

Додаток А

Міністерство освіти і науки України

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук

Кафедра Інформаційних управляючих систем

МАГІСТЕРСЬКА АТЕСТАЦІЙНА РОБОТА
ГРАФІЧНИЙ МАТЕРІАЛ

Дослідження методів кластеризації трас журналів подій в проектах
інтелектуального аналізу процесів

Студент УПІТМ-18-1

(шифр групи)

(підпис)

Волошин М.С.

(прізвище, ініціали)

Керівник роботи

(підпис)

проф. Чалий С.Ф.

(посада, прізвище, ініціали)

2019 р.

Цілі та задачі досліджень

Об'єкт дослідження	Процес формування моделей бізнес процесів «як є» за технологією process mining
Предмет дослідження	Методи кластеризації трас журналів подій.
Мета роботи	Удосконалення методу кластеризації трас на основі побудови ієрархії дій процесу
Научна новизна	Удосконалення методів побудови моделей багатоваріантних процесів на основі аналізу логів шляхом ієрархічної кластеризації трас з подальшим виділення окремих конфігурацій таких процесів
Практичні результати	Алгоритм кластеризації трас для побудови моделей багатоваріантних процесів засобами process mining
Задачі дослідження	<ol style="list-style-type: none"> 1. Аналіз задач process mining; 2. Дослідження методів кластеризації для інтелектуального аналізу процесів; 3. Удосконалення інтелектуального аналізу процесів на основі ієрархічної кластеризації трас журналу подій; 4. Розробка плану проекту розробки модулю ієрархічної кластеризації трас; 5. Експериментальна перевірка отриманих результатів

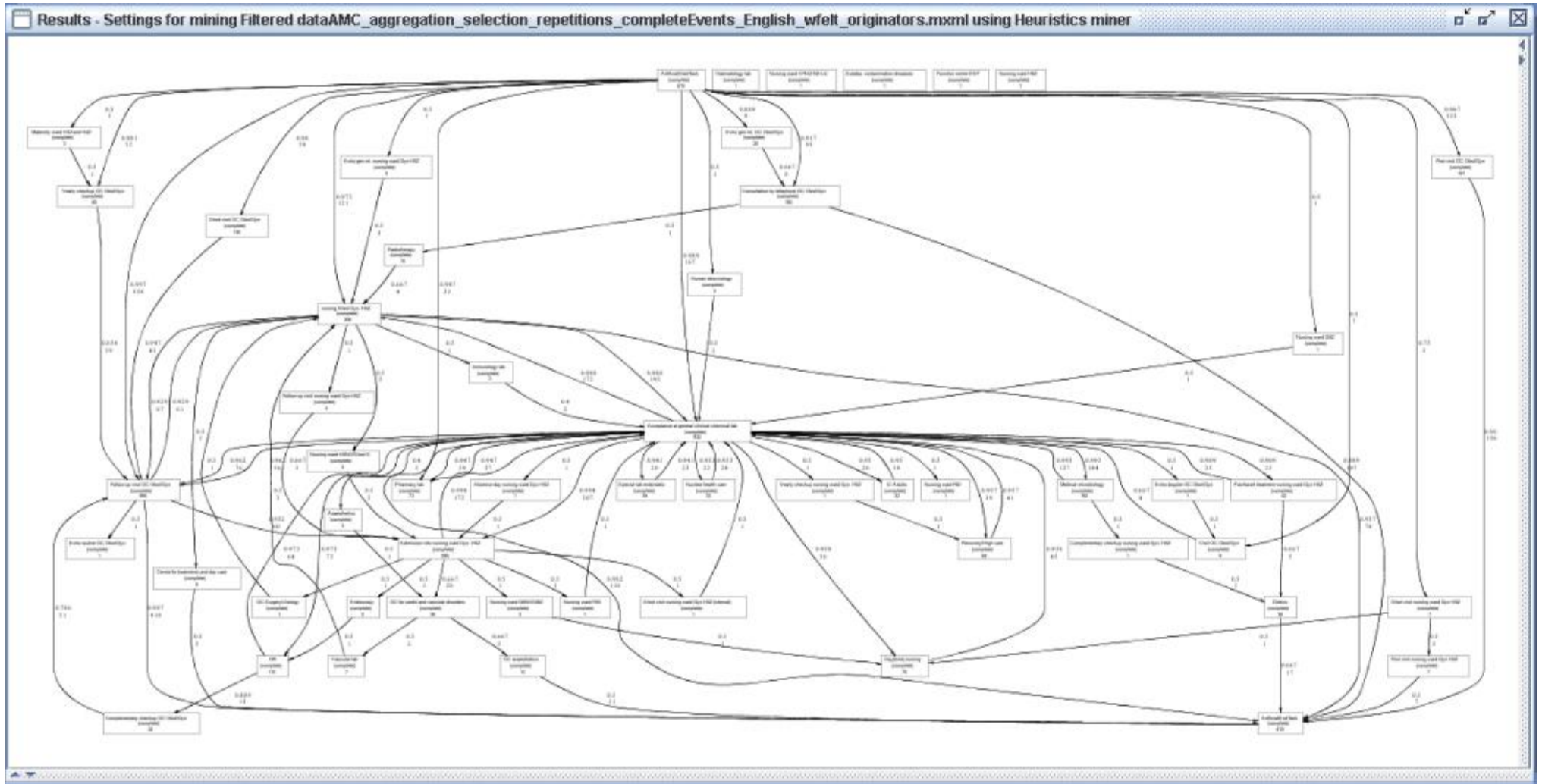
Задачі інтелектуального аналізу процесів



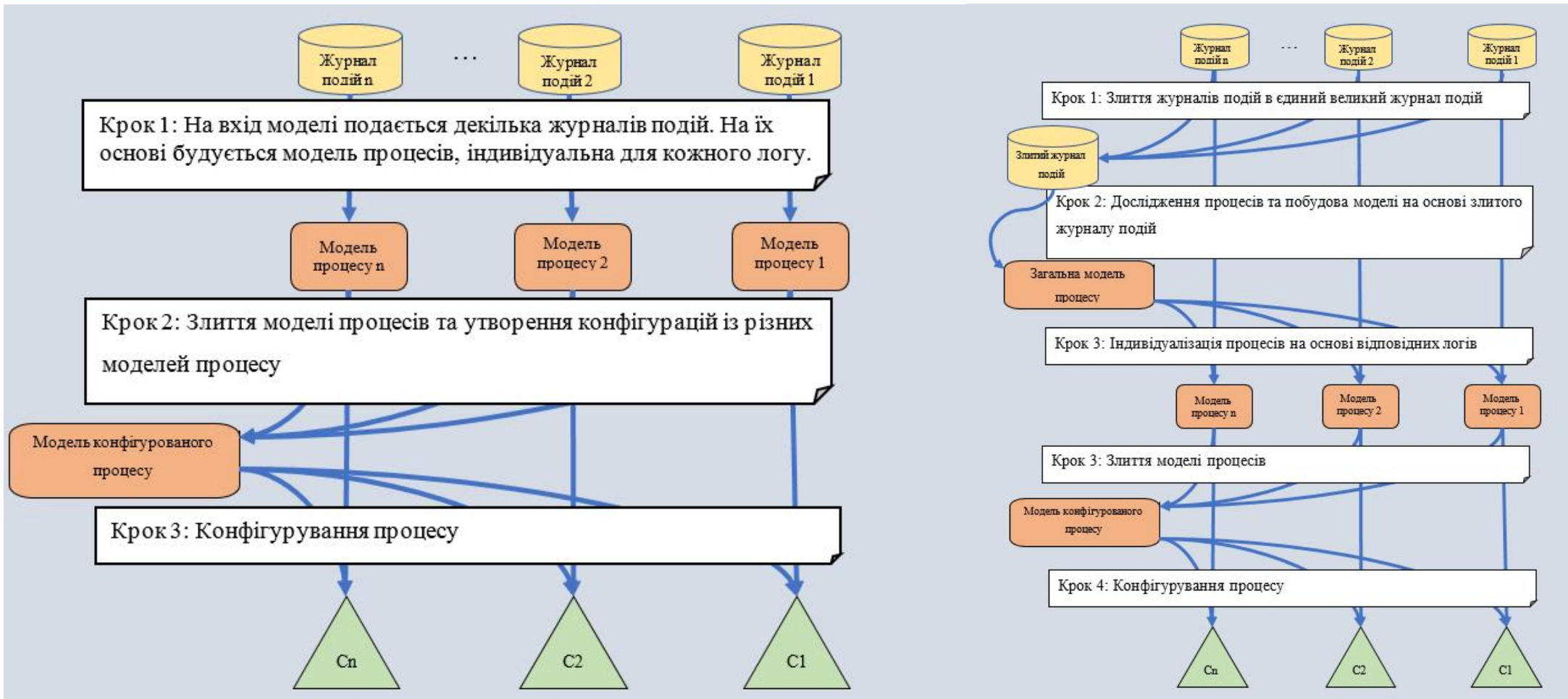
Приклад журналу подій

№	Атрибут				Кількість занять	Кількість незаліків	...	№ траєн	№ події	Атрибути		Ресурс	...
	Оплата	Студент	Інструктор	Екзаменатор						Часова мітка	Дія		
221	Абонемент	Артур	Сергій	Вікторія	3	0	...	221	41251	02.05.2014 13:05	Початок(A)	Микола	...
285	Після класу	Марія	Сергій	Микола	2	1	...		41252	02.05.2014 13:12	Придбати зняття + Екзамен(B)	Микола	...
...		41257	02.05.2014 13:31	Заняття в класі(C)	Сергій	...
									41258	02.05.2014 14:32	Потрібно ще(E)	Сергій	...
									41261	05.05.2014 15:14	Заняття в класі (C)	Сергій	...
									41262	05.05.2014 16:17	Потрібно ще (E)	Сергій	...
									41269	07.05.2014 13:01	Теоретичний Екзамен(F)	Сергій	...
									41271	10.05.2014 14:59	Заняття в класі (C)	Сергій	...
									41280	15.05.2014 15:57	Здати Практичний Екзамен(G)	Вікторія	...
									41295	21.05.2014 14:07	Отримати водійське посвідчення(I)	Микола	...
								258	53151	21.05.2014 14:07	Початок(A)	Микола	...
									53151	21.05.2014 14:07	Заняття в класі (C)	Микола	...
									53151	21.05.2014 14:07	Сплатити за заняття (D)	Сергій	...
									53151	21.05.2014 14:07	Потрібно ще(E)	Сергій	...
									53151	21.05.2014 14:07	Заняття в класі (C)	Сергій	...
									53151	21.05.2014 14:07	Потрібно ще (D)	Сергій	...
									53151	21.05.2014 14:07	Теоретичний Екзамен(F)	Сергій	...
									53151	21.05.2014 14:07	Заняття в класі (G)	Сергій	...
									53151	21.05.2014 14:07	Здати Практичний Екзамен(J)	Вікторія	...
							

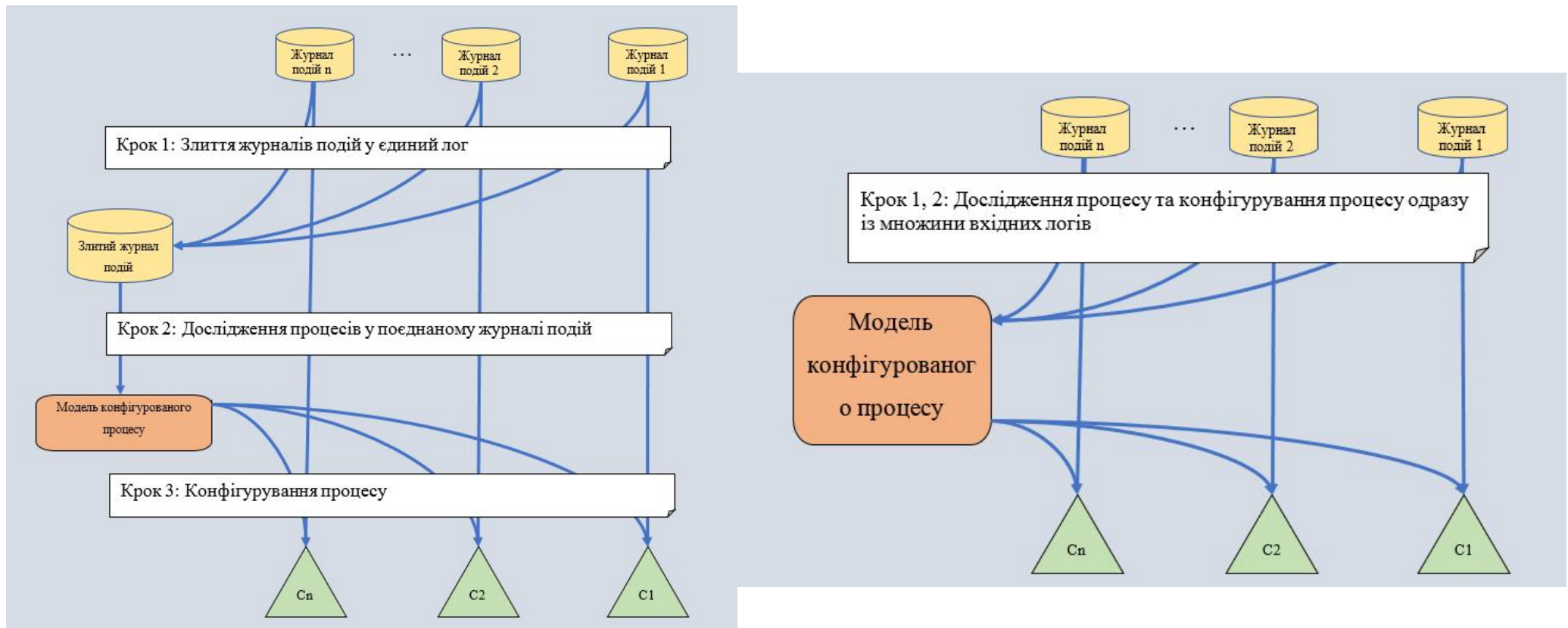
Спагеті – подібна модель



Підходи побудови моделей багатоваріантних процесів (1)

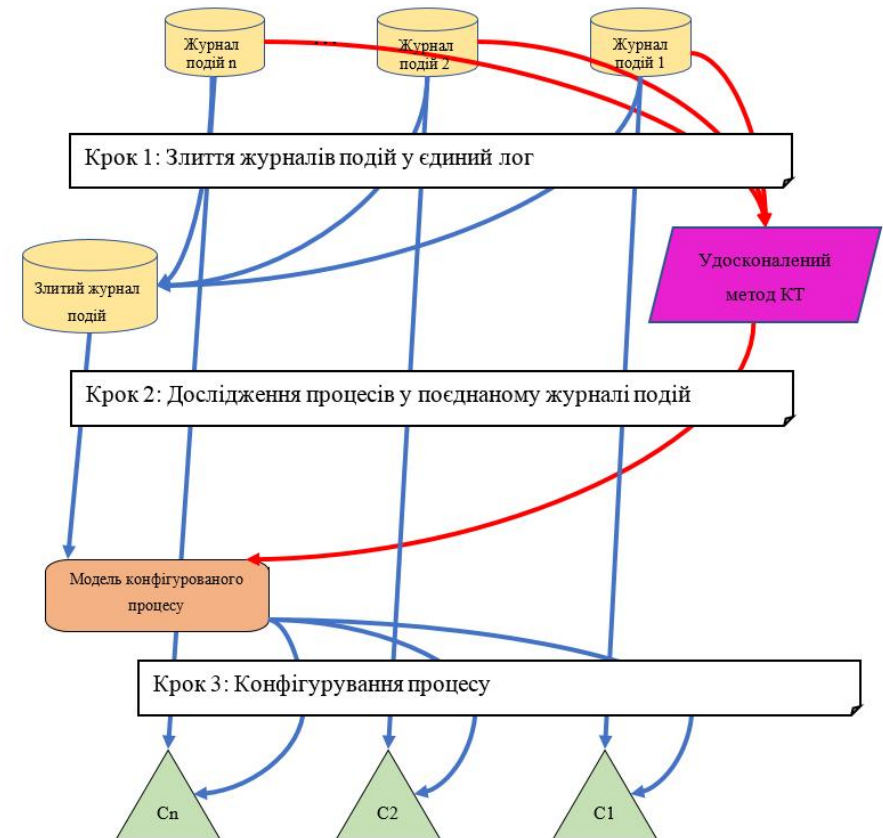


Підходи побудови моделей багатоваріантних процесів (2)



Загальна схема удосконаленого методу кластеризації трас

Покращення дослідження
конфігурацій процесів шляхом
отримання базової структури дерева
процесів за допомогою алгоритму
ієрархічної кластеризації



Етапи удосконаленого методу

Метод містить у собі наступні етапи

1. Розрахунок кількості повторів пар сусідніх подій на одній трасі та по трасам.

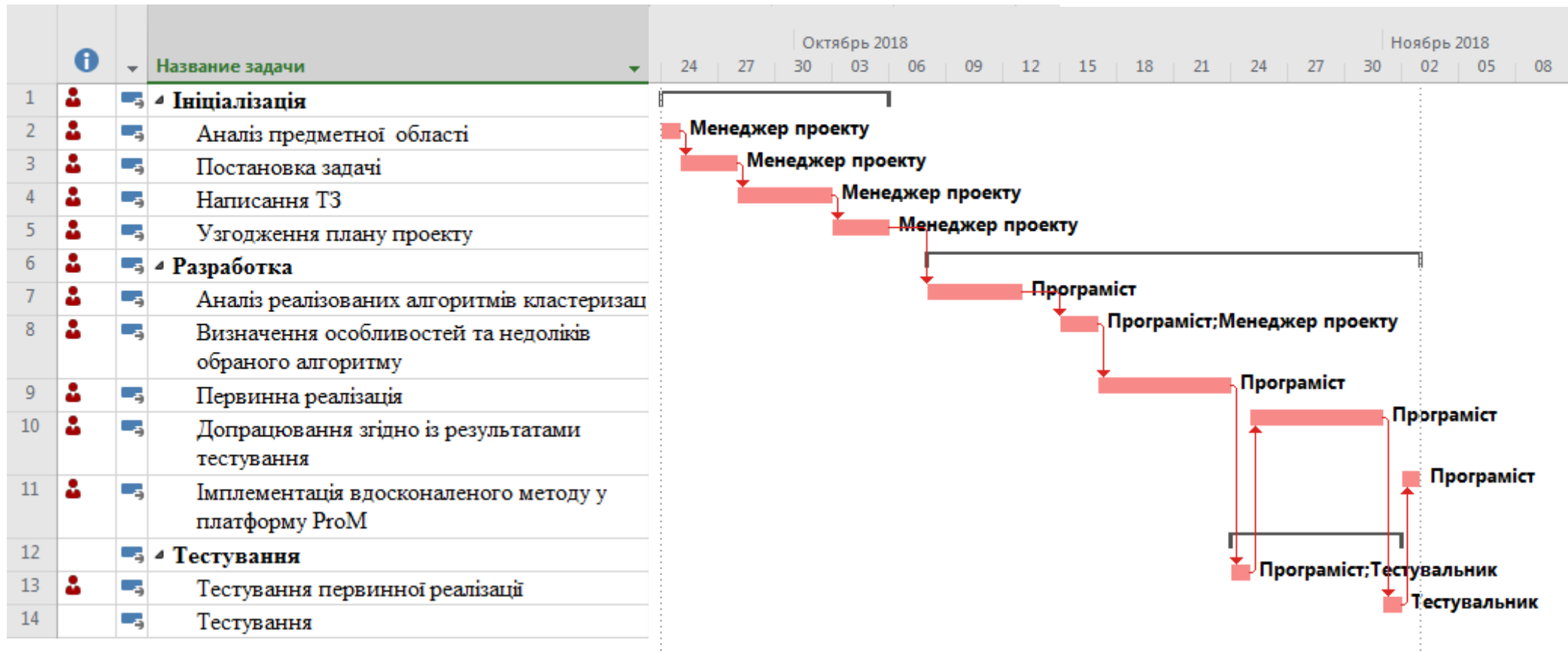
2. Відбір пар подій таких, що задовольняють умові $k_{ei} \geq k_t$, де k_{ei} – кількість пар (e_i, e_{i+1}) , k_t – кількість трас.

Отримана множина подій складає базову конфігурацію бізнес-процесу. Для зручності позначень введено поняття ваги пари

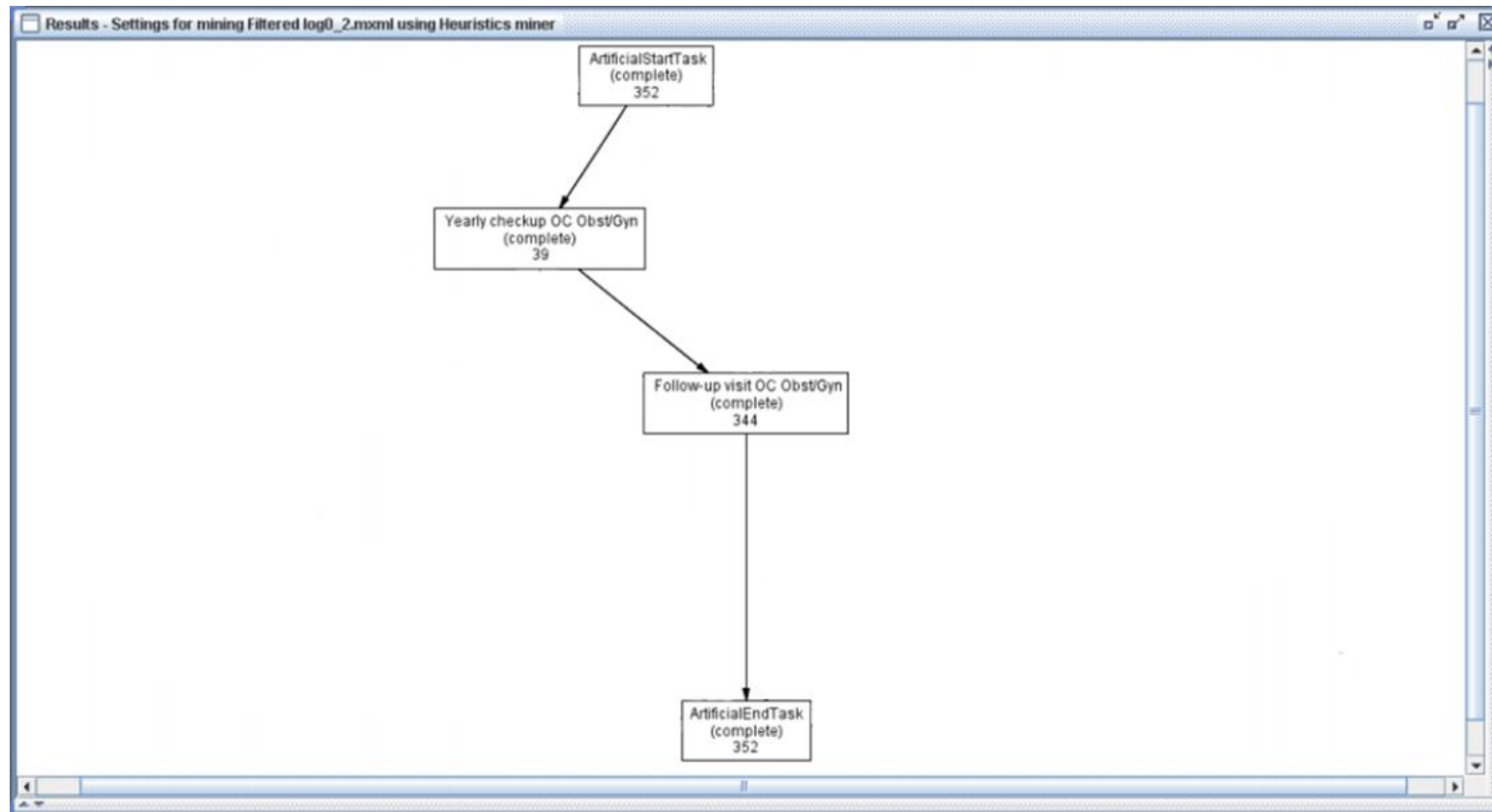
$$w_t = \frac{L}{T}$$

3. розбиття множини подій, що залишилися на задану кількість кластерів. Результатом буде множина додаткових конфігурацій багатоваріантного БП.

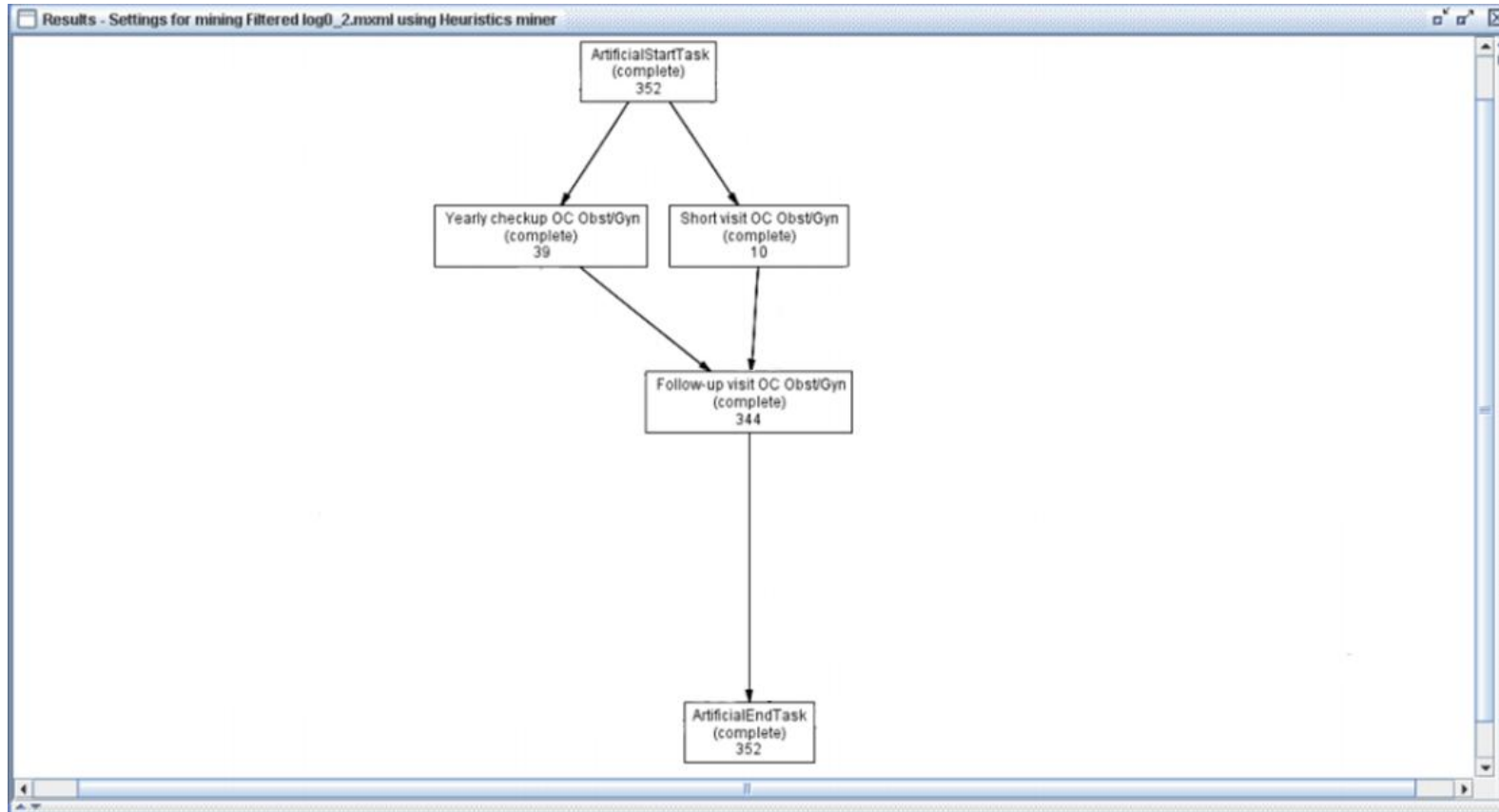
План проекту реалізації модернізованого методу



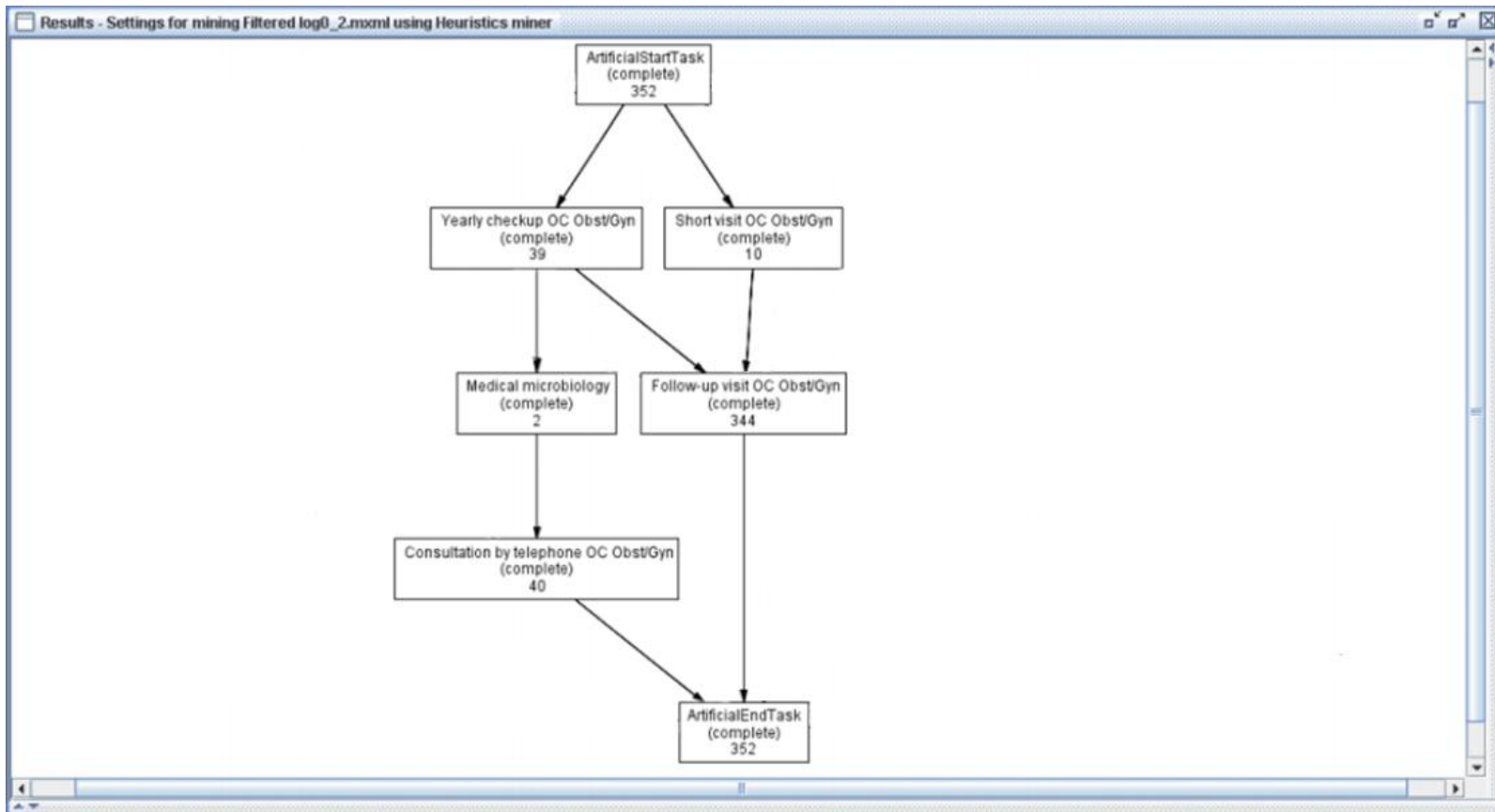
Демонстрація роботи методу при $w_t \geq 1$



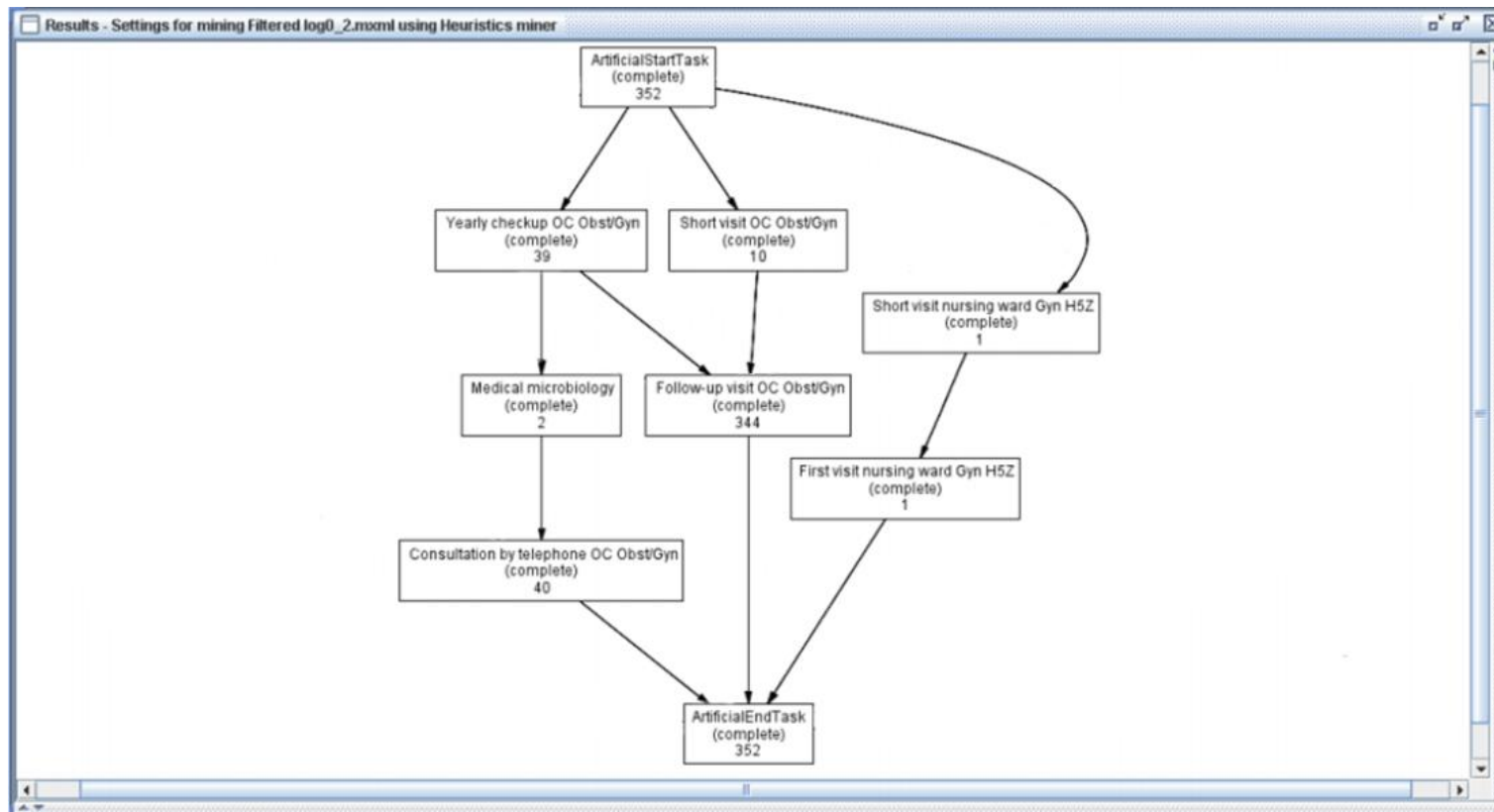
Демонстрація роботи методу $w_t \in (0,75; 1]$



Демонстрація роботи методу $w_t \in (0,50; 0,75]$



Демонстрація роботи методу $w_t \in (0; 0,50]$



Висновки

- Під час виконання атестаційної роботи біли виконані наступні завдання: Аналіз задач process mining; Дослідження методів кластеризації для інтелектуального аналізу процесів; Удосконалення інтелектуального аналізу процесів на основі ієрархічної кластеризації трас журналу подій; Розробка плану проекту розробки модулю ієрархічної кластеризації трас; Експериментальна перевірка отриманих результатів.
- Практичним результатом є алгоритм кластеризації трас для побудови моделей багатоваріантних процесів засобами