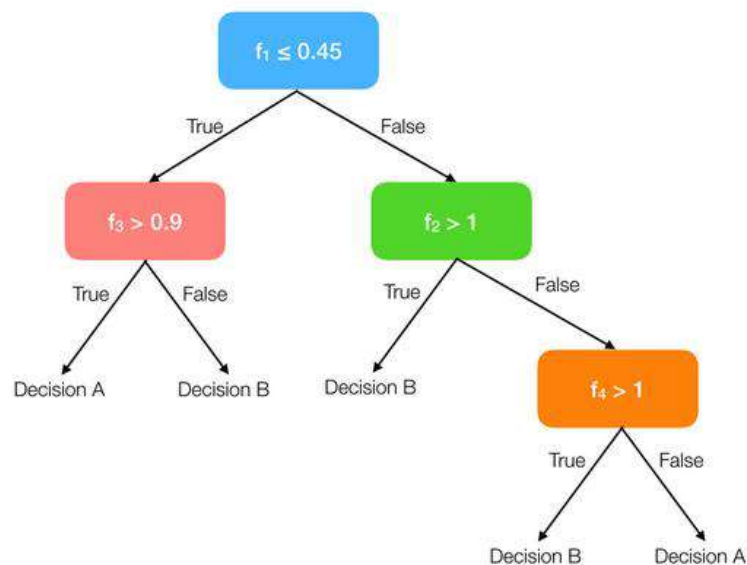


Рисунок 3.3 – DecisionTreeClassifier

Для навчання моделі класифікації необхідно інтерпретувати основу знань у навчальну вибірку X . Дані сенсорів у конкретний час, y – стан системи.

Далі необхідно виконати такі етапи:

- провести навчання моделі на навчальній вибірці;
- безліч предикатів отриманої моделі розбити на підмножини для кожного сенсора;
- відсортувати предикати за зростанням (предикати у дереві розташовані в хаотичному порядку),
- надіслати отриманий список предикатів сенсорю. Для функціонування алгоритму кожному сенсорю потрібно мати дані про предикати для локального ухвалення рішення про необхідність передачі даних.

Предикати сенсора можна доставити двома способами: бездротовою мережею або зберегти на локальному запам'ятовуючому пристрої до введення сенсора в експлуатацію.

Перший спосіб є енерговитратним, оскільки породжує додаткові сеанси зв'язку в мережі, але за умови, що для тривалого функціонування мережі буде потрібно наступне коригування предикатів, то залишається найзручнішим. Використання бази даних не потрібно, оскільки розмір списку предикатів має невеликий обсяг (порядку кілобайта) і не потрібно додаткових операцій до роботи зі списком.

3.2 Структурно-функціональна модель алгоритму

Перший рівень декомпозиції структурно-функціональної моделі визначає процеси, що відбуваються під час класифікації стану в БСМ за допомогою запропонованого алгоритму, а саме навчання дерева рішень та формування числових діапазонів для функції локального прийняття рішення лише на рівні сенсора. Оскільки етапи навчання моделі та формування

числових діапазонів виконуються на боці головного модуля, споживання автономного енергоресурсу вузла БСМ немає.

Одним з механізмів системи виступає база знань, сформована при попереднього функціонування БСМ або її моделювання. З цієї бази знань формуються параметри, які входять до безлічі навчальної вибірки. Після формування навчальної вибірки виконується навчання моделі класифікації "дерево рішень".

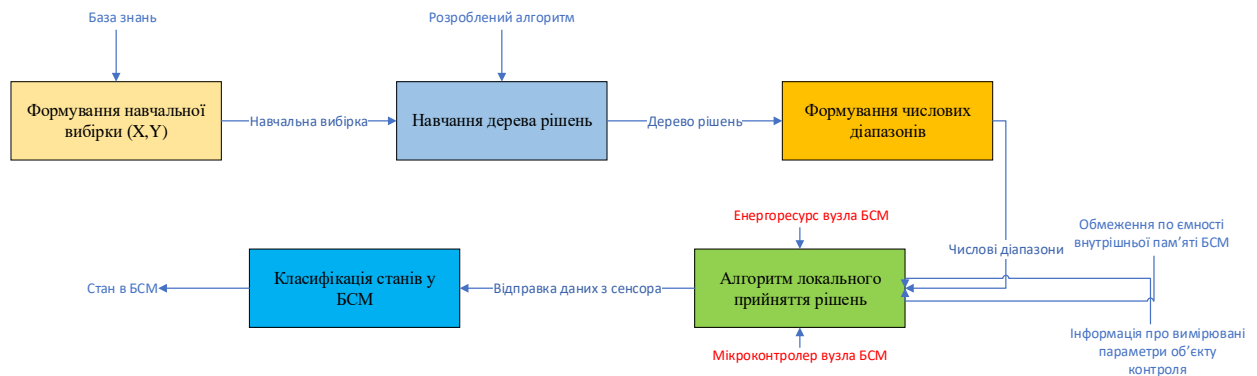


Рисунок 3.4 – Структурно-функціональна модель адаптивного енергозберігаючого алгоритму

Для генерації моделі використовують параметри регуляризації, які дозволяють контролювати глибину дерева, точність класифікації. Параметри вирішального дерева використовують як для класифікації станів у БСМ, так і для локального прийняття рішення про необхідності передачі даних через мережу. Оскільки сенсор може бути вимкнений, а дані про числові діапазони зберігаються в оперативній пам'яті, то також додано блок «Зберігання числових діапазонів», який підвищує відмовостійкість сенсора.

Специфікація ролей адаптивного енергозберігаючого алгоритму класифікації станів мережі Ролі описують категорії користувачів, відрізняються різними функціями та правами доступу залежно від розв'язуваної задачі. Базові ролі:

- "оператор". Користувачі, які використовують БСМ та його алгоритм

класифікації для цільових завдань Оператор за необхідності здійснює перевірку коректності класифікації, допомагає класифікувати нові ситуації, ініціює оновлення моделі класифікації;

- "аналітик". Користувач, відповідальний за підготовку навчальної вибірки, навчання моделі класифікації, оцінку точності та повноти класифікації;

- "інженер". Користувач, який підтримує функціонування БСМ, опрацьовує протокол оновлення параметрів моделі класифікації, контролює технічні характеристики системи, у тому числі – енергоспоживання.

Представлена діаграма (рисунок 3.5) розгортання системи, яка містить структуру та склад програмної частини БСМ, що виконує класифікацію станів при допомозі запропонованого алгоритму.

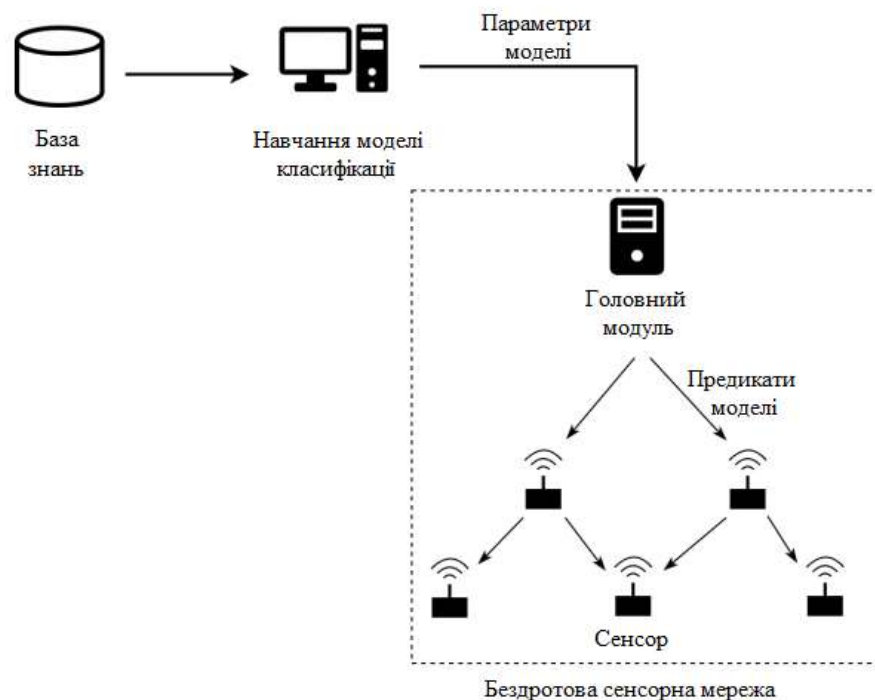


Рисунок 3.5 – Діаграма розгортання алгоритму енергозбереження класифікації станів

Програмна частина включає підготовку навчальної вибірки з бази знань, навчання та налаштування моделі класифікації, підготовку предикатів

до передачі на сенсори та алгоритм локального прийняття рішення про необхідність передачі даних по бездротовій мережі. З бази знань формується навчальна вибірка для моделі класифікації, далі параметри моделі передаються на головний модуль БСМ, який, у свою чергу, відправляє числові діапазони бездротової мережі. Після того, як дані про предикати моделі доставлені на всі сенсори, енергозберігаючий алгоритм може почати енергоефективну класифікації станів у БСМ.

3.3 Проведення експериментів

Для навчання моделі класифікації необхідно мати базу знань, представлену у форматі параметрів навколишнього середовища та прийнятих рішень про стан мережі. Ця вибірка може бути отримана або з набору історичних даних або при натурних або напівнатурних випробуваннях. Також важливо перевірити алгоритм щодо стійкості до перепусток у вхідних даних.

Проведення натурних випробувань є витратною процедурою, оскільки симуляції різних вхідних умов потрібна велика кількість ресурсів. Також натурні випробування мають на увазі наявність складного технічного забезпечення для функціонування БСМ, тим самим складно сконцентруватися винятково на дослідженні розробленого алгоритму. Тому у роботі розглянуто напівнатурну модель БСМ, в якій виконано генерацію вхідних даних із урахуванням поведінки реальних параметрів довкілля.

Рівень достовірності результатів дослідження розробленого алгоритму на напівнатурній моделі прямо пропорційно залежить від ступеня відповідності згенерованих вхідних даних параметрів реальних процесів. У разі якщо кореляція зазначених параметрів прагнути до 1, можна говорити про 80-90% достовірності результатів дослідження.

Напівнатурне моделювання є найбільш пріоритетним способом дослідження, оскільки його проведення вимагає значно менше часових та

ресурсних витрат. Беручи до уваги той факт, що рівень достовірності напівнатурної моделі багато в чому залежить від характеристик вхідних даних, то для підвищення достовірності напівнатурної моделі велика увага приділена способу генерації вхідних даних.

Розглянемо процес генерації тестових даних для оцінки енергоефективності алгоритму, що розробляється. Для цього накладемо обмеження на гауссівські процеси у вигляді значень реальних даних, щоб визначити параметри генерації – математичне очікування та дисперсію, які дозволяють згенерувати випадкові послідовності, найбільш схожі щодо поведінки з реальними даними (рисунок 3.5)

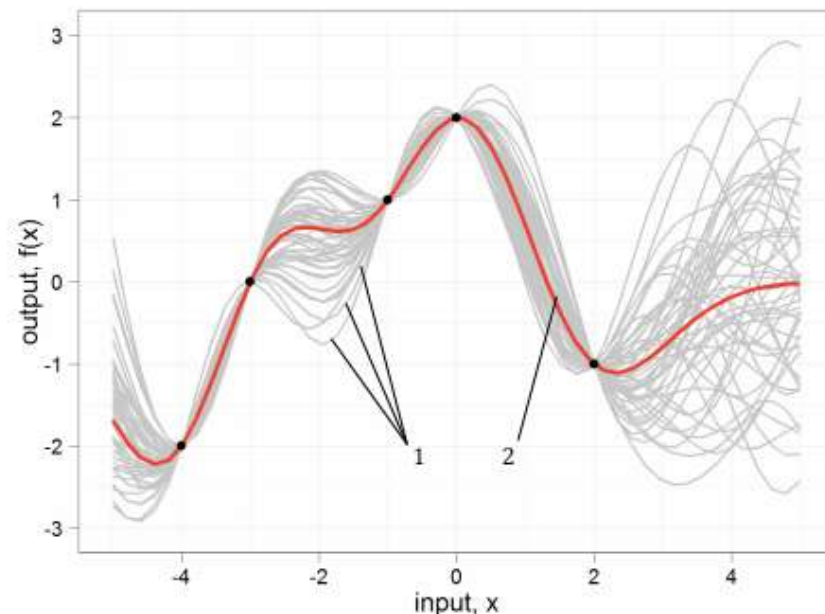


Рисунок 3.5 – Приклад накладених умов на гаусівський процес: 1 – гаусівські процеси з накладеними умовами, 2 – середнє значення безлічі гаусівських процесів

Використовуючи параметри, отримані при накладенні обмежень на гаусівський процес, згенеруємо випадкові послідовності для кожного параметри навколишнього середовища, що контролюють сенсори БСМ (рисунок 3.6).

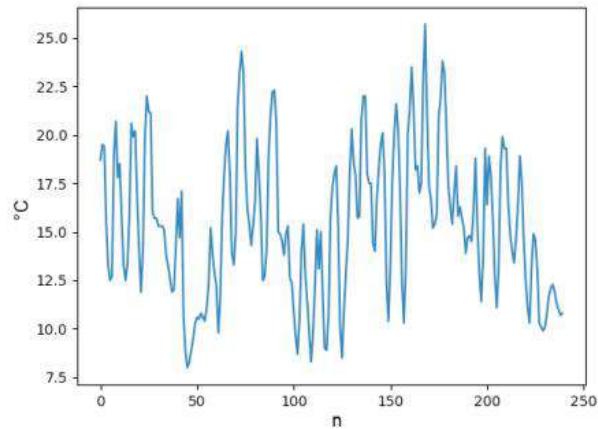


Рисунок 3.6 – Тестовий набір даних

Створення випадкової послідовності, побудованої за допомогою гаусівського процесу, показало, що в нескінченності значення випадкової величини значно віддаляються від вихідного математичного очікування. Дані значення мають явний тренд у негативну чи позитивну напіввісь, що дуже рідко відбувається у параметрів навколишнього середовища, які часто мають обмежену область значення. При додатковому Дослідженні було виявлено, що сума безлічі випадкових послідовностей прагне до первісного математичного очікування (рисунок 3.7).

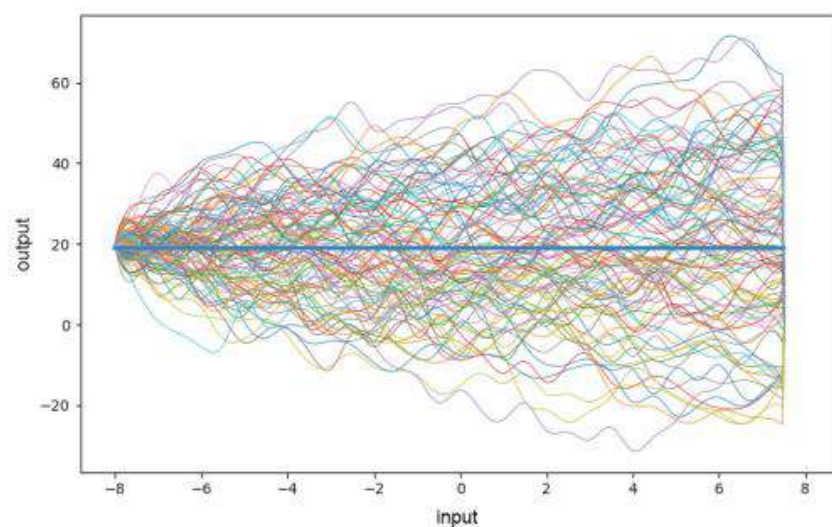


Рисунок 3.7 – Безліч гаусівських процесів з єдиним початковим значенням
(лінія – сума випадкових величин)

Отже, для обмеження зони значення випадкової послідовності можна використовувати суму кількох випадкових послідовностей за правилом: чим більше випадкових послідовностей бере участь у сумі, тим менше область значення підсумкової випадкової величини.

Для перевірки можливості генерації бажаної залежності було написано програмний код, завдання якого полягало в тому, щоб послідовно генерувати еталонну вибірку за допомогою нормального розподілу з заданою точністю, при цьому, якщо випадкова величина не потрапляла у потрібну точність, то спроба зараховувалася як невдала та відбувалася повторна генерація. Кількість генерацій сильно залежить від параметрів нормального розподілу: математичного очікування та дисперсії. Для цього за математичне очікування було прийнято перше значення еталонної послідовності, а дисперсія була встановлена емпіричним шляхом.

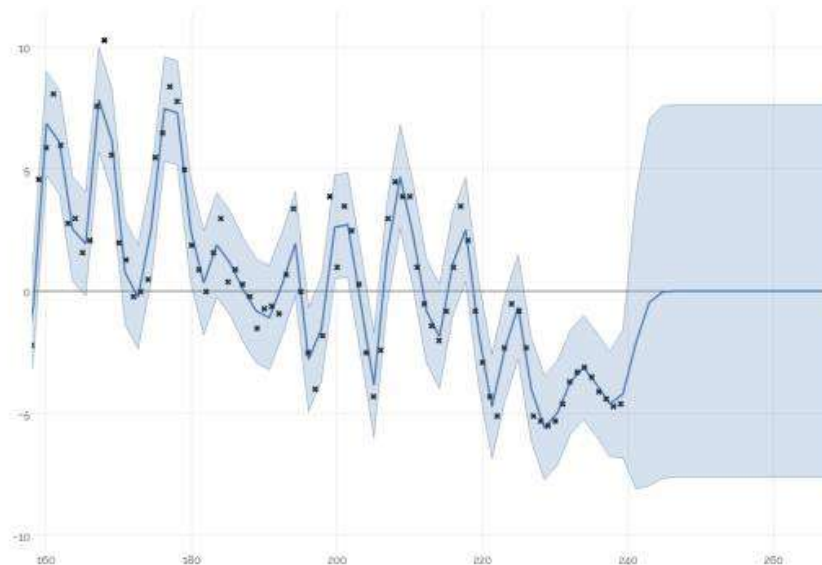


Рисунок 3.8 – Результат генерації випадкової величини

Результати генерації випадкової послідовності, що складається із суми випадкових послідовностей зі вставленими параметрами математичного очікування та дисперсії представлені на рисунку 3.8. На графіці відображається розкид значень випадкової величини, суцільна лінія –

математичне очікування випадкової величини у конкретний час. З графіка видно, що випадкова величина змінюється у межах області значення.

В результаті було встановлено, що гауссівські процеси здатні генерувати дані, при цьому генерація відбувається за певну кількість ітерацій, що описується нормальним розподілом. Як альтернативне рішення, було розглянуто генерацію за допомогою рівномірного розподілу. Результати подібної генерації показали, що для рівномірного розподілу потрібно більше ітерацій, а кількість ітерацій має сильні викиди, що перевершують математичне очікування у рази.

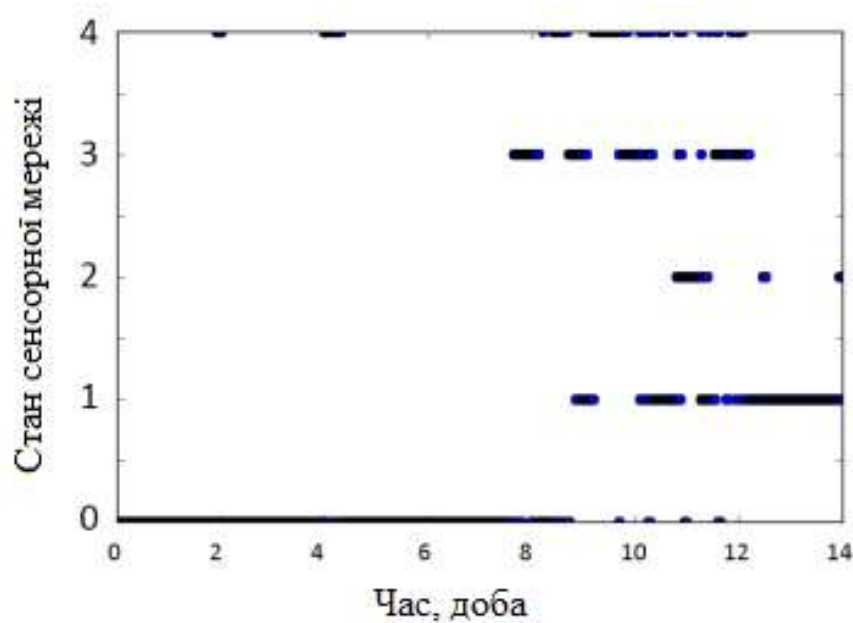


Рисунок 3.9 – Графік розподілу станів сенсорної мережі (0 – норма, 1 – перегрів, 2 – займання, 3 – підвищена вологість, 4 – переохолодження)

Для дослідження енергоефективності алгоритму було проведено експеримент із класифікації станів мережі. Для цього об'єкти тестуючої вибірки було подано на вхід моделі класифікації, а потім виконано порівняння результатів класифікації та еталонних значень.

ВИСНОВКИ

Розроблено алгоритми роботи сенсора БСМ та побудови бінарного дерева рішень, які можуть бути використані для підвищення енергозбереження функціонування вузлів в бездротових сенсорних мережах. Розглянуто методи машинного навчання. Для реалізації запропонованого алгоритму обрано метод дерева рішень. Вибір пов'язаний з тим, що даний метод має високу інтерпретацію та схожий із процесом прийняття рішення оператором. Грунтуючись на ідемпотентності предикатів дерева рішень, пропонується скоротити кількість сеансів зв'язку за рахунок динамічного визначення частоти передачі. Таким чином, цей підхід дозволить позбутися постійної передачі даних бездротовим каналом зв'язку, тим самим буде зекономлено енергоресурси вузлів мережі.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Гаптельманов А. В., Міхаль О. П., Щепка О. О. Підвищення енергозбереження бездротових сенсорних мереж з використанням методів машинного навчання // Системи управління навігації та зв'язку. Випуск 2 (72). Національний університет «Полтавська політехніка імені Юрія Кондратюка». Полтава 2023 р. С. 89-95.
2. Akyildiz, I. F., Weilian, S. A survey on sensor networks // IEEE Communications magazine. 2002. 40(8). P. 102–114. doi: 10.1109/MCOM.2002.1024422.
3. Abbas, I., Liu, J., Faheem, M., Noor, R. S. Different sensor based intelligent spraying systems in Agriculture // Sensors and Actuators A: Physical. 2020. Volume 316. P. 1–16.
4. Alias, F., Carrie, J. C., Alsina-Pages, R. WASN-Based day–night characterization of urban anomalous noise events in narrow and wide streets // Sensors. 2020. 20(17) 26 p.
5. Chen, Y., Nasser, N. Energy-balancing multipath routing protocol for wireless sensor networks // Quality of service in heterogeneous wired/wireless networks. New York: Qshine, 2006. Vol. 21. – P. 245–249. doi: 10.1145/1185373.1185401.
6. Das, K., Behera, R. N. A survey on machine learning: concept, algorithms and applications // International Journal of Innovative Research in Computer and Communication Engineering. 2017. 5(2). P. 1301–1309.
7. Gao, T., Greenspan, D., Welsh, M., Juang, R. R., Alm, A. Vital signs monitoring and patient tracking over a wireless network // Proceedings of the 27th IEEE EMBS Annual International Conference. 2005. P. 66–74.
8. Huang, J.H., Amjad, S., Mishra, S. CenWits: A sensor-based loosely coupled search and rescue system using witnesses // Proceedings of the Third International Conference on Embedded Networked Sensor Systems (Sensys), San

Diego, CA. 2005. P. 180–191.

9. Monitoring volcanic eruptions with a wireless sensor network / Werner-Allen, G., Johnson, J., Ruiz, M., Lees, J., Welsh, M. // Proceedings of the Second European Workshop on Wireless Sensor Networks, IEEE. Istanbul, 2005. 13 p.

10. Hardware design experiences in ZebraNet / Zhang, P., Sadler, C. M., Lyon, S. A., Martonosi, M. // Proceedings of the SenSys'04, Baltimore, MD. 2004. P. 227–238. Li Q.-L. (2016), “Nonlinear Markov processes in large networks”, Special Matrices, Vol. 4(1), doi: <https://doi.org/10.1515/spma-2016-0019>

11. Yick, J., Mukherjee, B., Ghosal, D. Analysis of a Prediction-based Mobility Adaptive Tracking Algorithm // Proceedings of the IEEE Second International Conference on Broadband Networks (BROADNETS), Boston. 2005. Vol. 1. P. 753–760

12. Yu, Z., Tsai J. J., Weigert T. An automatically tuning intrusion detection system // IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). 2007. 37(2). P. 373–384.

13. Design considerations for solar energy harvesting wireless embedded systems / Raghunathan, V., Kansai, A., Hse, J., Friedman, J., Srivastava, M. // Proceedings of the IPSN. 2005. P. 457–462.

14. Second Generation System-on-Chip Solution for 2.4 GHz IEEE 802.15.4 / RF4CE ZigBee [Electronic resource] // Texas Instruments. Available at: <http://www.ti.com/lit/ds/symlink/cc2530.pdf/>. 05.05.2023. 36 p

15. Pelckmans, K., Jos, D. B., Johan, A. S., Bart, D. M. Handling missing values in support vector machine classifiers // Neural Networks 18. 2005. №5-6. P. 684–692

16. Venkatesan, R., Meng, J. E. A novel progressive learning technique for multi-class classification // Neurocomputing. 2016. 207. P. 310–321.

17. Yu, Z., Tsai, J. J. P. A Framework of Machine Learning Based Intrusion Detection for Wireless Sensor Networks // 2008 IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing. 2008. P. 272–279.