

**М. М. Байас**<sup>1, 2, 3</sup>  
**В. М. Дубовой**<sup>1</sup>  
**М. Е. Дуда**<sup>1</sup>

## ДЕКОМПОЗИЦІЯ ЗАДАЧ УПРАВЛІННЯ МЕТОДОМ КЛАСТЕРИЗАЦІЇ

<sup>1</sup>Вінницький національний технічний університет, Україна;

<sup>2</sup>Universidad Estatal Península de Santa Elena, Еквадор;

<sup>3</sup>SENESCYT «Міністерство вищої освіти, науки, технологій і інновацій», Еквадор

*Розглянуто питання побудови критерію та алгоритму кластеризації для розв'язування задачі декомпозиції в проектуванні структури управління технологічним процесом. Запропоновано алгоритм на основі метода найближчого сусіда.*

**Ключові слова:** алгоритм кластеризації, координація задач, декомпозиція задач, витрати на обмін інформацією.

### Вступ

В сучасних умовах відбуваються об'єктивні зміни у функціонуванні промислових підприємств, що пов'язано зі зростанням їх розмірів і складності, виникає необхідність модернізації технологічних процесів, координації елементів виробництва, вдосконалення структури систем управління. Проблема науково обґрунтованої побудови систем ієрархічного управління стає актуальною в умовах безперервної адаптації сучасних виробництв до зовнішніх змін. Серед найзначущих напрямків в галузі побудови ієрархічних структур різної природи слід відмітити теорію управління ієрархічними багаторівневими системами, основи якої закладені М. Месаровичем [1]. Істотні результати отримані в роботах А. П. Афанасьєва, С. Л. Попцова, А. А. Вороніна, С. П. Мішина, В. Н. Буркова, Д. А. Новикова, М. В. Губко, М: J. Beckmann, P. Drucker, D. Мако, R; E. Miles, H. Mintzberg, S. Rosen, Y. Takahara, O. E. Вілліамсон, і ряду інших вчених [2—4]. В основу більшості цих робіт покладено класичні методи математичного програмування, теорії ігор, теорії управління динамічними системами.

Можна виділити низку основних задач функціонування та управління ієрархічними системами, розв'язання яких висвітлено у роботах [5—8], зокрема: задача декомпозиції системи, задача координації системи, задача врахування невизначеності параметрів і змінних в ієрархічних системах прийняття рішень.

Структура багаторівневої ієрархічної системи, безсумнівно, впливає на її найважливіші характеристики, такі як оперативність, надійність, вартість. Тому визначення оптимальної структури є актуальною задачею під час розробки складних систем управління технологічними процесами. Завдання вибору структури системи управління технологічним процесом ґрунтується на декомпозиції множини задач управління з подальшою їх координацією з урахуванням територіальної та функціональної розподіленості.

Концепція розподіленого управління технологічним процесом визначає ієрархічну структуру системи управління. Основними характеристиками структури, що підлягають вибору, є число рівнів і число підсистем на кожному рівні.

В існуючих розв'язках задачі побудови ієрархічних структур систем не враховується вартість обміну інформацією, що суттєво впливає на вибір структури системи управління, і тому метою статті є розв'язування задачі побудови ієрархічних структур систем управління на основі кластеризації задач управління з урахуванням вартості обміну інформацією.

### Постановка задачі

Нехай управління розподіленим технологічним процесом визначається як задача координації

підпроцесів і потребує ресурсів загальною вартістю  $r_i$  та обсягів інформації  $I_{ij}$ , забезпечення якої вимагає витрат  $c_{ij}$ , де  $i$  — керований підпроцес,  $j$  — підпроцес, з яким необхідна координація управління.

Таким чином, витрати на управління підпроцесом

$$q_i = r_i + \sum_{\substack{j=1 \\ j \neq i}}^n c_{ij},$$

де  $n$  — кількість керованих підпроцесів (задач управління).

Необхідно виконати декомпозицію множини задач так, щоб витрати на управління були мінімальними.

### Розв'язання задачі

Вартість управління підпроцесом  $r_i$  залежить від його складності  $S_i$ . Складність підпроцесу можна оцінити за допомогою алгоритмічної моделі [9]. З урахуванням цього витрати на управління підпроцесом можна визначити так:

$$q_i = r_i + c_i = f_1(S_i) + f_2\left(\sum_{v \in A_i} I_{iv}\right), \quad (1)$$

де  $c_i$  — витрати на обмін інформацією між задачами управління.

У зв'язку з тим, що управління різними процесами характеризується різними залежностями витрат від складності і кількості інформації, надзвичайно важко виконати аналіз рівняння (1) у загальному випадку. У більшості випадків достатньо проаналізувати лінеаризовану залежність, при цьому для окремих підпроцесів необхідно тільки визначити коефіцієнти рівняння (2):

$$q_i = k_1 S_i + k_2 \sum_{v \in A_i} I_{iv}, \quad (2)$$

де  $A_i$  — множина інформаційних зв'язків між задачами управління операціями  $i$ -го підпроцесу, задана у вигляді матриці суміжності,  $A_i \subset A_0$ ,  $A_0$  — множина інформаційних зв'язків усіх задач управління технологічного процесу;  $I_{iv}$  — інформаційний потік в  $v$ -му інформаційному зв'язку;  $k_1$ ,  $k_2$  — коефіцієнти.

Якщо управління підпроцесом забезпечується локальною підсистемою, то витрати на забезпечення отримання зовнішньої інформації  $c_{ext}$  значно вище витрат на забезпечення внутрішніх потоків  $c_{int}$ . Таким чином, кластеризація задач управління приводить до зменшення витрат на управління та координацію. Процес кластеризації ілюструє рис. 1.

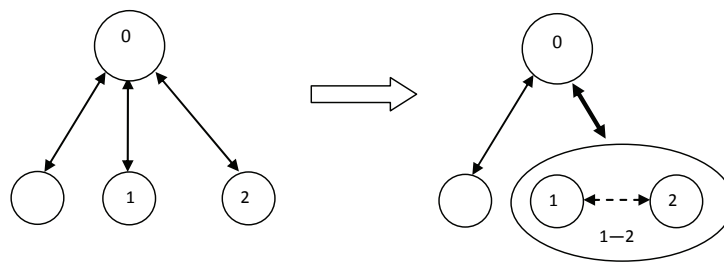


Рис. 1. Процес кластеризації задач управління

Відповідно, витрати на забезпечення необхідного обсягу інформації

$$c_i = c_{i\text{int}} + c_{i\text{ext}} = \sum_{\substack{j \neq i \\ j \in \{A_i\}}}^{n_i} C_{ij} = f_3\left(\sum_{\substack{j \neq i \\ j \in \{A_i\}}}^{n_i} I_{ij}\right); \quad (3)$$

$$n_i = \text{card}[A_i] = \text{card}[A_{i\text{int}}] + \text{card}[A_{i\text{ext}}]$$

або у лінеаризованому вигляді:

$$c_i = k_{3ext} \sum_{j \in \{A_i/A_{int}\}} I_{ij} + k_{3int} \sum_{j \in A_{int}} I_{ij}, \quad (4)$$

де «/» — операція доповнення.

Очевидно, між коефіцієнтами (2) і (4) виконуються співвідношення

$$k_2 = \frac{c_i}{\sum_{v \in A_i} I_{iv}};$$

$$k_{3ext} \geq k_2 \geq k_{3int}.$$

Витрати на управління процесами складаються з двох компонент: витрати на управління зовнішніми підпроцесами по відношенню до підсистеми, яка керується однією локальною системою управління (ЛСУ),  $q_{ext}$ , і витрати на управління підпроцесами всередині підсистеми  $q_{int}$ .

Кількість підпроцесів підсистеми  $n_j$  впливає на витрати на ресурси  $r_j$ , які необхідні ЛСУ [10]:

$$r_j = r_0 * S * n_j^\beta, \quad (5)$$

де  $r_0$  — витрати на ресурси, а  $\beta$  залежить від характеристик процесу.

Витрати на обмін інформацією залежать від кількості інформаційних зв'язків між завданнями управління підпроцесами. Загальна кількість зв'язків

$$N = k_0 n(n-1),$$

де  $n$  — загальна кількість підпроцесів;  $k_0$  — коефіцієнт заповнення матриці суміжності графа інформаційних зв'язків підпроцесів. Відповідно кількість зв'язків всередині підсистеми

$$N_{int} = k_0 n_{int} (n_{int} - 1),$$

а кількість зовнішніх зв'язків

$$N_{ext} = N - N_{int} = k_0 [n(n-1) - n_{int} (n_{int} - 1)].$$

Таким чином, витрати на обмін інформацією з урахуванням обсягу переданої інформації:

$$\begin{cases} c_{ext} = k_{3ext} I_0 k_0 [n(n-1) - n_{int} (n_{int} - 1)]; \\ c_{int} = k_{3int} I_0 k_0 [n_{int} (n_{int} - 1)], \end{cases} \quad (6)$$

де  $I_0$  — середня кількість переданої інформації між двома підпроцесами;  $k_{3int}$  — витрати на передачу одиниці інформації по внутрішнім зв'язкам;  $k_{3ext}$  — витрати на передачу одиниці інформації з зовнішніх зв'язків.

З урахуванням (5) і (6) лінеаризована модель витрат набуває вигляду:

$$\begin{cases} q = q_{ext} + q_{int}; \\ q_{ext} = r_0 * S_0 * (n - n_{int})^\beta + k_{3ext} I_0 k_0 [n(n-1) - n_{int} (n_{int} - 1)]; \\ q_{int} = r_0 * S_0 * n_{int}^\beta + k_{3int} I_0 k_0 [n_{int} (n_{int} - 1)]. \end{cases} \quad (7)$$

Дослідимо вплив кількості підпроцесів  $n_{int}$ , які керуються однією ЛСУ, на витрати. Рис. 2 показує значення  $q$  для різних значень параметрів процесу. Графіки побудовано для нормалізованих значень:

$$I_0 = 1; r_0 = 1; k_0 = 1; S_0 = 1; k_{3int} = 1; k_{3ext} = 10; n = 20.$$

Видно, що для значень  $\beta \geq 2,45$  об'єднання підпроцесів під управлінням однією ЛСУ дає суттєвий вигоду.

Задача кластеризації розв'язується на основі матриці відстаней між об'єктами кластеризації. Такими об'єктами є підзадачі управління, а відстані обернені до витрат (3)

$$d_{ij} = \frac{\Delta q_{\min}}{\Delta q_{ij}},$$

де  $d_{ij}$  — відстань між підзадачами;  $\Delta q_{ij}$  — зменшення витрат за умови об'єднання підзадач у одно-

му кластері;  $\Delta q_{\min}$  — мінімальне значення зменшення витрат (з урахуванням знаку).

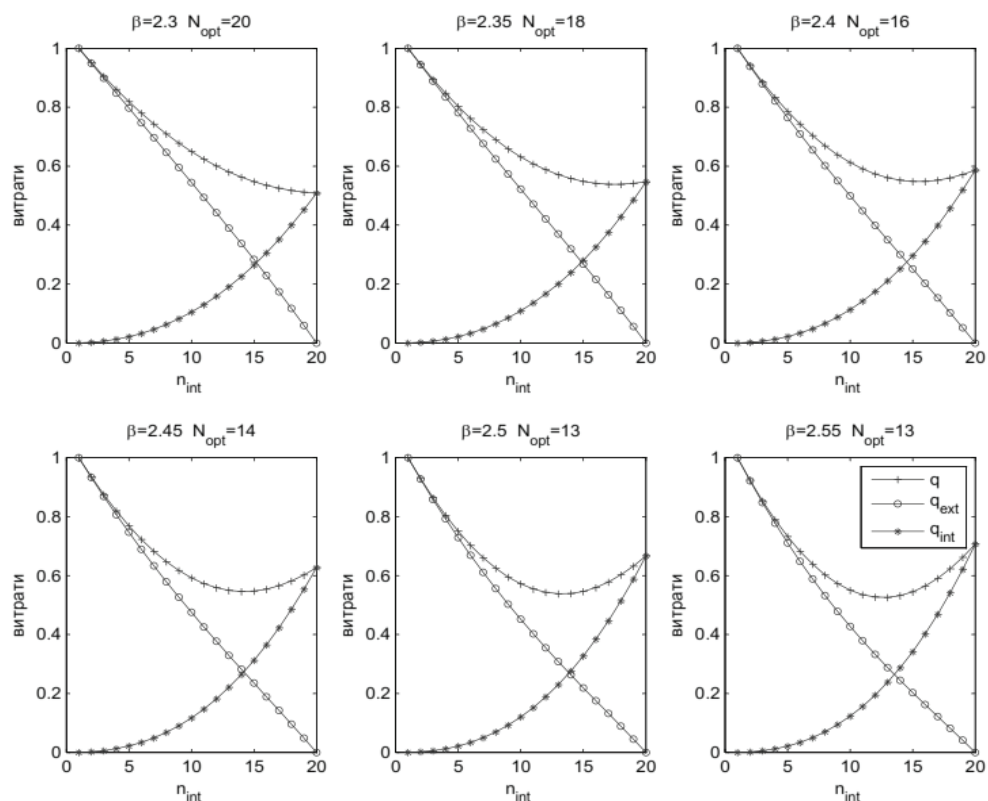


Рис. 2. Якісний характер впливу кількості задач ЛСУ на вартість виконання підпроцесу

Залежно від обраної метрики в просторі об'єктів, розмірності цього простору та інших факторів використовують різні методи кластеризації [11—12]. Перевагою ієрархічних методів кластеризації є гнучкість та простота. Серед недоліків відзначається неможливість внести виправлення після того, як рішення було прийнято, відсутність ясного тлумачення щодо кластера, і розпливчастість критерію припинення. Для подолання цих обмежень запропоновані алгоритми. CURE, BIRCH, MST та інші. У табл. 1 наведено переваги та недоліки деяких алгоритмів: Для однопараметричних задач кластеризації обчислювальна складність не виправдовує використання таких методів. У цих випадках метод «найближчого сусіда» є зручнішим варіантом, через його простоту використання і швидкість алгоритму. Обмеженням для методу «найближчого сусіда» є вимога унімодальності критерію кластеризації.

Таблиця 1

Порівняння ієрархічних алгоритмів кластеризації

Метод	Розмірність простору параметрів	Переваги	Недоліки
BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies)	Багатовимірний	Двоступенева кластеризація. Працює з великими обсягами даних. Кожне рішення кластеризації проводиться без перевірки всіх точок даних. Передбачається, що дані нерівномірно розподілені у просторі оброблюваної області, тому обробляються області з великою щільністю як єдиний кластер.	Добре виділяє тільки кластери опуклої або сферичної форми, необхідно задавати порогові значення
CURE (Clustering Using Representatives)	Багатовимірний	Виділяє кластери складної форми і різних розмірів, часова складність лінійно залежить від розмірності, виконує кластеризацію навіть за наявності викидів	Є необхідність у заданні порогових значень і кількості кластерів

MST (Minimum Spanning Trees)	Багатовимірний	Виділяє кластери довільної форми, вибирає з кількох оптимальних рішень найоптимальніше	Чутливий до викидів
Найближчого сусіда («nearest neighbour»)	Довільна	Простота використання	Існує складність вибору міри «близькості» (метрики), вирішує завдання невеликої розмірності по кількості класів та змінних
$k$ -середніх (K-means)	Довільна	Простота використання; швидкість; зрозумілість і прозорість алгоритму	Шукає локальний оптимум; результат залежить від початкового вибору центроїдів. Необхідно знати кількість кластерів. Чутливий до викидів. Не підходить для виявлення неопуклих кластерів
Алгоритм HCM (Hard C — Means)	Довільна	Легкість реалізації, обчислювальна простота	Необхідне задання кількості кластерів, відсутність гарантії знаходження оптимального рішення

Для декомпозиції задач управління розподіленої технологічної системою можна запропонувати модифікований алгоритм кластеризації за методом найближчого сусіда.

Задачу кластеризації вирішуватимемо шляхом мінімізації критерію (2).

На підготовчому етапі оцінюється доцільність кластеризації на основі дослідження моделі (7).

Основні кроки алгоритму кластеризації:

1. Виконання алгоритму починається зі стану, коли кожна вершина-задача є окремим кластером, тобто вирішується окремою ЛСУ, а всі дуги-зв'язки позначені як зовнішні.

2. Виконується об'єднання кластерів, допоки серед усіх пар кластерів є такі, об'єднання яких зменшує критерій загальних витрат.

2.1. Переглядаються всі наявні кластери.

2.1.1. Переглядаються всі сусідні кластери по відношенню до розглянутого кластера.

2.1.2. Розраховуються зміни витрат в результаті об'єднання кожної пари вершин.

2.1.3. Знаходиться варіант з найбільшим зменшенням витрат.

2.1.4. Якщо це зменшення додатне, то кластери об'єднуються

2.1.4.1. Здійснюється агрегування кластерів за правилами агрегування вершин графів.

2.1.4.2. Зменшується розмірність вектора вартості ресурсів і в нього замість вартостей ресурсів окремих кластерів, які об'єднувалися, записується вартість ресурсів об'єданого кластера.

2.1.4.3. Зв'язки між кластерами, які об'єднувалися, позначаються як внутрішні.

2.1.4.4. Кінець циклу.

2.1.5. Кінець циклу.

2.2. Кінець циклу.

3. Кінець.

Розглянемо приклад кластеризації множини з десяти задач [13—14]:  $T = \{T_1, \dots, T_i, \dots, T_{10}\}$ , і множини з трьох координаторів:  $P = \{p_1, p_2, p_3\}$ .

Мета полягає в мінімізації загальних витрат на виконання задач і обмін інформацією. Позначимо  $k_1 S_{ij}$  — витрати на виконання задачі  $T_i$ , коли вона виконується координатором  $p_j$ , і  $k_3 I_{iv}$  — витрати на обмін інформацією між задачами  $i$  і  $v$ . Дані про витрати виражені в одиницях часу на виконання обробки та обміну інформацією. Таке вимірювання дозволяє абстрагуватися від змін цін і може бути легко конкретизовано. Вихідні дані задачі зведені в табл. 2 і показані на рис. 3.

Застосовуючи запропонований алгоритм кластеризації до даних прикладу, отримуємо такий розподіл задач  $p_1$ :  $T_2, T_3, T_5$ ;  $p_2$ :  $T_4, T_7, T_8$ ;  $p_3$ :  $T_1, T_6, T_9, T_{10}$ .

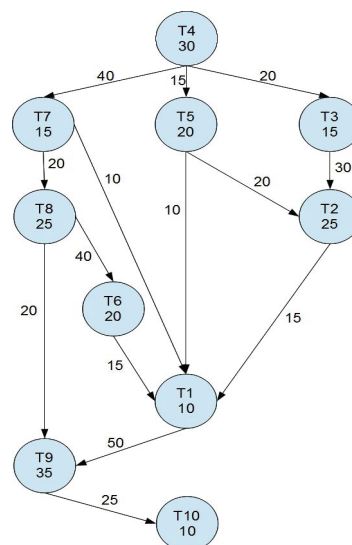


Рис. 3. Направлений ациклічний граф взаємозв'язку підзадач

Інформаційний потік  $I_{iv}$  між підзадачами

	T <sub>1</sub>	T <sub>2</sub>	T <sub>3</sub>	T <sub>4</sub>	T <sub>5</sub>	T <sub>6</sub>	T <sub>7</sub>	T <sub>8</sub>	T <sub>9</sub>	T <sub>10</sub>
T <sub>1</sub>	0	0	0	0	0	0	0	0	50	0
T <sub>2</sub>	15	0	0	0	0	0	0	0	0	0
T <sub>3</sub>	0	30	0	0	0	0	0	0	0	0
T <sub>4</sub>	0	0	20	0	15	0	40	0	0	0
T <sub>5</sub>	10	20	0	0	0	0	0	0	0	0
T <sub>6</sub>	15	0	0	0	0	0	0	0	0	0
T <sub>7</sub>	10	0	0	0	0	0	0	20	0	0
T <sub>8</sub>	0	0	0	0	0	40	0	0	20	0
T <sub>9</sub>	0	0	0	0	0	0	0	0	0	25
T <sub>10</sub>	0	0	0	0	0	0	0	0	0	0

Порівняння результатів, отриманих в роботі [14], з результатами, отриманими за допомогою запропонованого алгоритму кластеризації, показано в табл. 3.

Таблиця 3

Результати кластеризації підзадач

Координатор	Приклад Williams [14]		Модифікований алгоритм кластеризації	
	задачі	вартість	задачі	вартість
$p_1$	2, 5	115	2, 3, 5	120
$p_2$	3, 4, 6, 7	180	4, 7, 8	145
$p_3$	1, 8, 9, 10	160	1, 6, 9, 10	140
Максимум		180		145

Зауважимо, що наш алгоритм домігся кращого розподілу робочого навантаження.

### Висновки

Запропонований метод декомпозиції задач управління за допомогою кластеризації дозволяє звести до мінімуму міжкластерний обмін інформацією. Результати можуть бути використані для підвищення ефективності управління складними розподіленими технологічними процесами.

### Подяки

Ця робота була виконана за фінансової підтримки гранту «Міністерства вищої освіти, науки, технологій та інновацій «SENESCYT» Еквадор, і спонсорується університетом Santa Elena «UPSE», Еквадор.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Месарович М. Теория иерархических многоуровневых систем / Месарович М., Мако Д. Такахара И. — М. : Мир, 1973. — 344 с.
2. Бурков В. Н. Основы математической теории активных систем / В. Н. Бурков. — М. : Наука, 1977. — 255 с.
3. Волкович В. Л. Системный подход к исследованию иерархических систем управления / В. Л. Волкович, Н. Ф. Радомский : матер. Международного симпозиума по проблемам организационного управления и иерархическим системам, Баку, 1971. — С. 25—31.
4. Ладанюк А. П. Системна задача координації в технологічних комплексах неперервного типу [Електронний ресурс] / А. П. Ладанюк, Д. А. Шумигай, Р. О. Бойко. — Режим доступу : [http://dspace.nuft.edu.ua/jspui/bitstream/123456789/4444/1/Sh\\_3.pdf](http://dspace.nuft.edu.ua/jspui/bitstream/123456789/4444/1/Sh_3.pdf).
5. Berkhin P. Survey of Clustering Data Mining Techniques / P. Berkhin. — Accrue Software, 2002.
6. Ткаченко В. Н. Координация взаимодействия локальных систем регулирования процесса непрерывной разливки стали [Электронный ресурс] / В. Н. Ткаченко, О. С. Волуева. — Режим доступу : [http://ea.dgtu.donetsk.ua:8080/jspui/bitstream/123456789/7797/1/статья\\_Волуева\\_Ткаченко.pdf](http://ea.dgtu.donetsk.ua:8080/jspui/bitstream/123456789/7797/1/статья_Волуева_Ткаченко.pdf).
7. Sethi S. P. et al. Optimal and hierarchical controls in dynamic stochastic manufacturing systems : A survey / S. P. Sethi, H. Yan, H. Zhang, Q. Zhang // Manufacturing & Service Operations Management. — 2002. — № 2. (4). — С. 133—170.
8. Parmer G. West R. HIREs : A system for predictable hierarchical resource management / G. Parmer, R. West // Real-Time and Embedded Technology and Applications Symposium (RTAS), 2011 17th IEEE. — 2011. — С. 180—190.

9. Дубовой В. М., Никитенко О. Д. Оптимізація підсистем збору даних АСУТП в умовах комбінованої невизначеності / В. М. Дубовой, О. Д. Никитенко : моногр. — Вінниця : УНІВЕРСУМ-Вінниця, 2011. — 169 с.
10. Stephen Pierzchala. Java Enterprise Performance / Stephen Pierzchala & Andreas Grabner / Compuware Corporation. — 2012.
11. Berkhin P. Survey of Clustering Data Mining Techniques / P. Berkhin / Accrue Software. — 2002.
12. Murty J. Data Clustering : A Review / Jain Murty and Flynn : ACM Comp. Surv. — 1999.
13. Ucar B. Task assignment in heterogeneous computing systems / B. Ucar, C. Aykanat, K. Kaya, & M. İkinci // Journal of parallel and Distributed Computing. — 2006. — Vol. 66, No. 1. — С. 32—46.
14. Woodside C. M. Fast Allocation of Processes in Distributed and Parallel Systems / C. M. Woodside, and G. G. Monforton // IEEE Transactions on parallel and distributed systems. — 1993. — С. 164—174.

Рекомендована кафедрою комп'ютерних систем управління ВНТУ

Стаття надійшла до редакції 12.11.2014

**Байас Марсія Марисоль** — аспірантка кафедри комп'ютерних систем управління, email: marciabayas@gmail.com;

**Дубовой Володимир Михайлович** — д-р техн. наук, професор, завідувач кафедри комп'ютерних систем управління, e-mail: v.m.dubovoy@gmail.com;

**Дуда Маріна Ернестівна** — студентка Інституту автоматики, електроніки та комп'ютерних систем управління.

Вінницький національний технічний університет, Вінниця

**M. M. Bayas**<sup>1,2,3</sup>

**V. M. Dubovoi**<sup>1</sup>

**M. E. Duda**<sup>1</sup>

## Decomposition of control tasks by clustering method

<sup>1</sup>Vinnytsia National Technical University

<sup>2</sup>Universidad Peninsula de Santa Elena», Ecuador

<sup>3</sup>SENESCYT «Ministry of Higher Education, Science, Technology and Innovation», Ecuador

*This article discusses the construction of clustering criterion and algorithm for solving the task of decomposition in the formation of the structure of the process control. An algorithm based on the nearest neighbor method is offered.*

**Keywords:** clustering algorithm, task coordination, task decomposition, cost of information exchange.

**Bayas Marcia M.** — Post-Graduate Student of the Chair of Computer Control Systems, e-mail: marciabayas@gmail.com;

**Dubovoi Volodymyr M.** — Dr. Sc. (Eng.), Professor, Head of the Chair of Computer Control Systems, e-mail: v.m.dubovoy@gmail.com;

**Duda Maryna E.** — Student of the Institute of Automation, Electronics and Computer Systems

**М. М. Байас**<sup>1,2,3</sup>

**В. М. Дубовой**<sup>1</sup>

**М. Э. Дуда**<sup>1</sup>

## Декомпозиция задач управления методом кластеризации

<sup>1</sup>Винницкий национальный технический университет, Украина;

<sup>2</sup>Universidad Estatal Península de Santa Elena, Эквадор;

<sup>3</sup>SENESCYT «Министерство высшего образования, науки, технологий и инноваций», Эквадор

*Рассмотрены вопросы построения критерия и алгоритма кластеризации для решения задачи декомпозиции при формировании структуры управления технологическим процессом. Предлагается алгоритм на основе метода ближайшего соседа.*

**Ключевые слова:** алгоритм кластеризации, координация задач, декомпозиция задач, затраты на обмен информацией.

**Байас Марсія Марисоль** — аспірант кафедри комп'ютерних систем управління, e-mail: marciabayas@gmail.com;

**Дубовой Владимир Михайлович** — д-р техн. наук, профессор, заведующий кафедрой комп'ютерних систем управління, e-mail: v.m.dubovoy@gmail.com;

**Дуда Марина Эрнестовна** — студент Інституту автоматики, електроніки та комп'ютерних систем управління