DOI: 10.36910/6775-2524-0560-2019-36-20
УДК 681.516
Сироткіна О.І., к.т.н.; Алексєєв М.О., д.т.н., проф.; Мещеряков Л.І., д.т.н., проф.; Мороз Б.І., д.т.н., проф.
Національний технічний університет «Дніпровська політехніка»

МЕТОДИ РОБОТИ З «ВЕЛИКИМИ ДАНИМИ» НА ОСНОВІ ЗАСТОСУВАННЯ ТЕОРІЇ «*M*-АРНИХ КОРТЕЖІВ»

Сироткіна О.І., Алексєєв М.О., Мещеряков Л.І., Мороз Б.І. Методи роботи з «великими даними» на основі застосування теорії «*m*-арних кортежів». Статтю присвячено питанням створення і застосування математичних методів роботи з «великими даними» з метою мінімізації задіяних часових і обчислювальних ресурсів для структури організації даних типу «*m*-арні кортежі на основі впорядкованих множин довільної потужності». Проведено огляд і виконано аналіз математичних методів, які використовуються при вирішенні завдань даного класу. Сформульовані деякі властивості структури організації даних типу «*m*-арні кортежі на основі впорядкованих множин довільної потужності», які є наслідком логічних правил формування даної структури. Виведено набір функціональних залежностей між *m*-арними кортежами на основі їх розташування в структурі, яка визначається парою індексів (*j*, *m*) при приватних початкових умовах. Наведено наочну ілюстрацію графа булеана. Обкреслено вершини графа, визначені за допомогою виведених аналітичних залежностей, як елементи булеана, які включають в себе один і той же загальний елемент при приватних початкових умовах. Розраховано порівняльну оцінку часу виконання описаного в статті методу роботи з "великими даними" та методів-аналогів. Отримано логічні висновки про вплив досліджуваних властивостей та методів роботи зі структурою типу «*m*-арні кортежі на основі впорядкованих множин довільної потужності» на мінімізацію задіяних обчислювальних ресурсів.
Ключові слова: «великі дані», структура організації даних, впорядковані множини довільної потужності, *m*-арні кортежі, граф булеана, оцінка часу виконання алгоритму, мінімізація часових та обчислювальних ресурсів.

Сироткина Е.И., Алексеев М.О., Мещеряков Л.И., Мороз Б.И. Методы работы с «большими данными» на основе применения теории «*m*-арных кортежей». Статья посвящена вопросам создания и применения математических методов работы с «большими данными» с целью минимизации задействованных временных и вычислительных ресурсов для структуры организации данных типа «*m*-арные кортежи на основе упорядоченных множеств произвольной мощности». Проведен обзор и выполнен анализ математических методов, которые используются при решении задач данного класса. Сформулированы некоторые свойства структуры организации данных типа «*m*-арные кортежи на основе упорядоченных множеств произвольной мощности», которые являются следствием логических правил формирования данной структуры. Выведен набор функциональных зависимостей между *m*-арными кортежами на основе их местоположения в структуре, определяемого парой индексов (*j*, *m*) при частных начальных условиях. Приведена наглядная иллюстрация графа булеана. Очерченные вершины графа определены при помощи выведенных аналитических зависимостей как элементы булеана, включающие в себя один и тот же общий элемент при частных начальных условиях. Выполнена сравнительная оценки времени исполнения описанного в статье метода работы с «большими данными» и методов-аналогов. Получены логические выводы о влиянии исследуемых свойств и методов работы со структурой типа «*m*-арные кортежи на основе упорядоченных множеств произвольной мощности» на минимизацию задействованных вычислительных ресурсов.
Ключевые слова: «большие данные», структура организации данных, упорядоченное множество произвольной мощности, *m*-арные кортежи, граф булеана, оценка времени исполнения алгоритма, минимизация временных и вычислительных ресурсов.

O. Syrotkina, M. Alekseyev, L. Meshcheriakov, B. Moroz. Methods of working with "big data" based on the application of "m-tuple" theory. This article addresses the creation and application of mathematical methods to work with "big data". This allows us to minimize the time and computational resources involved for a data organization structure "*m*-tuples based on ordered sets of arbitrary cardinality (OSAC)". We reviewed and analysed mathematical methods used to solve problems of this class. We formulated several properties of this data organization structure which are a consequence of the logical rules for the formation of this structure. A set of functional dependencies between *m*-tuples is derived based on their location in the structure and is determined by a pair of indices (*j*, *m*) under particular initial conditions. It is also given the illustration of the Boolean graph. The outlined vertices of the graph are determined using the derived analytical dependencies as Boolean elements that include the same general element under particular initial conditions. We made a comparative evaluation between the execution time for the method to work with "big data" and the analogue methods described in the article. We obtained logical conclusions about the influence of the properties studied and methods to work with the structure "*m*-tuples based on OSAC" on the minimization of the computing resources involved.
Keywords: "big data", data organization structure, ordered sets of arbitrary cardinality, *m*-tuples, Boolean graph, estimation of algorithm execution time, minimization of time and computing resources.

Вступ. У зв'язку з сучасним соціально-економічним розвитком суспільства, глобальними змінами у всіх сферах діяльності, прискореним зростанням високотехнологічних і наукоємних виробництв, широкомасштабним впровадженням інформаційно-телекомунікаційних технологій ставляться підвищені вимоги до інформаційно-технологічних рішень, пов'язаних із завданнями обробки, зберігання, аналізу та управління «великими даними» в умовах часових, обчислювальних та інформаційних обмежень. Одним з основних напрямків розвитку сучасної інформаційно-технологічної інфраструктури є створення і впровадження різноманітних методів і алгоритмів достовірного і своєчасного аналізу великого потоку інформації, а також інструментарію для роботи з «великими даними». Як відомо [1], до основних характеристик «великих даних і методів їх обробки відносяться наступні параметри: обсяг даних, швидкість їх обробки, різноманіття типів, достовірність, життєздатність, цінність і мінливість даних. Таким чином, актуальним завданням є створення методів роботи з «великими даними», що забезпечують пропоновані вимоги до кількісних і якісних характеристик оброблюваних даних.

Аналіз досліджень. На даний час зберігання і обробка «великих даних» в системах на основі реляційних БД не завжди є високоефективною. Для зберігання даних в реляційної БД необхідно виконати кілька етапів: розробити структуру зберігання даних; профільтрувати та перетворити дані в необхідний формат; завантажити дані в БД. Кожен з цих етапів може бути тривалим і трудомістким процесом.

Один з напрямків для вирішення даної проблеми − створення нового класу NoSQL (Not Only SQL) систем. Перевагою NoSQL систем є гнучкість моделей даних, можливості горизонтального масштабування і паралельної обробки, швидкість отримання результатів.

З точки зору методів аналізу «великих даних» подальший розвиток отримали машинне навчання, штучний інтелект, розподілена обробка потоків і подій, візуальні методи дослідження даних.

У статті [2] описуються моделі інтероперабельності і навігації за «великими даними», що дозволяють витягувати інформацію з великих і складних колекцій слабоструктурованих даних. Відмінність запропонованих моделей від традиційної реляційної моделі даних полягає в тому, що інформація про структуру великих масивів даних визначена в термінах ідентифікатор домену та бізнес-ключа. Використання математичної моделі домену / ключа дозволяє створювати структуру, яка може усунути всі обмеження, крім обмежень домену та обмежень ключа. Але практична реалізація перетворення колекції слабоструктурованих даних в нормальну форму домену/ключ залишається відкритим питанням.

У роботі [3] розглянуто математичні методи та принципи побудови експертних систем. В роботі описуються такі поняття, як система альтернатив, модуль несумісності, модуль заборони, модуль імплікації, модуль продукції. Показано реалізацію логічного висновку в системі альтернатив. На думку авторів, метод систем альтернатив призначений для організації та обробки даних при вирішенні завдань, що пов'язані з перебором варіантів. Іншим способом застосування даного методу є можливість створення бази знань експертної системи у вигляді системи альтернатив для вирішення задач розпізнавання і класифікації. Також в роботі розглянуті методи відновлення регулярних і контекстно-вільних мов. Ці методи застосовуються автором при автоматизованому витягу знань з бази знань експертної системи.

В роботі [4] розглянуто принципи побудови експертної системи для діагностики об'єкта з неструктурованим діагностичним ознаками (НДО). В роботі показано, що значення НДО можуть бути задані набором еталонів. Суть методу діагностування, запропонованого в зазначеній роботі, полягає у визначенні коефіцієнта подібності об'єкта діагностики з еталоном. База знань експертної системи задається графом. Вершини графа − набор НДО, дуги − набор еталонів, що задають значення НДО. Після обчислення коефіцієнтів подібності об'єкта діагностики з еталоном, їх значення присвоюються дугам. Після цього розраховується ймовірність зв'язності початкової вершини графа з термінальними вершинами. У роботі наведено висновок: якщо ступінь помітності діагнозів відповідає заданому значенню, то процес діагностування завершується. Інакше – процес триває для наступного НДО.

В роботі [5] розглядається концепція діагностики роботи інформаційних систем (ІС) на основі LV (Latent Variable) моделей латентних змінних. На базі масиву архівних даних, зібраних SCADA системою, визначаються параметри процесу, які завищили статистику моніторингу. Ці параметри пов'язані з роботою обладнання, поведінкою технологічного процесу або порушеннями режимів роботи. Підхід засновано на здатності експлуатаційного персоналу підприємства сформулювати експертну оцінку тенденцій протікання технологічного процесу і роботи обладнання. Для автоматизації процесу формування експертної оцінки можуть бути застосовані системи підтримки прийняття рішень, засновані на знаннях. При цьому LV моделі повинні бути унікальні і інтерпретовані. Також потрібна обробка великих масивів відсутніх даних, контроль їх цілісності всередині області LV, що встановлюється за навчальною вибіркою. Обмеження застосування LV моделей, побудованих на архівних даних системи, полягає в обмеженні простору латентних змінних, що визначається самими архівними даними. LV моделі також можна використовувати для екстраполяції режимів роботи ІС, для яких відсутня статистика моніторингу.

Однією з найважливіших проблем застосування методів роботи з «великими даними» є їх обчислювальна осяжність (або неосяжність). Якщо із збільшенням кількості оброблюваних даних, число операцій зростає експоненціально, то таке явище називається «комбінаторним вибухом» [6]. Ще однією аналогічною проблемою, пов'язаною з «комбінаторним вибухом», є розв'язність методу (досягнення термінального стану з деякого ініціального стану) при експоненціальному зростанні числа станів досліджуваного об'єкта або процесу.

Для вирішення проблем «комбінаторного вибуху» виконуються, якщо є можливість, методи скорочення простору аналізованих станів, тобто розбиття його на досить незалежні підпростори з характерними неповними рішеннями.

Для скорочення простору аналізованих станів розроблено багато методів, в тому числі: використання симетрій при перевірці еквівалентності станів, абстракції на основі дослідження залежностей, абстракції предикатів, накладання обмежень на простір пошуку, спрямований пошук, евристичні методи [6, 7].

Формулювання мети дослідження. Метою нашого дослідження є мінімізації часових і обчислювальних ресурсів на обробку і аналіз «великих даних» шляхом розробки математичного методу скорочення простору аналізованих станів на основі виведення нових функціональних залежностей роботи зі структурою організацією даних (СОД) типу «*m*-арні кортежі на основі впорядкованих множин довільної потужності (ВМДП)».

Опис основних термінів і визначень, а також деяких властивостей і математичних методів роботи з СОД типу «*m*-арні кортежі на основі ВМДП» наведено в [8 − 12].

Постановка завдання в загальному вигляді. Необхідно розробити математичний метод визначення істинності виразу (1) для будь-якої коректної комбінації значень параметрів: n, m_1, m_2, j_1, j_2 з мінімізацією тимчасових і обчислювальних ресурсів на обробку та аналіз даних.

$$\begin{cases} y^n_{m_1,j_1} \quad op \quad y^n_{m_2,j_2} \\ op \in \{\subset, \cup, \cap\} \\ 1 \le m_1 < m_2 \le n \\ 1 \le j_1 \le \begin{pmatrix} n \\ m_1 \end{pmatrix} \\ 1 \le j_2 \le \begin{pmatrix} n \\ m_2 \end{pmatrix} \end{cases}, \qquad (1)$$

де

$y^n_{m,j}$	–	*m*-арний кортеж, елемент булеана 2^X;
n	–	потужність упорядкованої базової множини X;
m_1, m_2	–	довжини кортежів;
j_1, j_2	–	індекси (порядкові номери) *m*-арних кортежів в упорядкованих множинах $\boldsymbol{Y}^n_{m_1}$ та $\boldsymbol{Y}^n_{m_2}$.

Основна ідея даного методу полягає у виведенні набору функціональних залежностей $\{f_i(n, m_1, m_2, j_1, \eta)\}$ при різних початкових умовах (для всіх можливих коректних комбінацій значень аргументів) і визначенні істинності другого виразу системи (2).

$$\begin{cases} J2^n_{m_2}(y^n_{m_1,j_1}) = \{f_i(n, m_1, m_2, j_1, \eta)\} \\ j_2 \in J2^n_{m_2}(y^n_{m_1,j_1}) \\ J2^n(y^n_{m_1,j_1}) = \{J2^n_{m_2}(y^n_{m_1,j_1})\} \\ m_1 < m_2 \le n \end{cases}, \qquad (2)$$

де

η	–	розташування елемента $y^n_{m_1,j_1}$ у кортежі $y^n_{m_2,j_2}$;
$J2^n_{m_2}(y^n_{m_1,j_1})$	–	множина індексів j_2, що визначає елементи $y^n_{m_2,j_2}$ одної з підмножин $Y^n_{m_2}$ булеана 2^X, для яких виконується умова $m_1 < m_2 \le n$ та вираз (1) є істинним;
$J2^n(y^n_{m_1,j_1})$	–	множина індексів j_2, що визначає елементи $y^n_{m_2,j_2}$ всіх підмножин $Y^n_{m_2}$ булеана 2^X, для яких виконується умова $m_1 < m_2 \le n$ та вираз (1) є істинним.

Можливість виведення нових функціональних залежностей між елементами впорядкованої СОД випливає з властивостей та правил формування СОД. Функціональні залежності між m-арними кортежами виводяться на основі аналізу місця розташування елементів в СОД, що задається парою індексів (j, m).

У роботі [9] були розглянуті математичні методи виведення набору функціональних залежностей $\{f_i(n, m_1, m_2, j_1, \eta)\}$ та визначення істинності виразу (1), операндами якого є елементи булеана 2^X, які представлені «m-арними кортежами» на основі впорядкованої базової множини X довільної потужності n при заданих початкових умовах: m_1=1; j_1={1,2}.

У даній статті пропонуються до розгляду математичні методи вирішення поставленого завдання при заданих початкових умовах: m_1=1; j_1=3.

Виклад основного матеріалу та обґрунтування отриманих результатів. Як приклад дослідження будемо використовувати булеан 2^X символьної впорядкованої за алфавітом базової множини X ={a, b, c, d, e, f, g, h} потужності n = 8, яка була описана в [9].

Множина індексів $J2^n(y^n_{m_1,j_1}) = \{J2^n_{m_2}(y^n_{m_1,j_1})\}$ елементів $y^n_{m_2,j_2}$ впорядкованої множини кортежів $Y^n_{m_2}$, що задовольняють істинності виразу (1) для елемента булана $y^n_{m_1,j_1} = y^8_{1,3}$, наведено в таблиці 1.

Таблиця 1. Визначення множини $J2^8_{m_2}(y^8_{1,3})$

m_2	η	$J2^n_{m_2}(y^n_{m_1,j_1}) = \{J2^n_{m_2,\eta}(y^n_{m_1,j_1})\}$	$nj2_{m_2,\eta}(y^n_{m_1,j_1})$	$nj2_{m_2}(y^n_{m_1,j_1}) = \binom{n-m_1}{m_2-m_1}$
2	2	$J2^8_{2,2}(y^8_{1,3}) = \{2,8\}$	$nj2_{2,2}(y^8_{1,3}) = 2$	$nj2_2(y^8_{1,3}) = \binom{7}{1} = 7$
	1	$J2^8_{2,1}(y^8_{1,3}) = 14 \div 18$	$nj2_{2,1}(y^8_{1,3}) = 5$	
3	3	$J2^8_{3,3}(y^8_{1,3}) = 1$	$nj2_{3,3}(y^8_{1,3}) = 1$	$nj2_3(y^8_{1,3}) = \binom{7}{2} = 21$
	2	$J2^8_{3,2}(y^8_{1,3}) = 7 \div 11,\ \ 22 \div 26$	$nj2_{3,2}(y^8_{1,3}) = 10$	
	1	$J2^8_{3,1}(y^8_{1,3}) = 37 \div 46$	$nj2_{3,1}(y^8_{1,3}) = 10$	
4	3	$J2^8_{4,3}(y^8_{1,3}) = 1 \div 5$	$nj2_{4,3}(y^8_{1,3}) = 5$	$nj2_4(y^8_{1,3}) = \binom{7}{3} = 35$
	2	$J2^8_{4,2}(y^8_{1,3}) = 16 \div 25,\ \ 36 \div 45$	$nj2_{4,2}(y^8_{1,3}) = 20$	
	1	$J2^8_{4,1}(y^8_{1,3}) = 56 \div 65$	$nj2_{4,1}(y^8_{1,3}) = 10$	
5	3	$J2^8_{5,3}(y^8_{1,3}) = 1 \div 10$	$nj2_{5,3}(y^8_{1,3}) = 10$	$nj2_5(y^8_{1,3}) = \binom{7}{4} = 35$
	2	$J2^8_{5,2}(y^8_{1,3}) = 21 \div 30,\ \ 36 \div 45$	$nj2_{5,2}(y^8_{1,3}) = 20$	
	1	$J2^8_{5,1}(y^8_{1,3}) = 51 \div 55$	$nj2_{5,1}(y^8_{1,3}) = 5$	
6	3	$J2^8_{6,3}(y^8_{1,3}) = 1 \div 10$	$nj2_{6,3}(y^8_{1,3}) = 10$	$nj2_6(y^8_{1,3}) = \binom{7}{5} = 21$
	2	$J2^8_{6,2}(y^8_{1,3}) = 16 \div 20,\ \ 22 \div 26$	$nj2_{6,2}(y^8_{1,3}) = 10$	
	1	$J2^8_{6,1}(y^8_{1,3}) = 28$	$nj2_{6,1}(y^8_{1,3}) = 1$	
7	3	$J2^8_{7,3}(y^8_{1,3}) = 1 \div 5$	$nj2_{7,3}(y^8_{1,3}) = 5$	$nj2_7(y^8_{1,3}) = \binom{7}{6} = 7$
	2	$J2^8_{7,2}(y^8_{1,3}) = 7,8$	$nj2_{7,2}(y^8_{1,3}) = 2$	
8	3	$J2^8_{8,3}(y^8_{1,3}) = 1$	$nj2_{8,3}(y^8_{1,3}) = 1$	$nj2_8(y^8_{1,3}) = \binom{7}{7} = 1$

де

η – розташування елемента $y^n_{m_1,j_1}$ у кортежі $y^n_{m_2,j_2}$;

$nj2_{m_2}(y^n_{m_1,j_1})$ – потужність підмножини $J2^n_{m_2}(y^n_{m_1,j_1})$ у складі множини $J2^n(y^n_{m_1,j_1})$;

$nj2_{m_2,\eta}(y^n_{m_1,j_1})$ – потужність підмножини $J2^n_{m_2,\eta}(y^n_{m_1,j_1})$ у складі множини $J2^n_{m_2}(y^n_{m_1,j_1})$.

На рисунку 1 наведено граф булеана з потужністю базової множини n = 8. Замкнуті ламані $S1,...,S4$ визначені на $J2^n(y^n_{m_1,j_1})$ (див. табл. 1). Вони окреслюють набори елементів булеана, що представлені кортежами довжини $m_2 > m_1$ та задовольняють істинності виразу (1) для елемента булеана $y^n_{m_1,j_1} = y^8_{1,3}$.

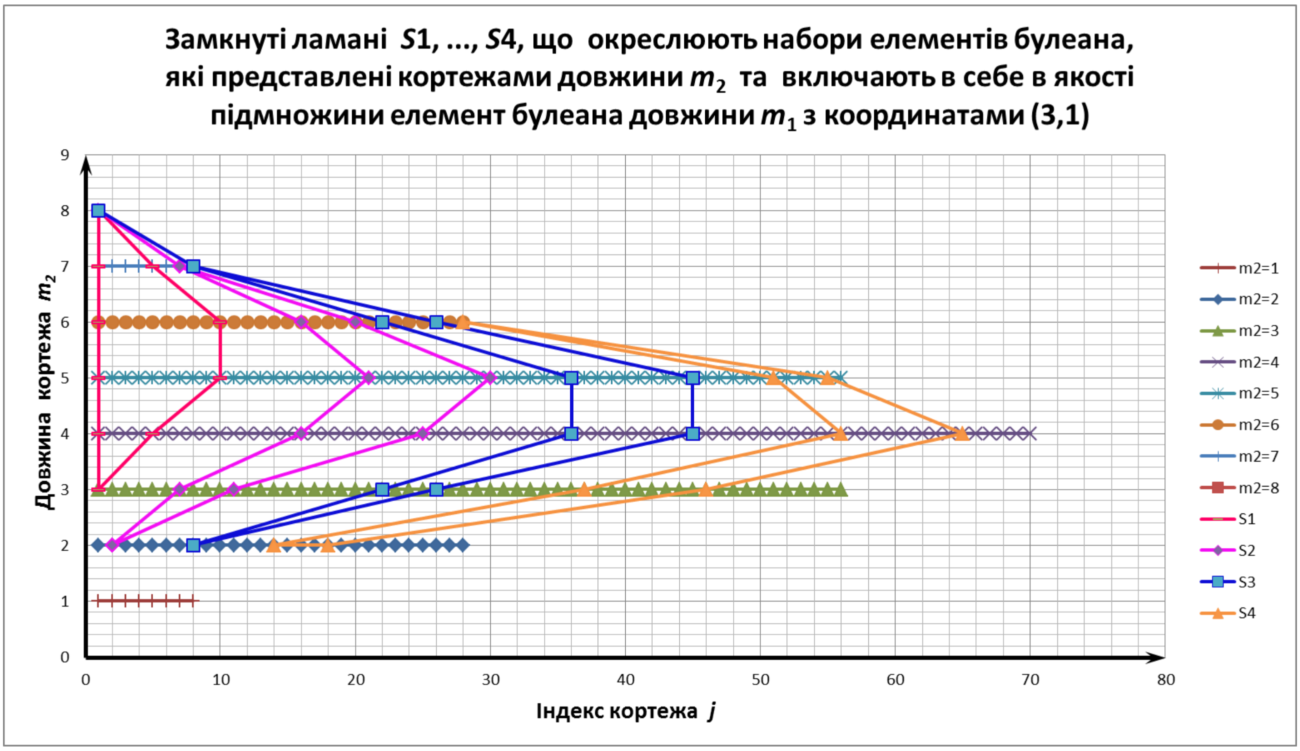

Рис. 1. Граф булеана з замкнутими ламаними для визначення множин $J2^8_{m_2}(y^8_{1,3})$

В якості наступного прикладу дослідження будемо використовувати булеан 2^X символьної впорядкованої за алфавітом базової множини $X = \{a, b, c, d, e, f, g, h, i\}$ потужності n = 9.

Множина індексів $J2^n(y^n_{m_1,j_1}) = \{J2^n_{m_2}(y^n_{m_1,j_1})\}$ елементів $y^n_{m_2,j_2}$ впорядкованої множини кортежів $Y^n_{m_2}$, що задовольняють істинності виразу (1) для елемента булана $y^n_{m_1,j_1} = y^9_{1,3}$, наведено в таблиці 2.

Таблиця 2. Визначення множини $J2^9_{m_2}(y^9_{1,3})$

m_2	η	$J2^n_{m_2}(y^n_{m_1,j_1}) = \{J2^n_{m_2,\eta}(y^n_{m_1,j_1})\}$	$nj2_{m_2,\eta}(y^n_{m_1,j_1})$	$nj2_{m_2}(y^n_{m_1,j_1}) = \binom{n-m_1}{m_2-m_1}$
2	2	$J2^9_{2,2}(y^9_{1,3}) = \{2{,}9\}$	$nj2_{2,2}(y^9_{1,3}) = 2$	$nj2_2(y^9_{1,3}) = \binom{8}{1} = 8$
	1	$J2^9_{2,1}(y^9_{1,3}) = 16 \div 21$	$nj2_{2,1}(y^9_{1,3}) = 6$	
3	3	$J2^9_{3,3}(y^9_{1,3}) = 1$	$nj2_{3,3}(y^9_{1,3}) = 1$	$nj2_3(y^9_{1,3}) = \binom{8}{2} = 28$
	2	$J2^9_{3,2}(y^9_{1,3}) = 8 \div 13,\ \ 29 \div 34)$	$nj2_{3,2}(y^9_{1,3}) = 12$	
	1	$J2^9_{3,1}(y^9_{1,3}) = 50 \div 64$	$nj2_{3,1}(y^9_{1,3}) = 15$	
4	3	$J2^9_{4,3}(y^9_{1,3}) = 1 \div 6$	$nj2_{4,3}(y^9_{1,3}) = 6$	$nj2_4(y^9_{1,3}) = \binom{8}{3} = 56$

	2	$J2^9_{4,2}(y^9_{1,3}) = 22 \div 36,\ \ 57 \div 71$	$nj2_{4,2}(y^9_{1,3}) = 30$	
	1	$J2^9_{4,1}(y^9_{1,3}) = 92 \div 111$	$nj2_{4,1}(y^9_{1,3}) = 20$	
5	3	$J2^9_{5,3}(y^9_{1,3}) = 1 \div 15$	$nj2_{5,3}(y^9_{1,3}) = 15$	$nj2_5(y^9_{1,3}) = \binom{8}{4} = 70$
	2	$J2^9_{5,2}(y^9_{1,3}) = 36 \div 55,\ \ 71 \div 90$	$nj2_{5,2}(y^9_{1,3}) = 40$	
	1	$J2^9_{5,1}(y^9_{1,3}) = 106 \div 120$	$nj2_{5,1}(y^9_{1,3}) = 15$	
6	3	$J2^9_{6,3}(y^9_{1,3}) = 1 \div 20$	$nj2_{6,3}(y^9_{1,3}) = 20$	$nj2_6(y^9_{1,3}) = \binom{8}{5} = 56$
	2	$J2^9_{6,2}(y^9_{1,3}) = 36 \div 50,\ \ 57 \div 71$	$nj2_{6,2}(y^9_{1,3}) = 30$	
	1	$J2^9_{6,1}(y^9_{1,3}) = 78 \div 83$	$nj2_{6,1}(y^9_{1,3}) = 6$	
7	3	$J2^9_{7,3}(y^9_{1,3}) = 1 \div 15$	$nj2_{7,3}(y^9_{1,3}) = 15$	$nj2_7(y^9_{1,3}) = \binom{8}{6} = 28$
	2	$J2^9_{7,2}(y^9_{1,3}) = 22 \div 27,\ \ 29 \div 34$	$nj2_{7,2}(y^9_{1,3}) = 12$	
	1	$J2^9_{7,1}(y^8_{1,3}) = 36$	$nj2_{7,1}(y^9_{1,3}) = 1$	
8	3	$J2^9_{8,3}(y^9_{1,3}) = 1 \div 6$	$nj2_{8,3}(y^9_{1,3}) = 6$	$nj2_8(y^9_{1,3}) = \binom{8}{7} = 8$
	2	$J2^9_{8,2}(y^9_{1,3}) = \{8{,}9\}$	$nj2_{8,2}(y^9_{1,3}) = 2$	
9	3	$J2^9_{9,3}(y^9_{1,3}) = 1$	$nj2_{9,3}(y^9_{1,3}) = 1$	$nj2_9(y^9_{1,3}) = \binom{8}{8} = 1$

На основі аналізу таблиць 1 і 2 була сформована таблиця 3, що описує метод A_{10} визначення істинності виразу (1) для елемента $y^n_{1,3}$ при довільному *n*. Для доказу істинності результату був застосований метод математичної індукції.

Таблиця 3. Метод A_{10} визначення множин $J2^n_{m_2}(y^n_{1,3}) = f(n, m_2, \eta)$

m_2	η	$J2^n_{m_2,\eta}(y^n_{m_1,j_1})$	$nj2_{m_2,\eta}(y^n_{m_1,j_1})$	$nj2_{m_2}(y^n_{m_1,j_1})$
2 ($m_2 < j_1$)	2	$J2^n_{2,2}(y^n_{1,3}) =$ $\{j2_{2,2,1}(y^n_{1,3}), j2_{2,2,last}(y^n_{1,3})\};$ $(\eta \le j_1)?$ $(j2_{2,2,1}(y^n_{1,3}) = nj2_{2,2}(y^n_{1,3})) =$ $= j_1 - 1;$ $(nj2_{2,2}(y^n_{1,3}) > 1)?$ $(j2_{2,2,last}(y^n_{1,3}) = n);$	$(\eta \le j_1)?$ $(nj2_{2,2}(y^n_{1,3}) = j_1 - 1 =$ $= 2;$	$nj2_2(y^n_{1,3}) =$ $= \binom{n - m_1}{m_2 - m_1} =$ $= n - 1;$
	1	$J2^n_{2,1}(y^n_{1,3}) =$ $= \{j2_{2,1,1}(y^n_{1,3}), \ldots, j2_{2,1,last}(y^n_{1,3})\};$ $j2_{2,1,1}(y^n_{1,3}) = (j_1 - 1) * (n - 1);$ $(nj2_{2,1}(y^n_{1,3}) > 1)?$ $j2_{2,1,last}(y^n_{1,3}) = j2_{2,1,1}(y^n_{1,3}) +$ $+ nj2_{2,1} - 1 =$ $= j2_{2,1,1}(y^n_{1,3}) + n - 4 =$ $= (j_1 - 1) * (n - 1) + n - 4;$	$nj2_{2,1}(y^n_{1,3}) =$ $= n - j_1 = n - 3;$	

3 (m_2==j_1)	3	$(m_2 == \eta == j_1)?$ $(J2^n_{3,3}(y^n_{1,3}) =$ $j2_{3,3,1}(y^n_{1,3}) = 1);$	$(m_2 == \eta == j_1)?$ $nj2_{3,3}(y^n_{1,3}) = 1;$	$nj2_3(y^n_{1,3}) =$ $= \binom{n-1}{2}$
	2	$J2^n_{3,2}(y^n_{1,3}) = \{J2^n_{3,2,g}(y^n_{1,3})\} =$ $= \{J2^n_{3,2,1}(y^n_{1,3}), J2^n_{3,2,2}(y^n_{1,3})\};$ $J2^n_{3,2,1}(y^n_{1,3}) = \{j2_{3,2,1,1}(y^n_{1,3}),...,$ $j2_{3,2,1,last}(y^n_{1,3})\};$ $J2^n_{3,2,2}(y^n_{1,3}) =$ $\{j2_{3,2,2,1}(y^n_{1,3}),...,$ $j2_{3,2,2,last}(y^n_{1,3})\};$ $j2_{3,2,1,1}(y^n_{1,3}) = n-1;$ $(nj2_{3,2}(y^n_{1,3}) > 1)?$ $j2_{3,2,1,last}(y^n_{1,3}) =$ $= j2_{3,2,1,1}(y^n_{1,3}) +$ $+ nj2_{3,2,1}(y^n_{1,3}) - 1 =$ $= j2_{3,2,1,1}(y^n_{1,3}) + n - 4 =$ $= 2*n-5;$ $j2_{3,2,2,1}(y^n_{1,3}) = nj2_3(y^n_{1,3}) + 1 =$ $= \binom{n-1}{2} + 1;$ $(nj2_{3,2,2}(y^n_{1,3}) > 1)?$ $j2_{3,2,2,last}(y^n_{1,3}) =$ $= j2_{3,2,2,1}(y^n_{1,3}) + nj2_{3,2,2}(y^n_{1,3})$ $-1 = j2_{3,2,2,1}(y^n_{1,3}) + n - 4 =$ $= \binom{n-1}{2} + n - 3;$	$nj2_{3,2}(y^n_{1,3}) = 2*$ $nj2_{2,1}(y^n_{1,3});$ $nj2_{m_2,\eta,g}(y^n_{1,3}) =$ $= nj2_{3,2,1}(y^n_{1,3}) =$ $= nj2_{3,2,2}(y^n_{1,3}) =$ $= nj2_{2,1}(y^n_{1,3}) =$ $= n - j_1 = n - 3;$	
	1	$J2^n_{3,1}(y^n_{1,3}) = \{j2_{3,1,1}(y^n_{1,3}),...,$ $j2_{3,1,last}(y^n_{1,3})\};$ $j2_{3,1,1}(y^n_{1,3}) =$ $= \binom{n}{m_2} - \binom{n-j_1+1}{m_2} + 1;$ $(nj2_{3,1}(y^n_{1,3}) > 1)?$ $j2_{3,1,last}(y^n_{1,3}) = j2_{3,1,1}(y^n_{1,3}) +$ $+ nj2_{3,1}(y^n_{1,3}) - 1 =$ $= \binom{n}{m_2} - \binom{n-j_1}{m_2};$	$nj2_{3,1}(y^n_{1,3}) =$ $= \binom{n-j_1+1}{m_2} -$ $- \binom{n-j_1}{m_2};$	

...	**...**	**...**	**...**	**...**
$m_2=\mu$, ($j_1<m_2$, $m_2<n$-1)	3	$J2^n_{\mu,3}(y^n_{1,3})=\{j2_{\mu,3,1}(y^n_{1,3}),...,$ $j2_{\mu,3,last}(y^n_{1,3})\};$ $j2_{\mu,3,1}(y^n_{1,3})=1;$ $(nj2_{\mu,3}(y^n_{1,3})>1)?$ $j2_{\mu,3,last}(y^n_{1,3})=$ $nj2_{\mu,3}(y^n_{1,3})=$ $=\binom{n-j_1}{\mu-m_1-\eta+1};$	$nj2_{\mu,3}(y^n_{1,3})=$ $=\binom{n-j_1}{\mu-m_1-\eta+1};$	$nj2_{\mu}(y^n_{1,3})=$ $=\binom{n-m_1}{\mu-m_1}$
	2	$J2^n_{\mu,2}(y^n_{1,3})=\{J2^n_{\mu,2,g}(y^n_{1,3})\}=$ $=\{J2^n_{\mu,2,1}(y^n_{1,3}),J2^n_{\mu,2,2}(y^n_{1,3})\}=$ $=\{j2_{\mu,2,1,1}(y^n_{1,3}),...,j2_{\mu,2,1,last}(y^n_{1,3}),$ $j2_{\mu,2,2,1}(y^n_{1,3}),...,j2_{\mu,2,2,last}(y^n_{1,3})\};$ $j2_{\mu,2,1,1}(y^n_{1,3})=nj2_{\mu,3}(y^n_{1,3})+$ $+nj2_{\mu,2,1}(y^n_{1,3})+1;$ $(nj2_{\mu,2,1}(y^n_{1,3})>1)?$ $j2_{\mu,2,1,last}(y^n_{1,3})=j2_{\mu,2,1,1}(y^n_{1,3})+$ $+nj2_{\mu,2,1}(y^n_{1,3})-1;$ $j2_{\mu,2,2,1}(y^n_{1,3})=nj2_{\mu}(y^n_{1,3})+1;$ $(nj2_{\mu,2,2}(y^n_{1,3})>1)?$ $j2_{\mu,2,2,last}(y^n_{1,3})=j2_{\mu,2,2,1}(y^n_{1,3})+$ $+nj2_{\mu,2,2}(y^n_{1,3})-1;$	$nj2_{\mu,2,g}(y^n_{1,3})=$ $nj2_{\mu,2,1}(y^n_{1,3})=$ $=nj2_{\mu,2,2}(y^n_{1,3})=$ $=\binom{n-j_1}{\mu-m_1-\eta+1};$ $nj2_{\mu,2}(y^n_{1,3})=$ $=2*nj2_{\mu,2,g}(y^n_{1,3});$	
	1	$J2^n_{\mu,1}(y^n_{1,3})=\{j2_{\mu,1,1}(y^n_{1,3}),...,$ $j2_{\mu,1,last}(y^n_{1,3})\};$ $j2_{\mu,1,1}(y^n_{1,3})=$ $=\binom{n}{m_2}-\binom{n-j_1+1}{m_2}+1;$ $(nj2_{\mu,1}(y^n_{1,3})>1)?$ $j2_{\mu,1,last}(y^n_{1,3})=j2_{\mu,1,1}(y^n_{1,3})+$ $+nj2_{\mu,1}(y^n_{1,3})-1;$	$nj2_{\mu,1}(y^n_{1,3})=$ $=\binom{n-j_1}{\mu-m_1-\eta+1};$	
...	**...**	**...**	**...**	**...**
n-1	3	$J2^n_{n-1,3}(y^n_{1,3})=\{j2_{n-1,3,1}(y^n_{1,3}),...,$ $j2_{n-1,3,last}(y^n_{1,3})\};$ $j2_{n-1,3,1}(y^n_{1,3})=1;$ $(nj2_{n-1,3}(y^n_{1,3})>1)?$	$nj2_{n-1,3}(y^n_{1,3})=$ $=\binom{n-j_1}{n-1-m_1-\eta+1}=$ $=n-j_1;$	$nj2_{n-1}(y^n_{1,3})=$ $=\binom{n-m_1}{n-1-m_1}=$ $=n-1;$

		$j2_{n-1,3,last}(y_{1,3}^n)=$ $nj2_{n-1,3}(y_{1,3}^n)=$ $=\begin{pmatrix} n-j_1 \\ n-1-m_1-\eta+1 \end{pmatrix}=n-j_1;$		
	2	$j2_{n-1,2,1}(y_{1,3}^n)=n-1;$ $j2_{n-1,2,last}(y_{1,3}^n)=n;$	$nj2_{n-1,2}(y_{1,3}^n)=$ $j_1-1=2;$	
n	3	$j2_{n,1,1}(y_{1,3}^n)=1;$	$nj2_{n,3}(y_{1,3}^n)=$ $nj2_n(y_{1,3}^n)=1;$	$nj2_n(y_{1,3}^n)=1$

де

$j2_{m_2,\eta,1}$ – мінімальний порядковий номер j_2 елемента y_{m_2,j_2}^n у множині $Y_{m_2}^n$, для якого істинний вираз (1) у разі, коли елемент y_{m_1,j_1}^n розташований на η-ому місці у кортежі y_{m_2,j_2}^n;

$j2_{m_2,\eta,last}$ – максимальний порядковий номер j_2 елемента y_{m_2,j_2}^n у множині $Y_{m_2}^n$, для якого істинний вираз (1) у разі, коли елемент y_{m_1,j_1}^n розташований на η-ому місці у кортежі y_{m_2,j_2}^n;

$j2_{m_2,\eta,g,1}$ – мінімальний порядковий номер j_2 елемента y_{m_2,j_2}^n у множині $Y_{m_2}^n$, для якого істинний вираз (1) у разі, коли елемент y_{m_1,j_1}^n розташований в g-ої підгрупі на η-ому місці у кортежі y_{m_2,j_2}^n;

$j2_{m_2,\eta,g,last}$ – максимальний порядковий номер j_2 елемента y_{m_2,j_2}^n у множині $Y_{m_2}^n$, для якого істинний вираз (1) у разі, коли елемент y_{m_1,j_1}^n розташований в g-ої підгрупі на η-ому місці у кортежі y_{m_2,j_2}^n.

Приклад. Для булеана 2^X, сформованого на основі символьної впорядкованої за алфавітом базової множини $X=\{a, b, c, d, e, f, g, h, i, j\}$ потужності $n = 10$ визначити істинність виразу $y_{1,3}^{10} \subset y_{5,187}^{10}$.

Дано: n=10, m_1=1, j_1=3, m_2=5, j_2=187.

Виконаємо перевірку коректності вхідних даних:

$$\begin{cases} m_1 < m_2 \le n, \\ j_2 \le \begin{pmatrix} n \\ m_2 \end{pmatrix} \end{cases} \Rightarrow \begin{cases} 1 < 5 \le 10, \\ 187 \le 252 \end{cases}$$

Вхідні дані коректні.

Варіант рішення 1. Визначимо елементи булеана: $y_{1,3}^{10}=c$ та $y_{5,187}^{10}=(b,e,f,i,j)$ із застосуванням методу A_1 [8]. Виконаємо метод $A_6: y_{m,j}^n = y_{m_1,j_1}^n \bigcap y_{m_2,j_2}^n$, описаний в [10].

В результаті маємо:

$$y_{1,3}^{10n} \bigcap y_{5,187}^{10} = \varnothing \Rightarrow y_{1,3}^{10} \not\subset y_{5,187}^{10}$$

Варіант рішення 2. З таблиці 3 вибираємо єдиний рядок, для якого умова стає істинною при підстановці вхідних даних. В нашому випадку, істинною стає умова $j_1 < m_2 < n-1$.

Дійсно, при підстановці вхідних даних маємо: 3 < 5 < 9.

$$nj2_{\mu}(y_{1,3}^{n}) = nj2_{5}(y_{1,3}^{10}) = \begin{pmatrix} n-m_1 \\ \mu-m_1 \end{pmatrix} = \begin{pmatrix} 10-1 \\ 5-1 \end{pmatrix} = \begin{pmatrix} 9 \\ 4 \end{pmatrix} = 126 \quad (3)$$

$$nj2_{\mu,3}(y_{1,3}^{n}) = nj2_{5,3}(y_{1,3}^{10}) = \begin{pmatrix} n-j_1 \\ \mu-m_1-\eta+1 \end{pmatrix} = \begin{pmatrix} 10-3 \\ 5-1-3+1 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \end{pmatrix} = 21 \quad (4)$$

$$nj2_{\mu,2,g}(y_{1,3}^{n}) = nj2_{5,2,g}(y_{1,3}^{10}) = \begin{pmatrix} n-j_1 \\ \mu-m_1-\eta+1 \end{pmatrix} = \begin{pmatrix} 10-3 \\ 5-1-2+1 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \end{pmatrix} = 35 \quad (5)$$

$$nj2_{\mu,2}(y_{1,3}^{n}) = nj2_{5,2}(y_{1,3}^{10}) = 2*nj2_{5,2,g}(y_{1,3}^{10}) = 2*35 = 70 \quad (6)$$

$$nj2_{\mu,1}(y_{1,3}^{n}) = nj2_{5,1}(y_{1,3}^{10}) = \begin{pmatrix} n-j_1 \\ \mu-m_1-\eta+1 \end{pmatrix} = \begin{pmatrix} 10-3 \\ 5-1-1+1 \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \end{pmatrix} = 35 \quad (7)$$

$$j2_{\mu,3,1}(y_{1,3}^{n}) = j2_{5,3,1}(y_{1,3}^{10}) = 1 \quad (8)$$

$$j2_{\mu,3,last}(y_{1,3}^{n}) = j2_{5,3,last}(y_{1,3}^{10}) = nj2_{5,3}(y_{1,3}^{10}) = 21 \quad (9)$$

$$j2_{\mu,2,1,1}(y_{1,3}^{n}) = j2_{5,2,1,1}(y_{1,3}^{10}) = nj2_{\mu,3}(y_{1,3}^{n}) + nj2_{\mu,2,1}(y_{1,3}^{n}) + 1 = 21+35+1 = 57 \quad (10)$$

$$j2_{\mu,2,1,last}(y_{1,3}^{n}) = j2_{5,2,1,last}(y_{1,3}^{10}) = j2_{\mu,2,1,1}(y_{1,3}^{n}) + nj2_{\mu,2,1}(y_{1,3}^{n}) - 1 = 57+35-1 = 91 \quad (11)$$

$$j2_{\mu,2,2,1}(y_{1,3}^{n}) = j2_{5,2,2,1}(y_{1,3}^{10}) = nj2_{\mu}(y_{1,3}^{n}) + 1 = 126+1 = 127 \quad (12)$$

$$j2_{\mu,2,2,last}(y_{1,3}^{n}) = j2_{5,2,2,last}(y_{1,3}^{10}) = j2_{\mu,2,2,1}(y_{1,3}^{n}) + nj2_{\mu,2,2}(y_{1,3}^{n}) - 1 = 127+35-1 = 161 \quad (13)$$

$$j2_{\mu,1,1}(y_{1,3}^{n}) = j2_{5,1,1}(y_{1,3}^{10}) = \begin{pmatrix} n \\ m_2 \end{pmatrix} - \begin{pmatrix} n-j_1+1 \\ m_2 \end{pmatrix} + 1 = \begin{pmatrix} 10 \\ 5 \end{pmatrix} - \begin{pmatrix} 10-3+1 \\ 5 \end{pmatrix} + 1 = 252-56+1 = 197 \quad (14)$$

$$j2_{\mu,1,last}(y_{1,3}^{n}) = j2_{5,1,last}(y_{1,3}^{10}) = j2_{\mu,1,1}(y_{1,3}^{n}) + nj2_{\mu,1}(y_{1,3}^{n}) - 1 = 197+35-1 = 231 \quad (15)$$

$$J2_{m_2}^{n}(y_{m_1,j_1}^{n}) = J2_{5}^{10}(y_{1,3}^{10}) = \{1,...,21, \quad 57,...,91, \quad 127,...,161, \quad 197,...,231\} \quad (16)$$

$$j_2 \notin J2_{m_2}^{n}(y_{m_1,j_1}^{n}) \quad (17)$$

Відповідь: $y_{1,3}^{10} \not\subset y_{5,187}^{10}$.

При вирішенні даного прикладу множина $J2_{m_2}^{n}(y_{m_1,j_1}^{n})$ була визначена повністю. При цьому було виконано максимальну кількість обчислень, в результаті чого було розраховано 13 параметрів. Однак, для отримання відповіді немає необхідності повністю визначати множину $J2_{m_2}^{n}(y_{m_1,j_1}^{n})$. У даному прикладі досить було розрахувати 5 параметрів: $nj2_{\mu}(y_{1,3}^{n})$, $nj2_{\mu,2,g}(y_{1,3}^{n})$, $j2_{\mu,2,2,1}(y_{1,3}^{n})$, $j2_{\mu,2,2,last}(y_{1,3}^{n})$, $j2_{\mu,1,1}(y_{1,3}^{n})$. При цьому час виконання алгоритму зменшується у 2.6 рази. Оптимізація алгоритму обчислень полягає в тому, щоб правильно обрати необхідні для розрахунку параметри в залежності від початкових умов.

Як бачимо з таблиці 3, для елемента $y_{1,3}^{n}$ в загальному випадку маємо r=4 групи параметрів визначення індексу j_2:

$$J2_{m_2}^{n}(y_{1,3}^{n}) = \{ j2_{m_2,3,1},..., j2_{m_2,3,last}, \quad j2_{m_2,2,1,1},..., j2_{m_2,2,1,last},$$
$$j2_{m_2,2,2,1},..., j2_{m_2,2,2,last}, \quad j2_{m_2,1,1},..., j2_{m_2,1,last}\};$$

Індекс $j_2 max$ останнього елемента множини $Y_{m_2}^{n}$ розраховується наступним чином:

$$j_2 \max = \begin{pmatrix} n \\ m_2 \end{pmatrix}$$

Таким чином, послідовність індексів $j_2 \in [1; j_2\max]$ множини $Y^n_{m_2}$ можна також розбити на r інтервалів $h = j_2\max / r$:

$$\{1,\ldots,h,\quad h+1,\ldots,2*h,\quad \ldots,\quad (i\text{-}1)*h+1,\ldots,i*h,\quad \ldots,\quad (r-1)*h+1,\ldots,r*h\ \}$$

Обчислюючи номер i інтервалу входження індексу j_2 згідно з умовою $(i\text{-}1)*h+1 \le j_2 \le i*h$, вибираємо групу початкових параметрів для розрахунку.

Визначимо початкові параметри розрахунку для нашого прикладу:

$$j_2\max = \binom{n}{m_2} = \binom{10}{5} = 252$$

$$h = j_2\max / r = 252 / 4 = 63$$

Послідовність індексів j_2 з розбиттям на r=4 групи має вигляд:

$$\{1,\ldots,63,\ 64,\ldots,126,\ \ 127,\ldots,\ 189,\ \ 190,\ldots,252\}$$

Для j_2 =187 обчислюємо номер інтервалу входження індексу i=3 $(127 \le 187 \le 189)$. Таким чином, для нашого прикладу вибираємо i=3 групу початкових параметрів для розрахунку: $j2_{m_2,2,2,1}$, $j2_{m_2,2,2,last}$.

В результаті аналізу таблиці 3 були визначені деякі додаткові до [10] властивості СОД типу «m-арні кортежі на основі ВМДП».

Властивість 7. Якщо елемент $y^n_{m_1,j_1}$ задовольняє умовам

$$(m_1 == 1)\ \&\ (1 < j_1 < n)\ \&\ (m_2 \ge j_1),$$

то вираз (1) завжди буде істинним для перших (j_2=1) елементів множин $Y^n_{m_2}$:

$$((m_1 == 1)\ \&(1 < j_1 < n)\ \&\ (m_2 \ge j_1))\ \ ?\ \ j_2 = 1.$$

Властивість 8. Якщо виконується умова

$$(m_2 == 2)\ \&\ (m_1 == 1)\ \&(1 < \ j_1 < n)\ ,$$

то вираз (1) завжди буде істинним для елементів множин $Y^n_{m_2}$ з індексом $j_2 = j_1 - 1$:

$$((m_2 == 2)\ \&\ (m_1 == 1)\ \&(1 < \ j_1 < n))\ ?\ j_2 = \ j_1 - 1.$$

Властивість 9. Якщо виконується умова

$$(m_2 == n - 1)\ \&\ (m_1 == 1)\ \&(1 < \ j_1 < n),$$

то вираз (1) завжди буде помилковим тільки для єдиного елемента множини Y^n_{n-1} з індексом $j_2 = n - j_1 + 1$:

$$((m_2 == n - 1)\ \&\ (m_1 == 1)\ \&(1 < \ j_1 < n))\ ?\ j_2 \ne \ n - j_1 + 1.$$

Для оцінки продуктивності методів роботи з СОД типу «m-арні кортежі на основі ВМДП» будемо використовувати асимптотичну оцінку $O(f(n))$ швидкості зростання кількості операцій $f(n)$ при збільшенні n.

В [8, 10] були розраховані оцінки часу виконання методів A_1 та A_6. У результаті був зроблений висновок, що методи A_1 та A_6 є алгоритмами кубічного часу виконання:

$$O_6(f_6(n)) = O_6(n^3)$$

За аналогією з [9], розглянутий в даній статті метод A_{10} є алгоритмом лінійного часу виконання:

$$O_{10}(f_{10}(n)) = O_{10}(n)$$

Враховуючи, що сучасні процесори працюють з тактовою частотою 3 Ггц і що в середньому МАС операція (МАС — Multiplier/Accumulator операція множення з накопиченням типу) виконується за 4 такти, приймаємо в якості розрахункового часу виконання однієї операції для розглянутих методів:

$$\tau = 4/(3*10^9 \text{ Гц}) = 1{,}33*10^{-9}\text{с} = 1{,}33\text{нс}$$

Складемо таблицю 4 зміни (з зростанням n) кількості операцій та оцінок часу їх виконання для методів лінійного та кубічного часу виконання.

Таблиця 4. Кількість операцій і оцінка часу виконання методів

n	n^3	$O(n)$, c	$O(n^3)$, c
10	$1{,}0*10^3$	$1{,}33*10^{-8}$	$1{,}33*10^{-8}$
50	$1{,}25*10^3$	$6{,}65*10^{-8}$	$1{,}6625*10^{-6}$
100	$1{,}0*10^6$	$1{,}33*10^{-7}$	$1{,}33*10^{-3}$
500	$1{,}25*10^8$	$6{,}65*10^{-7}$	$1{,}6625*10^{-1}$
1 000	$1{,}0*10^9$	$1{,}33*10^{-6}$	1,33
5 000	$1{,}25*10^{11}$	$6{,}65*10^{-6}$	$1{,}6625*10^2 \approx 2{,}77$с
10 000	$1{,}0*10^{12}$	$1{,}33*10^{-5}$	$1{,}33*10^3 \approx 22$ мин

Описаний в статті метод A_{10} роботи із СОД типу «m-арні кортежі на основі ВМДП» у порівнянні з методами A_2 та A_6 змінює функціональну залежність оцінки часу виконання методу від кількості вхідних даних n з кубічної $O(n^3)$ на лінійну $O(n)$, що дозволяє значно (на кілька порядків) зменшити час обробки даних. Різниця в оцінці часу виконання методів особливо помітна при $n \geq 1000$ (див. табл. 4).

Висновки та перспективи подальшого дослідження. На даному етапі досліджень були визначені деякі властивості і виведені деякі функціональні залежності для СОД типу «m-арні кортежі на основі ВМДП» для приватного випадку при заданих початкових умовах: $m_1=1$ & $j_1 \in [1; 3]$. У перспективі передбачається продовжити дослідження в цьому напрямку з визначенням нових властивостей і виведенням нових функціональних залежностей між m-арними кортежами при будь-яких можливих початкових умовах: $m_1 \in [1; n-1]$ & $j_1 \in [1; n]$.

СОД типу «m-арні кортежі на основі ВМДП» – це булеан, упорядкований (в нашому випадку) за правобічним перебором елементів базової множини X потужності n від нижньої межі можливої зміни значення індексу для кожного елемента кортежу до верхньої межі. Розглянуті в статті властивості та метод роботи з СОД типу «m-арні кортежі на основі ВМДП» виникають із правил формування впорядкованих за зростанням значень елементів СОД.

Наукова новизна одержаних результатів полягає у виведенні набору функціональних залежностей між m-арними кортежами на основі їх розташування у структурі, яка визначається парою індексів (j, m). Це дозволяє отримувати результати деяких операцій над елементами СОД на основі цих функціональних залежностей, а не в результаті виконання громіздких обчислювальних алгоритмів. Використання описаного у статті методу змінює функціональну залежність оцінки часу отримання результатів з кубічної на $O(n^3)$ на лінійну $O(n)$.

Практичне значення одержаних в роботі результатів обумовлюється тим, що даний підхід дозволяє мінімізувати тимчасові і обчислювальні ресурси, задіяні при обробці «великих даних», до масштабу реального часу.

Список бібліографічних посилань

- Min Chen. Big Data. Related Technologies, Challenges, and Future Prospects / Min Chen, Shiwen Mao, Yin Zhang, Victor C.M. Leung. Spinger. – 2014. – 100 p.
- Петрова С.Ю. Проблема навигации в больших данных / С.Ю. Петрова. Экспериментальные и теоретические исследования в современной науке: сб. ст. по матер. III междунар. науч.-практ. конф. – № 3(3). – Новосибирск: СибАК, 2017. – С. 5-8.
- Соловьев С. Ю. Математические методы и принципы построения автоматизированных систем инженерии знаний: дисс. доктора техн. наук: 05.13.16 / Соловьев Сергей Юрьевич. – Тверь, 1996. – 272 с.

- Тоценко В. Г. Экспертная система диагностирования по неструктурированным признакам / В. Г. Тоценко, Е. А. Петрова, А. А. Чернявская. Реєстрація, зберігання і обробка даних. – 2005. – т. 7. – № 2. – С. 94-103.
- MacGregor J. Monitoring, Fault Diagnosis, Fault-Tolerant Control and Optimization / J. MacGregor, A. Cinar // Data Driven Methods. Computers & Chemical Engineering. – 2012. – Vol. 47. – P. 111-120.
- Колчин А. В. Обзор современных систем и методов верификации формальных моделей / А.В. Колчин, А.А. Летичевский, С.В. Потиенко, В.С. Песчаненко. Проблеми програмування. – 2012. – № 4. – С. 75-88.
- Яцкив И. В. Проблема валидации имитационной модели и ее возможные решения / И. В. Яцкив. Материалы конференции. ИММОД. – 2003. – С. 211-217.
- Syrotkina O. The Application of Specialized Data Structures for SCADA Diagnostics / O. Syrotkina. System technologies. Regional interuniversity collection of scientific papers. – 2015. – Vol. 4. – P. 72−81.
- Syrotkina O. Graphical and Analytical Methods for Processing «Big Data» Based on the Analysis of Their Properties / O. Syrotkina, M. Alekseyev, I. Udovik . System technologies. Regional interuniversity collection of scientific papers. – 2019. – Vol. 4. – P. 78−90.
- Syrotkina, O., Alekseyev, M., Asotskyi, V., Udovik I. Analysis of How the Properties of Structured Data Can Influence the Way These Data are Processed / O. Syrotkina, M. Alekseyev, V. Asotskyi, I. Udovik. Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu. – 2019. – 3(171). – P. 119−128.