

Інтелектуальний аналіз транспортних процесів

Юлія Сілантьєва

Академія технічних наук України

Київ, Україна

Матеріали розділу подані для самостійної підготовки здобувачів до практичних занять в рамках освітніх компонентів, програми яких передбачають вивчення інформаційних технологій для модернізації транспортних процесів в системі автомобільних вантажних перевезень.

На сьогодні прискорення виконання зовнішньоторговельних угод щодо постачання товарів вимагає більш уважного й чіткого узгодження роботи великої кількості учасників, задіяних в процесі укладання договорів купівлі-продажу, підготовки товару до відвантаження, перевезення за участю різних видів транспорту, інформаційного та документального обміну, митних, фінансових і страхових операцій тощо. Тобто всіх тих етапів, які забезпечують успішність виконання транспортних процесів для системи вантажних перевезень (як міжнародних, так і внутрішніх). З огляду на спрощення такого узгодження між учасниками транспортного процесу через підвищення рівня автоматизації процесів, використання нових транспортних та інформаційних технологій, зменшення навантаження нетарифних заходів регулювання зовнішньоекономічної діяльності, що значно покращує реалізацію угод, все ще залишаються невизначеності, які постійно загрожують зривами строків й умов постачання або додатковими витратами ресурсів. Такі невизначеності можна встановити й спрогнозувати, застосовуючи методи науки про дані та процеси. Одним з розділів науки про процеси є інтелектуальний аналіз процесів, який пропонує широкий спектр інформаційних технологій й інструментів для моделювання та пошуку вирішення поставлених задач модернізації певного процесу в заданих умовах.

I. ТЕОРЕТИЧНІ ОСНОВИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ПРОЦЕСІВ (PROCESS MINING)

Вражаюче зростання цифрового всесвіту, відображене у терміні «Big Data», дає змогу фіксувати, демонструвати й аналізувати *події*. Події можуть відбуватися всередині механізму (наприклад, банкомату або системи сканування багажу), в межах інформаційної системи підприємства (наприклад, розміщення клієнтом замовлення або подання митної декларації в митні органи), в лікарні (наприклад, аналіз крові), в соціальній мережі (наприклад, обмін повідомленнями електронною поштою або у Telegram), в елементах транспортної системи (наприклад, покупка проїзних білетів, перетин пункту пропуску з автоматичною фіксацією автомобільних номерів чи стягування плати за проїзд автомобільними шляхами) тощо. Подіями можуть бути «події з життя», «події в машині» або «події в організації». Термін Інтернет подій (IoE) відноситься до всіх відкритих даних про події (рис.1, [1]).

IoE складається з:

- *Інтернет контенту* (IoC), тобто вся інформація, створена людьми для розширення знань щодо конкретних предметів. IoC включає веб-сторінки, статті, енциклопедії (наприклад, Вікіпедія), YouTube, електронні книги, новинні стрічки тощо.
- *Інтернет людей* (IoP), тобто всі дані, пов'язані із соціальною взаємодією. IoP включає електронну пошту, Facebook, Twitter, форуми, LinkedIn тощо.
- *Інтернет речей* (IoT), тобто всі фізичні об'єкти, підключені до мережі. IoT включає всі предмети, які мають унікальний ідентифікатор і присутні в структурі, подібної до Інтернету.
- *Інтернет розташування або локацій* (IoL), який відноситься до всіх даних, що мають геопросторовий вимір. З поширенням мобільних пристроїв дедалі більше подій мають атрибути місцезорозташування або руху.

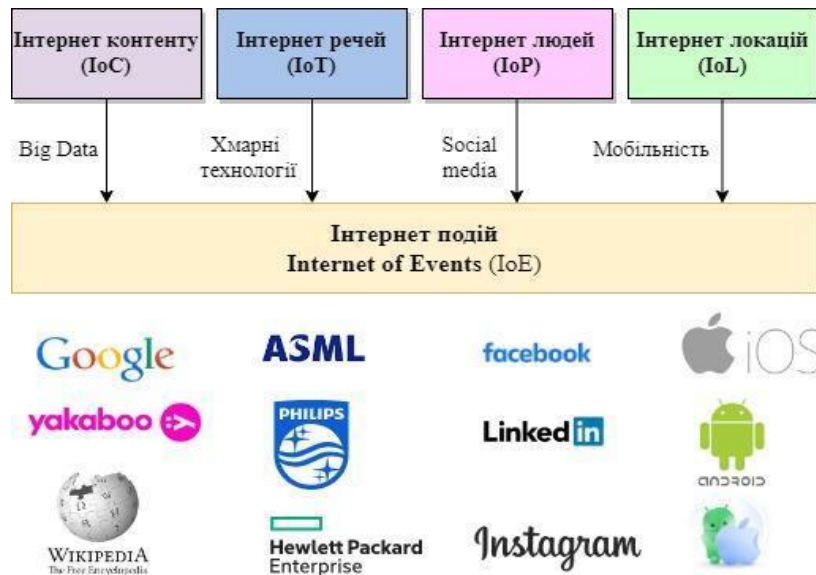


Рисунок 1 – Приклади Internet-джерел, що генерують дані про події

Складові ІоЕ перетинаються (наприклад, посилання на місцерозташування об'єкту на веб-сторінці або географічне розташування пристрою, з якого було надіслано пост в соціальну мережу).

Інтелектуальний аналіз процесів спрямований на осмислене використання даних про події, наприклад, для надання інформації, виявлення «вузьких місць», прогнозування проблем, реєстрації порушень, рекомендацій щодо контрзаходів та оптимізації процесів. Вхідні дані для інтелектуального аналізу процесів – це *журнал подій*, який можна розглядати як окремий обсяг наявних даних про певні події. Наприклад, журнал подій може містити перелік всіх подій, пов'язаних з усіма клієнтами компанії, і використовуватися для визначення дорожньої карти співпраці з окремим клієнтом чи певною підмножиною клієнтів.

Наука про дані – це міждисциплінарна область (рис. 2, [2]), метою якої є перетворення даних на реальну цінність. Дані можуть бути структурованими або неструктурованими, великими за обсягом або обмеженими, статичними або поточковими. Дані можуть бути представлені у вигляді прогнозів, автоматизованих рішень, моделей. Наука про дані включає отримання, підготовку, дослідження, перетворення, зберігання та пошук даних, обчислювальну

інфраструктуру, представлення пояснень та прогнозів, а також використання результатів роботи з даними з урахуванням етичних, соціальних, юридичних та ділових аспектів.



Рисунок 2 – Складові науки про дані

На рис. 2 орієнтовно наведені основні складові науки про дані. Дисципліни перетинаються і розрізняються за розміром. Крім того, їхні кордони не є чіткими та змінюються з часом. Їхні коріння дуже різні, наприклад, інтелектуальний аналіз даних виник поруч із базами даних, а машинне навчання розпочало свій розвиток у співтоваристві з штучним інтелектом, виникнення ж обох дисциплін не пов'язано зі статистикою. Наукова фантастика першою познайомила світ із концепцією роботів із розумом людини. Незважаючи на різне коріння, три дисципліни виразно перетинаються. Умовні складові науки про дані наступні:

- *Статистику* можна розглядати як джерело науки про дані. Дисципліна, як правило, поділяється на описову статистику (для узагальнення вибіркового даних, використовуючи такі поняття, як середнє значення, стандартне відхилення, частота) та логічну

статистику (використання вибіркового даних для оцінки характеристик усіх даних або для перевірки гіпотези).

- *Алгоритми* мають вирішальне значення в будь-якому підході до аналізу даних. Спрощено, алгоритм – це скінченний набір кроків для досягнення бажаного результату [3]. Алгоритми бувають рандомізовані і ймовірнісні, детерміновані й недетерміновані, примітивні і складені, онлайн і офлайн, евристичні тощо. Для розробки програмного забезпечення використовують алгоритми для баз даних, розподілених обчислень, виділення пам'яті, роботи з дисками та комп'ютерною мережею. Для машинного навчання розроблені алгоритми для навчання з вчителем/без вчителя, з підкріпленням, напіваавтоматичне, глибинне навчання та ін. Комбінаторика, інформатика, обчислювальна математика також використовують власний широкий спектр алгоритмів. Коли набори даних збільшуються, складність алгоритмів стає першочерговою проблемою. Apriori, MapReduce, PageRank – приклади реалізації різних типів алгоритмів.

- *Інтелектуальний аналіз даних* можна визначити як «аналіз (зазвичай, великих) наборів даних для пошуку непередбачуваних зв'язків та узагальнення даних новими способами, зрозумілими та корисними для власника даних» [4]. Вхідні дані, зазвичай, наводяться у вигляді таблиці, а вихідними можуть бути правила, кластери, деревоподібні структури, графіки, рівняння, шаблони тощо. Очевидно, що інтелектуальний аналіз даних базується на статистиці, базах даних та алгоритмах. Порівняно зі статистикою, тут увага зосереджена на масштабованості та практичному застосуванні.

- *Машинне навчання* стосується питання, як побудувати комп'ютерні програми, які автоматично покращуються з досвідом. Різниця між інтелектуальним аналізом даних і машинним навчанням неоднозначна. Сфера машинного навчання виникла в рамках штучного інтелекту за допомогою таких методів, як нейронні мережі. Термін «машинне навчання» позначає алгоритми, які дають комп'ютерам можливість навчатися без явного програмування.

- *Інтелектуальний аналіз процесів (process mining)* додає до машинного навчання та інтелектуального аналізу даних перспективи з точки зору процесу. Інтелектуальний аналіз процесів зіставляє дані про події (тобто спостерігає поведінку) і моделі процесів (наприклад, з використанням мереж Петрі або моделей BPMN).

- *Предиктивна аналітика (прогнозування)* – це практика вилучення інформації з існуючих наборів даних для визначення закономірностей та прогнозування майбутніх результатів і тенденцій. Для створення прогнозів, існуючі підходи щодо отримання корисної інформації та навчання застосовуються у бізнес-контексті. Предиктивна аналітика тісно пов'язана з бізнес-аналітикою.

- *Бази даних* використовуються для зберігання даних. Системи управління базами даних (СУБД) мають дві основні цілі: структурування даних для легкого керування ними та забезпечення масштабованості й надійної роботи. Донедавна нормою були реляційні бази даних і SQL. Через зростаючий обсяг різномірних даних з'явилися розподілені бази даних і NoSQL.

- *Розподілені системи* забезпечують інфраструктуру щодо аналізу. Розподілена система складається з взаємодіючих компонентів, які координують свої дії задля досягнення спільної мети. Деякі завдання аналізу надто великі або надто складні для виконання на одному комп'ютері. Такі завдання можна розділити на безліч дрібних завдань, які можна виконувати одночасно на різних обчислювальних вузлах. Масштабованість може бути реалізована шляхом спільного використання та/або розширення набору обчислювальних вузлів.

- *Візуалізація/ візуальна аналітика* поєднує методи автоматизованого аналізу з інтерактивною візуалізацією для ефективного розуміння проблем та прийняття рішень на основі дуже великих і складних наборів даних.

- *Бізнес-моделювання* надає економічної цінності даним для досягнення компаніями успіху.

- *Наука про поведінку (біхевіористика)* – це системний аналіз поведінки людини. Соціальні науки вивчають процеси соціальної системи та відносини між людьми в суспільстві.

- *Конфіденційність, захищеність, етика* є ключовими складовими захисту окремих осіб та організацій від «некоректних» методів дослідження даних. Дані повинні бути точними та безпечно зберігатися, не допускаючи несанкціонованого доступу. Поряд з конкретними порушеннями конфіденційності та безпеки, тут також досліджуються поняття і правила етичні поведінки.

Наука про процеси – це загальний термін для більш широкої дисципліни, яка поєднує знання з інформаційних технологій та знання з управлінських наук для покращення та управління операційними процесами. Підходи науки про дані, як правило, не залежать від процесу, тоді як підходи науки про процеси, як правило, засновані на моделях з урахуванням «доказів», прихованих у даних. Умовні складові науки про процеси наступні:

- *Стохастика* пропонує набір методів аналізу випадкових процесів. Поведінка процесу чи системи моделюється з допомогою випадкових величин, що дозволяє проводити аналіз. Добре відомі підходи включають моделі Маркова, системи масового обслуговування та імітаційне моделювання. Їх можна використовувати для аналізу часу очікування, надійності, використання тощо у контексті стохастичних процесів.

- *Методи оптимізації* спрямовані на надання «кращої» альтернативи (наприклад, найдешевшої чи найшвидшої) з великого або нескінченного набору альтернатив. Розглянемо, наприклад, наступне питання: за наявності переліку міст та відстаней між кожною парою міст, який можливий найкоротший маршрут, що проходить через кожне місто рівно один раз і повертається у вихідне місто? Для максимально ефективною відповіді на подібні питання було розроблено безліч методів оптимізації. Добре відомі підходи включають лінійне програмування (Linear Programming), цілочисельне лінійне програмування (Integer Linear Programming, ILP), задачі виконання обмежень (Constraint Satisfaction Problem) та динамічне програмування (Dynamic Programming).

- *Управління операціями й дослідження операцій* пов'язані з проектуванням, контролем та управлінням продуктами, процесами,

послугами та ланцюжками постачань. *Дослідження операцій* зосереджується на аналізі математичних моделей. *Управління операціями* ближче до промислового інжинірингу та бізнес-адмініструванню.

- *Управління бізнес-процесами* – це дисципліна, яка поєднує підходи до розробки, виконання, контролю, вимірювання та оптимізації бізнес-процесів. Зусилля з управління бізнес-процесами (BPM), як правило, акцентовані на явних моделях процесів (наприклад, мережі Петрі або моделі BPMN), які описують потік управління, а також інші аспекти процесів (ресурси, дані, функції тощо).

- *Інтелектуальний аналіз процесів* використовується для визначення моделей процесів з урахуванням даних про події. Відтворюючи ці дані, можна виявити вузькі місця та наслідки недотримання вимог. Порівняно з основними підходами BPM, основна увага приділяється не моделюванню процесів, а використанню даних про події. Терміни Workflow Mining (WM), Business Process Intelligence (BPI) і Automated Business Process Discovery (ABPD) використовуються для позначення орієнтованих на процеси підходів, заснованих на даних.

- *Удосконалення бізнес-процесів* – це загальний термін для багатьох підходів, спрямованих на покращення процесів. Прикладами є загальне управління якістю (TQM), Kaizen, (Lean) Six Sigma, теорія обмежень (TOC) та реінжиніринг бізнес-процесів (BPR). Насправді, можна стверджувати, що вся наука про процеси спрямована на удосконалення процесів.

- *Автоматизація процесів та управління робочими процесами* зосереджені на розробці інформаційних систем, що підтримують операційні бізнес-процеси, включаючи маршрутизацію та розподіл роботи. Система управління робочими процесами (WFM) керується моделлю, тобто моделі процесу достатньо для налаштування інформаційної системи та запуску процесу. В результаті, змінити процес можна, змінивши відповідну модель процесу.

- *Формальні методи та теорія паралелізму* засновані на теоретичній інформатиці, зокрема, логічних обчисленнях, формальних мовах, теорії автоматів та семантики мов програмування. Формальні

методи використовують низку мов для опису процесів. Прикладами є системи переходів, мережі Петрі, обчислення процесів, такі як CSP, CCS та π -числення, темпоральна логіка (Linear Temporal Logic, LTL та Computation Tree Logic, CTL) та діаграми станів. Інструменти перевірки моделей, такі як SPIN, можна використовувати для перевірки логічних властивостей, наприклад, відсутність взаємоблокувань. Паралелізм ускладнює аналіз, але також необхідний: елементи процесу чи системи можуть виконуватися одночасно та потенційно взаємодіяти один з одним. Мережі Петрі були одним з перших способів формалізації для моделювання та аналізу паралельних процесів. Багато підходів до BPM, WFM та інтелектуального аналізу процесів засновані на подібній формалізації.

Інтелектуальний аналіз процесів можна розглядати як засіб подолання розриву між наукою про дані та наукою про процеси (рис. 3, [2]).

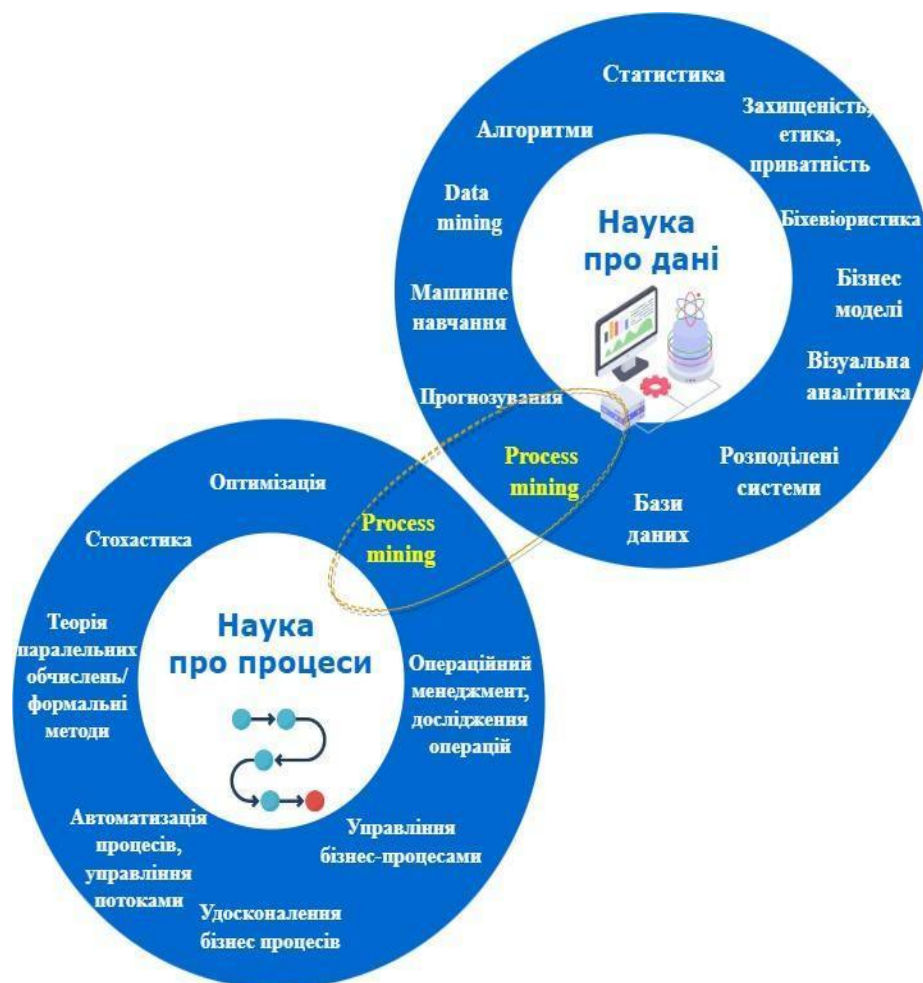


Рисунок 3 – Інтелектуальний аналіз процесів(process mining) як зв'язок між наукою про дані та наукою про процеси

Тобто, інтелектуальний аналіз процесів шукає зіставлення між даними про події (тобто поведінкою, що спостерігається) і моделями процесів (створеними вручну або виявленими автоматично). Інтелектуальний аналіз даних, статистика та методи машинного навчання не враховують повномірні й багатофункціональні моделі процесів. Підходи науки про процеси орієнтовані на процес, але часто зосереджені на моделюванні, а не на вивченні даних про події. Унікальне положення складової науки про процеси – інтелектуального аналізу процесів – робить її потужним інструментом для використання зростаючої доступності даних і удосконалення процесів на всіх його етапах.

Існують різні нотації моделювання процесів: модель станів і переходів (transition system), мережі Петрі, мережі робочих процесів (WF-nets), YAWL, BPMN, EPCs, причинно-наслідкові мережі (Causal nets), дерево процесів (Process Trees). Відносно легко автоматично перевести результати аналізу процесу в потрібну нотацію. Наприклад, α -алгоритм генерує мережу Петрі, яку легко перетворити в модель BPMN, модель BPEL або діаграму активності UML.

II. ЗАСТОСУВАННЯ ТЕОРІЇ МЕРЕЖ ПЕТРІ ДЛЯ МОДЕЛЮВАННЯ ПРОЦЕСІВ

Мережі Петрі є найстарішою та найкраще розвиненою мовою моделювання паралельних процесів. Графічні позначення мережі Петрі є інтуїтивно зрозумілими та простими (рис. 4), а також виконуваними, і для їх аналізу можна використовувати різні методи. Мережа Петрі – це двочастковий граф, що складається з позицій і переходів. Структура мережі є статичною, але, для мережі можна використати маркери (іноді можна зустріти інші назви – токени або мітки), які будуть переміщатися мережею згідно із заданими правилами. Стан мережі Петрі визначається розподілом маркерів на позиціях і називається її маркуванням. У початковому маркуванні, показаному на рис. 4, є лише один маркер на стартовій позиції, хоча їх можна задати в необмеженій кількості в різних позиціях. Крім того, у початковій позиції (позиціях)

маркери можуть з'являтися із заданим інтервалом. На рисунку 4 за допомогою мережі Петрі зображений процес отримання, обробки і прийняття рішення щодо звернення клієнта для отримання компенсації за повернений залізничний квиток.

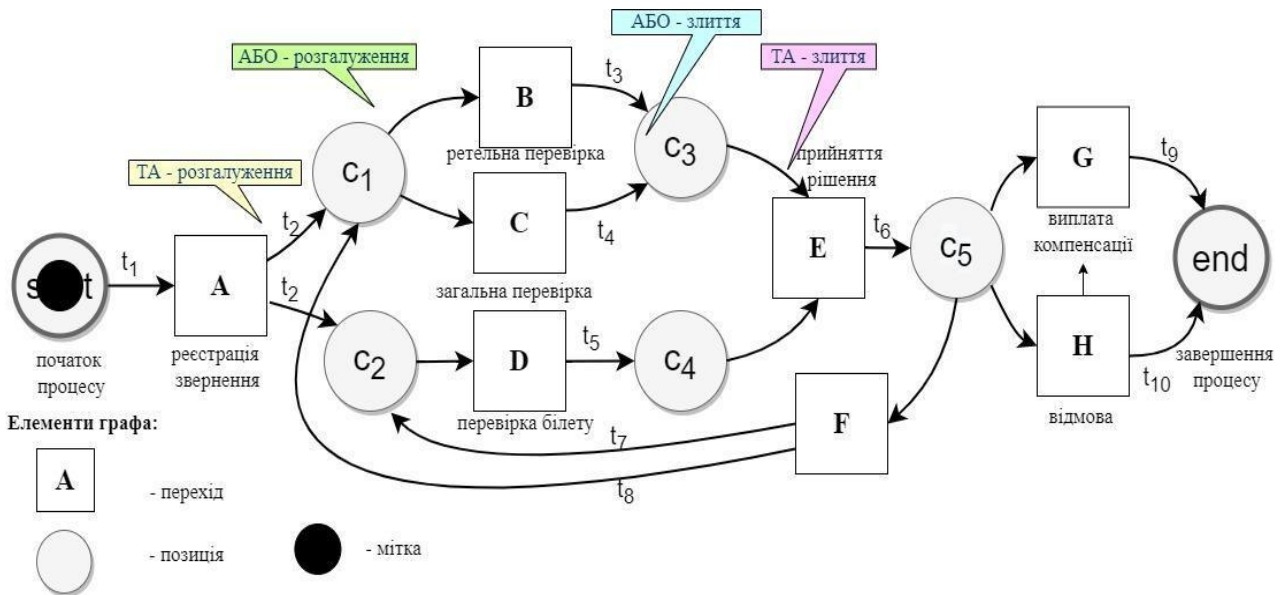


Рисунок 4 – Приклад мережі Петрі з початковим маркуванням

В даному випадку процес починається із реєстрації звернення клієнта (A) в електронну систему продажу залізничних квитків для повернення раніше придбаного квитка і отримання грошової компенсації. Далі відбувається ретельна (B) або стандартна (C) перевірка як ситуації в цілому, так і дійсність квитка (D). Після обох перевірок й отримання необхідних підстав приймається рішення (E) або відправити справу на доопрацювання з уточненням даних, або виплатити компенсацію, або відмовити.

Формалізованою мовою мережа Петрі – це триплет $N = (P, T, F)$, де P – скінченна множина позицій, T – скінченна множина переходів, і $F \subseteq (P \times T) \cup (T \times P)$ – набір спрямованих дуг. Промаркована мережа Петрі – це пара (N, M) , де $N = (P, T, F)$ – мережа Петрі, а $M \in B(P)$ – мультимножина на множині P , що позначає маркування мережі.

Відповідно, мережу Петрі, показану на рис. 4, можна формалізувати наступним чином: позиції $P = \{start, c_1, c_2, c_3, c_4, c_5, end\}$,

переходи $T = \{A, B, C, D, E, F, G, H\}$ і направлені дуги $F = \{(start, A), (A, c_1), (A, c_2), (c_1, B), (c_1, C), (c_2, D), (B, c_3), (C, c_3), (D, c_4), (c_3, E), (c_4, E), (E, c_5), (c_5, F), (F, c_1), (F, c_2), (c_5, G), (c_5, H), (G, end), (H, end)\}$. Початковим маркуванням, показаним на рис. 4, є $[start]$, тобто набір, що містить лише один маркер (мітку). Динамічна поведінка такої промаркованої мережі Петрі визначається наступним правилом спрацьовування: *перехід увімкнений, якщо кожна з його вхідних позицій містить маркер*. Увімкнений перехід може запускатися, споживаючи один маркер з кожної позиції входу і створюючи один маркер для кожної вихідної позиції (якщо над дугами не вказана їх кількість). Отже, у вищенаведеному прикладі перехід А вмикається при позначенні $[start]$. Запуск А призводить до маркування $[c_1, c_2]$, тобто споживається один маркер, а виробляються два. При позначенні $[c_1, c_2]$ перехід А більше не вмикається. Тепер стали активними переходи В, С і D. З позначення $[c_1, c_2]$, запуск В призводить до маркування $[c_2, c_3]$. Тут D все ще увімкнений, але В і С – більше ні. Через конструкцію циклу, що включає F, існує багато послідовностей запуску, які починаються з $[start]$ і закінчуються $[end]$.

Мультимножина (ще зустрічається «сумка» чи комплект) – набір елементів, в якому кожен елемент може зустрічатися кілька разів. Наприклад, $[A^1, B^2, C^3, D^2, E^1]$ – це мультимножина з дев'ятьма елементами: один А, два В, три С, два D і один Е.

Припустимо, що початкове маркування – $[start^5]$. Запуск А призведе до позначення $[start^4, c_1^1, c_2^1]$. При цьому позначення А все ще увімкнене, оскільки досі ще має 4 мітки. Повторний запуск А призводить до позначення $[start^3, c_1^2, c_2^2]$. Перехід А може спрацювати п'ять разів поспіль, що призводить до позначення $[c_1^5, c_2^5]$. Після першої активації А також активуються В, С і D, які можуть запускатися одночасно.

Для формалізації правила спрацьовування введемо позначення вхідних (вихідних) позицій (переходів). Нехай $N = (P, T, F)$ – мережа Петрі. Елементи $P \cup T$ називаються вузлами. Вузол x є вхідним вузлом іншого вузла y тоді і тільки тоді, коли існує спрямована дуга від x до y (тобто $(x, y) \in F$).

Якщо модель містить велику кількість паралельних дій або кілька маркерів знаходяться в одному місці, то система переходів набагато більша, ніж мережа Петрі. Насправді, промаркована мережа Петрі може мати нескінченну кількість доступних станів.

Так, позначена мережа Петрі на рис. 5 (а) складається лише з однієї позиції та одного переходу. Тим не менш, відповідна перехідна система має нескінченно багато станів. У цьому прикладі перехід t постійно ввімкнено, оскільки він не має місця введення. Тому він може помістити будь-яку кількість маркерів у позицію p . Мережа Петрі на рис. 5 (b) має дві дуги, а не одну, і тепер єдиним доступним станом є $[p]$. Позначена мережа Петрі на рис. 5 (c) показує ефект паралельності. Відповідна система переходів має $6^5 = 7776$ станів і 32 400 переходів.

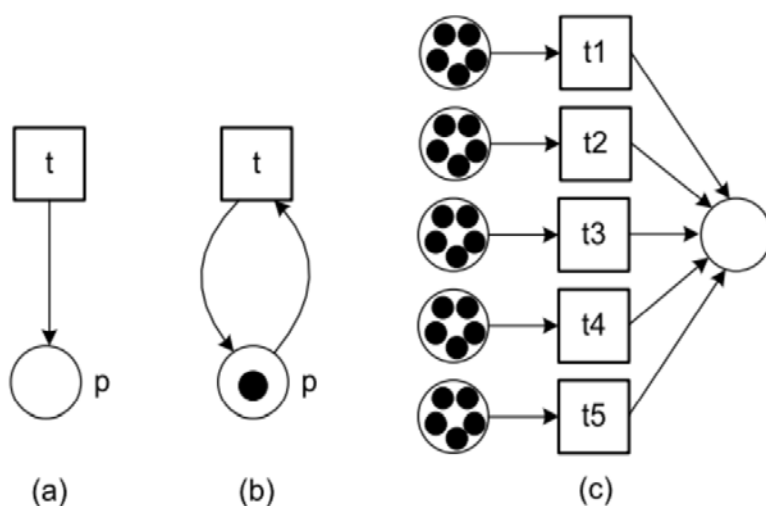


Рисунок 5 – Стани мережі Петрі: (а) із нескінченним простором станів; (b) лише з єдиним досяжним станом; (c) із 7776 досяжними станами

На рис. 6 показано, яким чином модель на рис. 4 може бути розширена за допомогою додавання іншої інформації (наприклад, тривалість події та/або експертність виконавця операції).

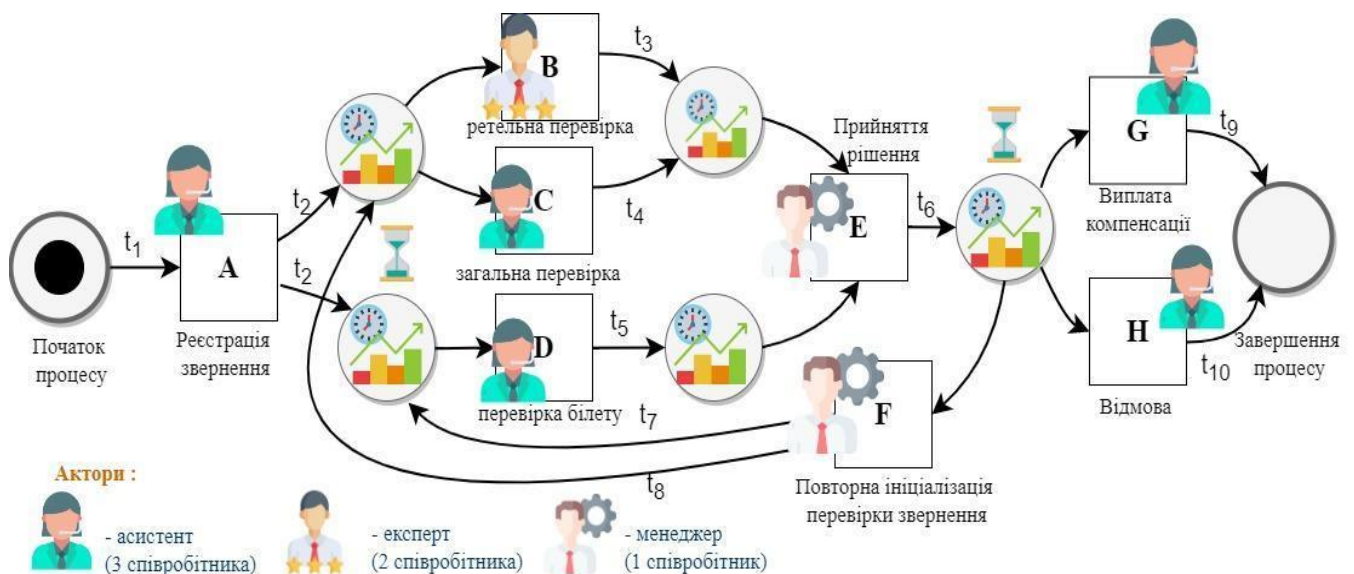


Рисунок 6 – Розширена модель процесу виплати компенсації за повернення залізничного квитка

Модель процесу розширюється на цьому прикладі (рис. 6) наступними додатковими аспектами:

- ❖ з точки зору організації – які організаційні ролі та які ресурси забезпечують певну діяльність;
- ❖ з точки зору події – які характеристики події вплинули на конкретне рішення;
- ❖ з точки зору витрат часу – де вузькі місця в процесі;

Журнал подій може визначити ролі в організації (співробітники з подібними повноваженнями). Ці ролі можна використати, щоб зв'язати конкретних людей із певною роботою. З журналу подій також можна витягнути й візуалізувати інформацію про характеристики процесу (наприклад, мінімальний чи середній час між двома послідовними діями), а також правила й умови прийняття певних рішень під час реалізації технологічного процесу. Журнал подій може бути сформований для різних цілей:

1. *Перевірка відповідності*: розбіжності між журналом і моделлю можна виявити та визначити кількісно шляхом повторного відтворення журналу. Наприклад, відтворення певної поведінки на моделі, показує, що окремий необхідний етап періодично пропускається.

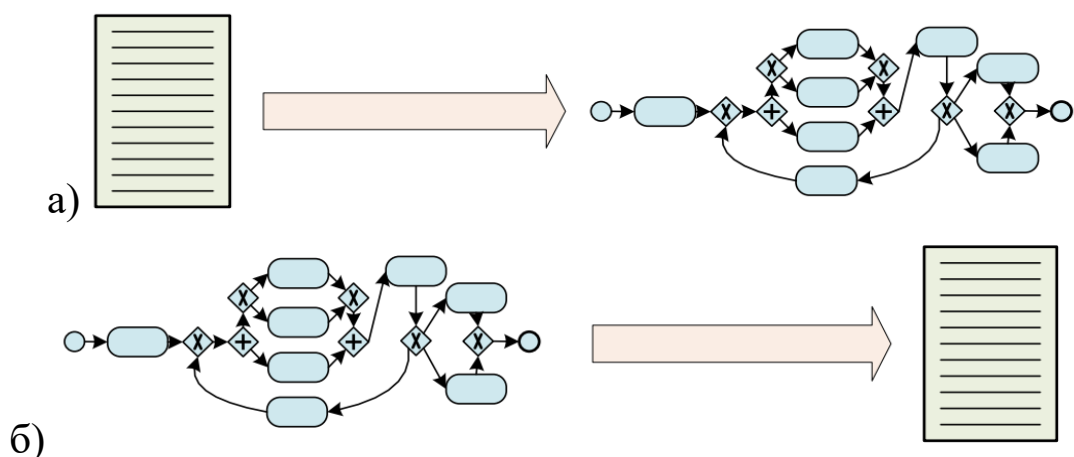
2. *Розширення моделі для визначення частоти:* відтворюючи журнал, можна побачити, які частини моделі відвідують частіше. Такий спосіб моделювання можна використовувати для виявлення «вузьких місць» процесу.

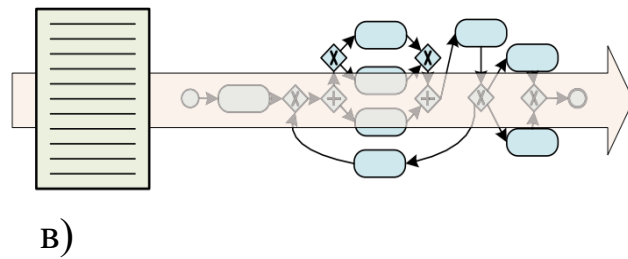
3. *Тимчасове виділення окремих підпроцесів із включенням додаткової інформації і виключенням нерелевантної інформації:* відтворюючи тільки проблемні ділянки моделі, які масштабуються за рахунок додавання тимчасової інформації, можна більш коректно визначити причини виникнення "вузьких місць".

4. *Побудова прогнозних моделей.* Відтворюючи журнали подій, для різних станів моделі можна робити конкретні передбачення. Наприклад, прогнозна модель, отримана шляхом відтворення багатьох випадків, може показати, що очікуваний час до завершення після ввімкнення стану E (рис. 6) становить 5 годин.

5. *Оперативна підтримка.* Відтворення не обмежується історичними даними про події. Можна також відтворити часткові сліди запусчених справ для виявлення відхилень під час виконання. Отже, може бути згенероване попередження про відхилення справи до завершення процесу. Так само можна передбачити час обробки, що залишився, або ймовірність відхилення справи. Такі прогнози також можна використовувати для рекомендації наступних кроків щодо удосконалення процесу.

Способи поєднання журналів подій (або інших джерел інформації, що містять приклад поведінки) і моделей процесів наведено на рис.7.





Діагностика
 Прогнозування
 Рекомендації
 Частота виконання
 певної послідовності
 Оперативна підтримка

Рисунок 7 – Способи поєднання журналів подій і моделей процесів:
 а) побудова моделі на основі прикладів поведінки (журналу подій);
 б) створення поведінки на основі побудованої моделі; в) використання журналу подій і моделі процесу як вхідних даних для аналізу процесу

Основна ідея моделювання на основі журналу подій полягає в тому, щоб багаторазово запускати модель і таким чином збирати статистичні дані та довірчі інтервали. В останньому випадку (рис. 7, в) журнал подій відтворюється поверх моделі процесу.

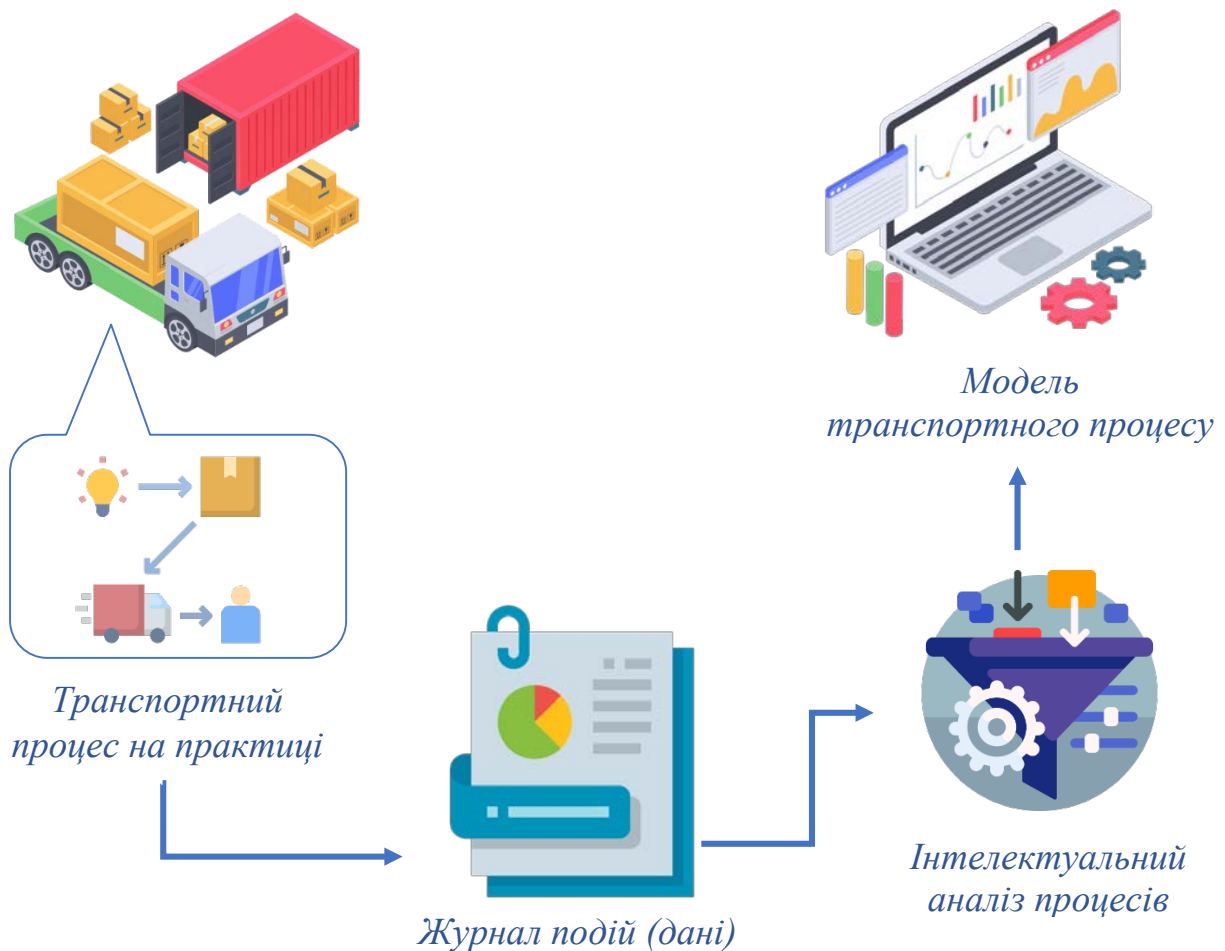


Рисунок 8 – Схема використання журналу подій для моделювання транспортних процесів

В результаті використання журналу подій під час моделювання процесів (рис. 8) можна вирішити наступні проблеми:

- ✓Виявлення неефективних операцій;
- ✓Демонстрація відповідності моделі реальному процесу;
- ✓Затримки, які виникають під час процесу;
- ✓Існування відхилень від нормованих значень параметрів процесу;
- ✓Чіткий аналіз деталей процесу;
- ✓Контроль за виконанням поставлених цілей (наприклад, останній етап процесу має завершитись через 5 діб).

Мережі Петрі мають міцну теоретичну основу та можуть добре враховувати паралелізм. Крім того, існує широкий спектр потужних методів та інструментів аналізу [5]. Очевидно, що ця спрощена модель має проблеми із врахуванням аспектів, пов'язаних з даними та часом. Тому було запропоновано різні типи мереж Петрі високого рівня. Кольорові мережі Петрі (CPN) є найбільш широко використовуваним формалізмом на основі мереж Петрі, який може мати справу з аспектами, пов'язаними з даними та часом [6]. Маркери CPN несуть значення даних (різні типи даних візуалізують через кольори). Позначка часу вказує ранній час, коли маркер (мітка) може бути використаний. Переходи можуть призначати затримку для маркерів. Таким чином можна змодельовати час очікування та обслуговування. CPN може бути ієрархічною, тобто переходи можуть бути розкладені на підпроцеси. Таким чином, можна структурувати великі моделі. CPN Tools – це набір інструментів, який забезпечує підтримку моделювання та аналізу CPN.

Інтелектуальний аналіз процесів на основі теорії мереж Петрі також використовується в транспортній галузі. Широкий спектр міжнародних дослідницьких компаній пропонують удосконалювати технологічні процеси в різних напрямках діяльності логістичних, транспортних, посередницьких компаній і державних установ.

Якщо мільйони людей/вантажовласників щодня користуються різними видами транспорту, таким компаніям важко проаналізувати і зробити висновки про те, як найкраще їх підтримати. Process Mining може допомогти транспортним агентствам і організаторам дорожнього руху, наприклад,:

- визначити оптимальний розподіл між видами транспорту
- визначити кращі місця для паркування
- моделювати зміни моделей подорожей
- оптимізувати автобусні маршрути
- звести до мінімуму потребу в зміні транспортних засобів або видів транспорту.

Транспортні компанії можуть визначити закономірності транспортних процесів та основні причини:

- порожнього пробігу,
- часу простою транспортних засобів,
- інших затримок,
- збільшення використання ресурсів.

Індивідуальні мандрівники також можуть відстежувати свої власні подорожі та отримати вигоду, наприклад, від:

- пропозицій щодо спільного використання автомобілів, якщо дані вказують на те, що інші мандрівники можуть підвезти;
- попереджень про зайнятість, щоб уникнути заторів або переповнених транспортних засобів громадського транспорту під час пандемії COVID-19;
- пропозицій щодо маршрутів, які залежать від багатьох критеріїв, таких як швидкість, безпека, зручність, вплив на клімат;
- пропозицій щодо інтермодальних маршрутів з урахуванням доступності видів транспорту на основі даних попередніх поїздок. Наприклад, у Google Maps фіксується інформація, що мандрівник використав велосипед для подолання «останньої милі» до пункту призначення, тому на зворотному шляху також пропонує скористатись велосипедом.

III. АНАЛІЗ СУЧАСНОГО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ РОЗРОБКИ ТЕХНОЛОГІЧНИХ ПРОЦЕСІВ З ВИКОРИСТАННЯМ МЕРЕЖ ПЕТРІ

Для реалізації інтелектуального аналізу процесів на сьогодні існує близько 40 програмних продуктів, представлених в різних регіонах світу у комерційних, академічних й відкритих версіях. Вивчення можливостей цих інструментів, а також наявність версій з відкритим кодом дозволяє обрати необхідне програмне забезпечення для моделювання різноманітних транспортних процесів.

Огляд поширеного програмного забезпечення для моделювання процесів з використанням інтелектуального аналізу процесів наведений в таблиці 1 [5].

Таблиця 1 – Коротка характеристика програмного забезпечення, доступного для користувачів у сфері інтелектуального аналізу технологічних процесів

Програмне забезпечення	Версії	Коротка характеристика
<p style="text-align: center;">Apromore</p> 	<p>комерційна академічна <i>відкрита</i></p>	<p>Програмне забезпечення з відкритим кодом створено експертами в області процесного майнінгу у 2009 року. Переваги: інтуїтивно зрозумілий дизайн, розширена функціональність процесів і підключення до різноманітних джерел даних.</p>
<p style="text-align: center;">ARIS Process Mining</p> 	<p>комерційна академічна</p>	<p>ARIS Process Mining автоматично відтворює реальний процес на основі даних про процес. Візуальні компоненти та параметри вибору дозволяють аналізувати процес під різними кутами.</p>
<p style="text-align: center;">bupaR</p> 	<p><i>відкрита</i></p>	<p>Інтегрований пакет з відкритим вихідним кодом для обробки та аналізу даних бізнес-процесів. bupaR забезпечує підтримку різних етапів аналізу процесу, таких як імпорт і попередня обробка даних про події, статистичні обчислення, візуалізація процесу та перевірка відповідності.</p>

Продовження таблиці 1

Програмне забезпечення	Версії	Коротка характеристика
<p>BusinessOptix</p> 	академічна	<p>Моделювання та аналіз процесів, сценарії трансформації, дошки Kanban, документація. BusinessOptix надає інформацію, необхідну для виявлення та визначення процесів, визначення можливостей для вдосконалення та моделювання майбутніх процесів. Дані, отримані BusinessOptix, сприяють прийняттю кращих рішень і забезпечують впевненість у тому, що майбутні вдосконалення процесу базуються на емпіричних даних.</p>
<p>Celonis</p> 	комерційна академічна	<p>Програмне забезпечення для аналізу й підвищення продуктивності бізнес процесів з використанням методів інтелектуального аналізу процесів, системи управління виконанням (EMS)</p>
<p>Datricks</p> 	комерційна	<p>Використовуючи технологію штучного інтелекту, яка розкриває раніше невідомі ризики, фінансові компанії можуть ретельніше проводити.</p>
<p>DCR Portal</p> 	комерційна академічна <i>відкрита</i>	<p>Динамічне реагування на умови (DCR) – це графічна нотація процесу на основі обмежень для адаптивного керування справами. На відміну від традиційних нотацій моделювання бізнес-процесів, які зазвичай засновані на потоках, що описують переходи між станами, що викликаються подіями, DCR натомість зосереджується на зборі правил бізнесу та відповідності, що обмежують порядок подій. Таким чином, DCR відображає логіку процесу, а не просто кілька можливих шляхів, які практично рідко відповідають реальності.</p>
<p>Disco</p>	комерційна академічна	<p>Програмне забезпечення для моделювання мереж Петрі, BPMN</p>
<p>eDromos</p> 	комерційна академічна <i>загально-доступна</i>	<p>Програмне забезпечення документує, моделює та прогнозує бізнес-процеси для підвищення продуктивності, зниження витрат, кращого розуміння неефективності.</p>

Продовження таблиці 1

Програмне забезпечення	Версії	Коротка характеристика
<p>EverFlow</p> 	комерційна	Інструмент технологічного майнінгу, який дозволяє аналізувати процес у реальному житті, виявляти вузькі місця, обґрунтовувати заходи подолання різних типів неефективності.
<p>Explora Process</p> 	комерційна	Збирання та інтерпретація даних для прогнозування бізнес-можливостей та формування нових підходів, орієнтованих на клієнтів.
<p>FortressIQ</p> 	комерційна	Платформа виявлення процесів. Об'єднує комп'ютерний зір, обробку природної мови (NLP), штучний інтелект (AI) та машинне навчання (ML), щоб автоматично виявляти, відображати, моделювати та документувати всі процеси та завдання, що виконуються на підприємстві співробітниками організації.
<p>Kofax</p> 	комерційна	Поєднання моніторингу та аналізу процесів, з візуалізацією, аналітикою та інтеграцією даних в єдине рішення для наскрізної демонстрації продуктивності та відповідності.
<p>Livejourney</p> 	комерційна	Зв'язування даних зі своїми ІТ-системами, візуалізація того, що насправді відбувається, отримання сповіщення в режимі реального часу про відхилення процесу, аналіз першопричин, моделювання майбутніх шляхів за допомогою модуля прогнозування, вимірювання рентабельності інвестицій та ефективності змін.

ProM (Process Mining framework) – це розширюваний фреймворк з відкритим вихідним кодом, який підтримує широкий спектр методів формування процесів у формі плагінів. Він не залежить від платформи, оскільки реалізований на Java. Розробники постійно підтримують практичні застосування ProM, а також запрошуюють дослідників і розробників зробити власний внесок у вигляді нових плагінів. [7] Структура ProM (рис. 9):

- *Об'єкти даних (1)*, які розміщені на вкладці «workspace» (2),
- *Плагіни*, доступ до яких можна отримати через вкладку «run» (3),
- *Візуалізатори*, доступ до яких можна отримати через вкладку «visualizer» (4).

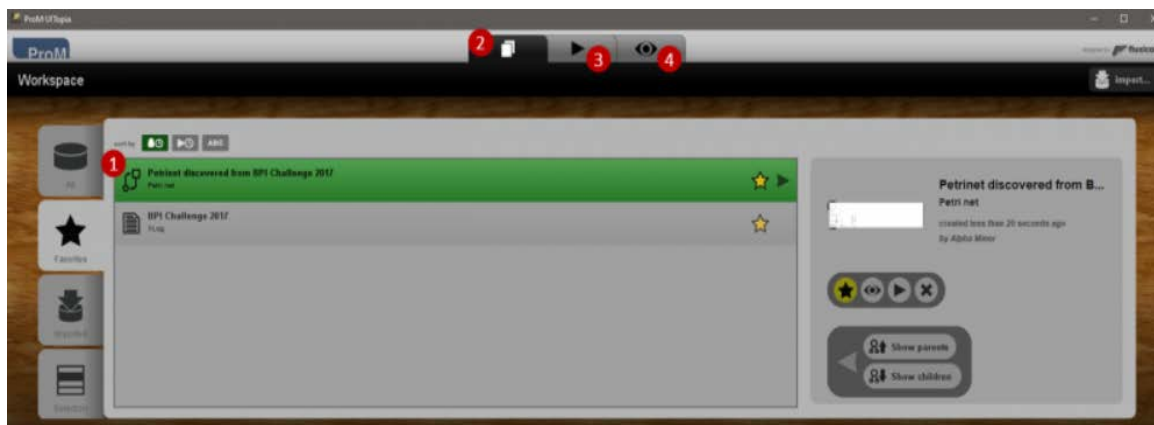


Рисунок 9 – Структура ProM

Об'єкти даних, плагіни та візуалізатор організовані в пакети, які встановлюються за допомогою менеджера пакетів. Об'єкти складаються з журналу подій і моделей процесів. ProM надає кілька типів моделей процесів: мережі Петрі та діаграми BPMN. Плагіни створені для фільтрування журналу подій, виявлення процесів, підтвердження відповідності, удосконалення моделей. Візуалізатори створені для всіх типів об'єктів.

Інструмент RapidMiner дозволяє виконувати робочі процеси у зв'язці з фреймворком ProM6. Таким чином, будь-який алгоритм виявлення, встановлення відповідності або розширення ProM6 може використовуватися RapidMiner в процесі інтелектуального аналізу даних.

Інструмент моделювання Disco виконує наступні функції (рис. 10):

- 1) Автоматичне виявлення процесу.
- 2) Анімація технологічної карти.
- 3) Перегляд статистики в Disco із використанням інтерактивних діаграм і докладної інформації про кожну діяльність, ресурс і значення атрибута.

- 4) Перегляд повної історії відповідних кейсів із визначенням стандартних і нестандартних випадків.
- 5) Фільтрування даних для зосередження на певній проблемі.
- 6) Управління проектами.

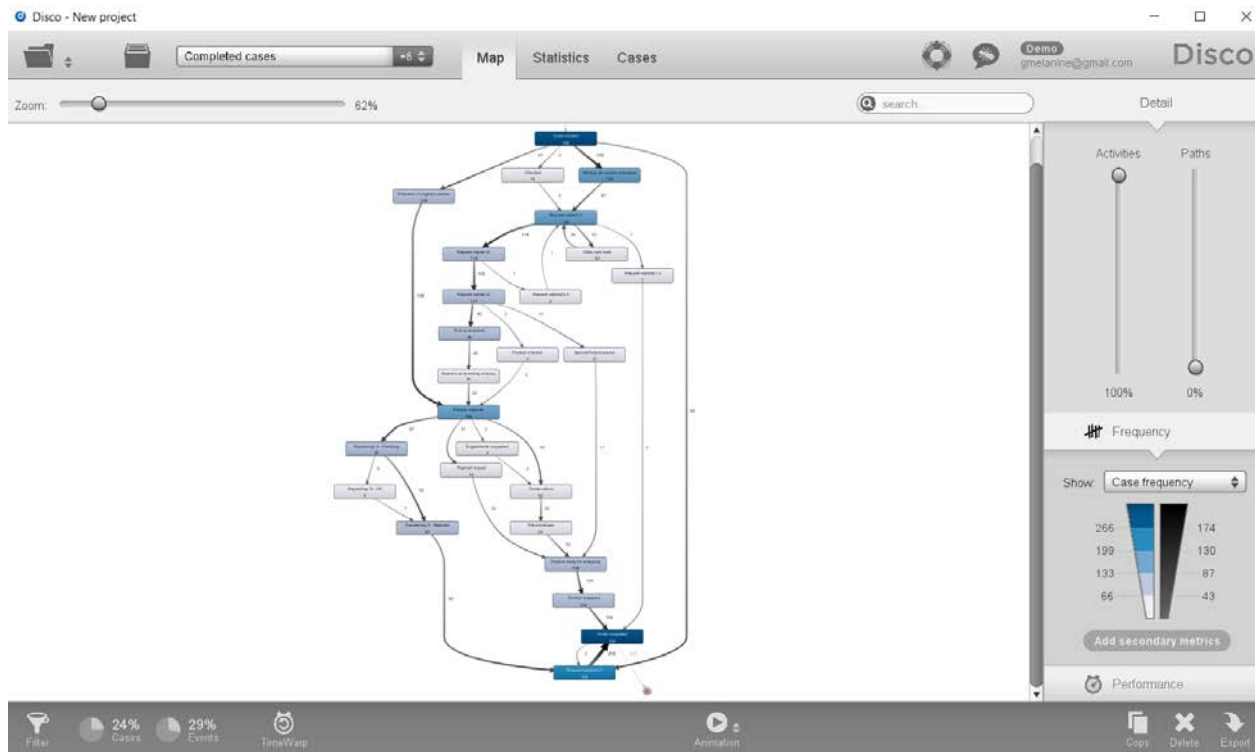


Рисунок 10 – Робоче вікно Disco

Отже, інструменти й нотації інтелектуального аналізу процесів надають можливість бізнес-аналітикам, менеджерам, та виконавчим директорам широкі можливості щодо підвищення ефективності бізнес процесів в організації, а також:

- ✓ виявлення неефективних операцій;
- ✓ демонстрація відповідності моделі реальному процесу;
- ✓ витрати ресурсів під час реалізації процесу;
- ✓ існування відхилень від нормованих значень параметрів процесу;
- ✓ чіткий аналіз деталей процесу;
- ✓ контроль за виконанням поставлених цілей;
- ✓ прогнозування та ін.

Існують різні нотації моделювання процесів: модель станів і переходів (transition system), мережі Петрі, мережі робочих процесів (WF-nets), YAWL, BPMN, EPCs, причинно-наслідкові мережі (Causal nets), дерево процесів (Process Trees). Відносно легко перевести результати моделювання з однієї нотації в іншу.

Надалі для моделювання транспортних процесів була обрана нотація мереж Петрі, яка має всі вище перелічені функції. Заснована в 40-х роках минулого століття, ця теорія має розвиток й практичне застосування для різних сфер діяльності людини – там, де складні процеси потребують постійного моніторингу, ретельного аналізу, прогнозування ефективності й результативності їх виконання.

Вивчення можливостей інструментів, а також наявність версій з відкритим кодом обумовило вибір програмного забезпечення для подальшого моделювання транспортних процесів - ProM, Disco [7].

IV. ДОСЛІДЖЕННЯ ПАРАМЕТРІВ ТРАНСПОРТНИХ ПРОЦЕСІВ НА ОСНОВІ МЕРЕЖ ПЕТРІ

Для прикладу виділимо з транспортно-технологічної схеми всього процесу доставки вантажів автомобільним транспортом одну ділянку цього процесу – переміщення товарів на транспортному засобі в режимі імпорту від міжнародного пункту пропуску на території України до митниці призначення.

Важливим етапом доставки вантажів у міжнародному сполученні є митне оформлення вантажів і транспортних засобів, що перетинають митний кордон України. Розвиток митної справи України передбачає інтенсифікацію автоматизації процесу документального оформлення зовнішньоторговельних угод, у тому числі для митних цілей за допомогою Єдиної Автоматизованої Інформаційної Системи (ЄАІС) митних органів України. Спрощена процедура митного оформлення товарів, що переміщують від митниці відправлення на міжнародному пункті пропуску України (МПП) до митниці призначення, наведена на рисунку 11.

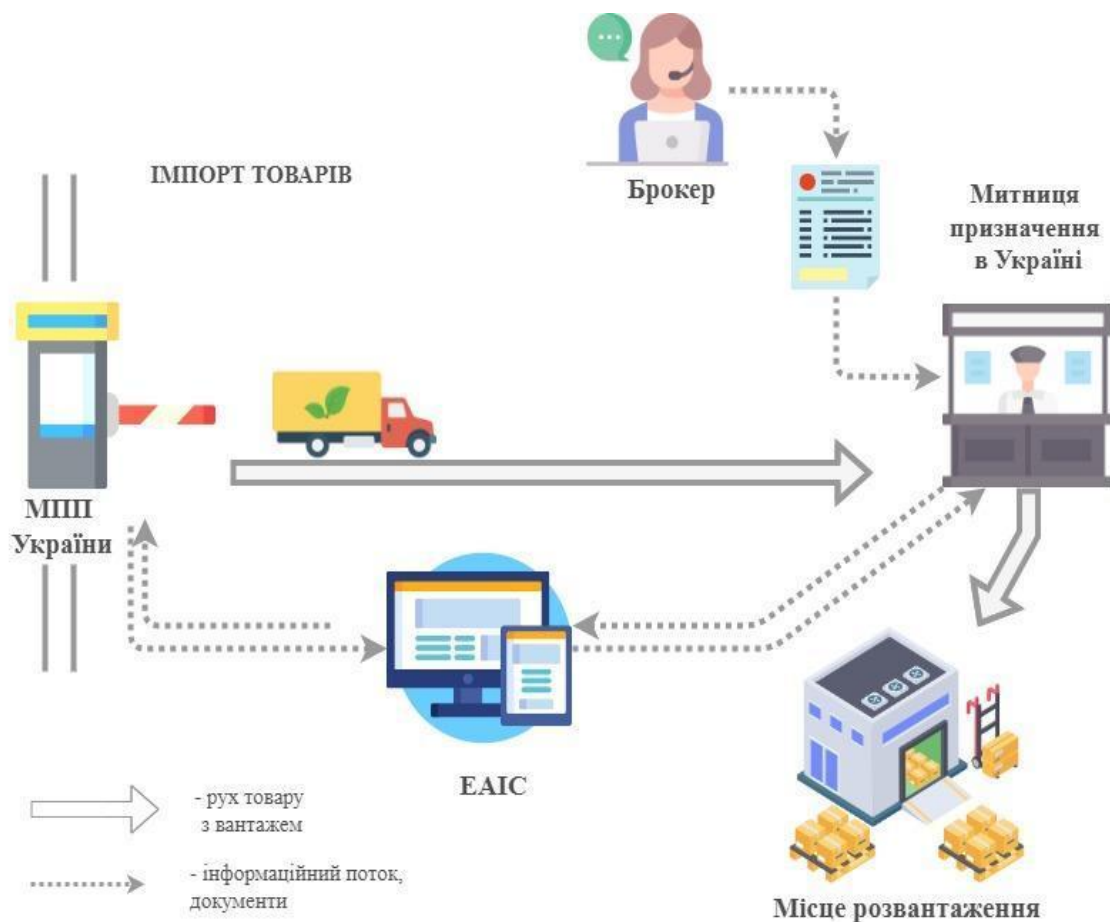


Рисунок 11 – Спрощена схема контролю за переміщенням товарів на митній території України в режимі імпорту

Від ефективності процедур оформлення товарів залежить ефективність всього процесу доставки, успішність завершення зовнішньоторговельних угод, і, відповідно, привабливість ведення бізнесу в Україні.

Основний параметр для дослідження – середня тривалість виконання операцій в рамках вищенаведеного процесу. На основі зібраних точних і приблизних характеристик про окремі операції для 7 однотипних кейсів був сформований журнал подій (рис. 12, [8]). У такому журналі ідентифікаторами визначені код кейсу, назва операції в кейсі, час початку і завершення виконання даної операції, виконавець операції і його/її експертність.

Case	Activity	StartTime	EndTime	EmployeeRole	Employee
1	Обробка запиту на декларування вантажу в режимі і...	1/12/21 10:15	1/12/21 12:25	Агент	Ганна
2	Підготовка комерційних документів	1/12/21 13:00	1/13/21 16:00	Агент	Ганна
3	Заповнення МД, електронний підпис	1/13/21 16:00	1/13/21 17:30	Агент	Ганна
4	Відправка документів і МД у митницю призначення	1/13/21 17:40	1/13/21 18:00	Агент	Ганна
5	Форматологічний контроль декларації	1/14/21 10:00	1/14/21 10:45	Інспектор	123
6	Підвѣдження прийняття МД до оформлення	1/14/21 10:45	1/14/21 11:00	Інспектор	123
7	Отримання копій товаросупровідних документів	1/12/21 17:00	1/12/21 18:00	Агент	Ганна
8	Отримання інформації про прибуття ТЗ з вантажем н...	1/13/21 13:00	1/13/21 13:15	Агент	Ганна
9	Митне оформлення ТЗ з вантажем на МПП	1/13/21 13:00	1/13/21 15:00	Інспектор МПП	222
10	Рух ТЗ з вантажем від МПП до митниці призначення	1/13/21 15:00	1/13/21 23:00	Водій	Олександр
11	Митний огляд на митниці призначення	1/14/21 13:00	1/14/21 14:00	Інспектор	123
12	Митний контроль ТЗ з вантажем на митниці призначе...	1/14/21 12:00	1/18/21 15:00	Інспектор	123
13	Санітарний контроль	1/14/21 16:00	1/14/21 16:30	Санітарний інспектор	124
14	Огляд тари, перевірка документів виробника тари	1/14/21 12:00	1/14/21 12:30	Інспектор	123
15	Подання заяви на фітосанітарний контроль	1/14/21 13:00	1/14/21 13:30	Агент	Ганна
16	Оплата фітосанітарного контролю	1/14/21 13:30	1/14/21 13:40	Агент	Ганна
17	Проведення фітосанітарного контролю, видача серти...	1/15/21 11:30	1/15/21 12:30	Фіто інспектор	125
18	Перевірка фітосанітарного сертифікату	1/15/21 12:30	1/15/21 12:40	Інспектор	123
19	Перевірка умов контракту	1/14/21 15:00	1/14/21 15:30	Інспектор	123
20	Перевірка товаросупровідних документів	1/14/21 15:30	1/14/21 16:00	Інспектор	123
21	Видача водію оформленого пакету документів	1/18/21 15:00	1/18/21 15:30	Інспектор	123
22	Виїзд ТЗ з вантажем з митниці призначення до місця...	1/18/21 15:30	1/18/21 15:40	Водій	Олександр
23	Сплата митних платежів	1/14/21 10:00	1/14/21 10:30	Агент	Ганна
24	Перевірка сплати митних платежів	1/14/21 16:30	1/14/21 17:00	Інспектор	123
25	Перевірка кодів УКТЗЕД	1/14/21 16:00	1/14/21 16:30	Інспектор	123
26	Обробка запиту на декларування вантажу в режимі і...	1/13/21 11:30	1/13/21 12:30	Агент	Віктор
27	Підготовка комерційних документів	1/13/21 14:00	1/13/21 18:00	Агент	Віктор
28	Заповнення МД, електронний підпис	1/14/21 9:00	1/14/21 11:30	Агент	Віктор
29	Відправка документів і МД у митницю призначення	1/14/21 11:30	1/14/21 12:00	Агент	Віктор
30	Форматологічний контроль декларації	1/14/21 14:00	1/14/21 15:00	Інспектор	223
31	Підвѣдження прийняття МД до оформлення	1/14/21 15:00	1/14/21 15:15	Інспектор	223

Рисунок 12 – Журнал подій для моделювання процесу декларування товарів (витяг)

На основі схеми процесу (рис. 11), журналу подій обраних для аналізу кейсів (рис. 12) і визначених параметрів транспортних процесів розроблені моделі (рис. 13, 14, 15), які дозволили дослідити проблемні випадки виконання цих процесів в ситуаціях, коли вони не є очевидними.

Складність зв'язків між етапами спрощеного для задач моделювання процесу демонструє рисунок 13. Застосунок також дає можливість спростити процес, максимально узагальнюючи процес і відкидаючи ті етапи, які зустрічаються рідше (рисунок 14).

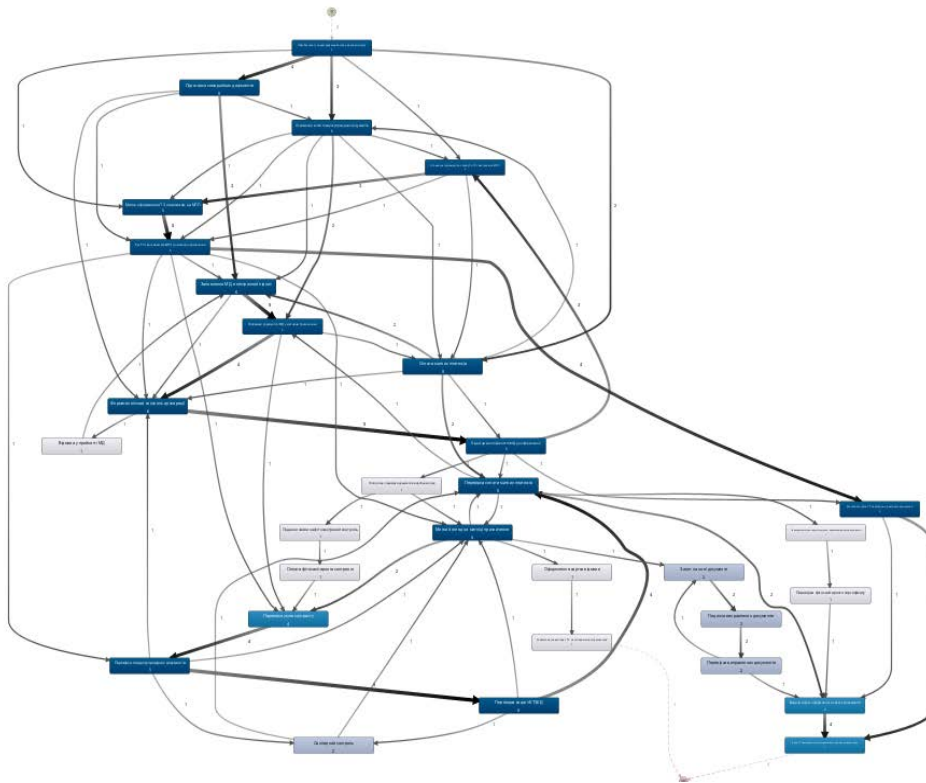


Рисунок 13 – Аналіз процесу контролю за доставкою на основі журналу подій (загальний вигляд зменшеної у розмірі карти 5 кейсів)

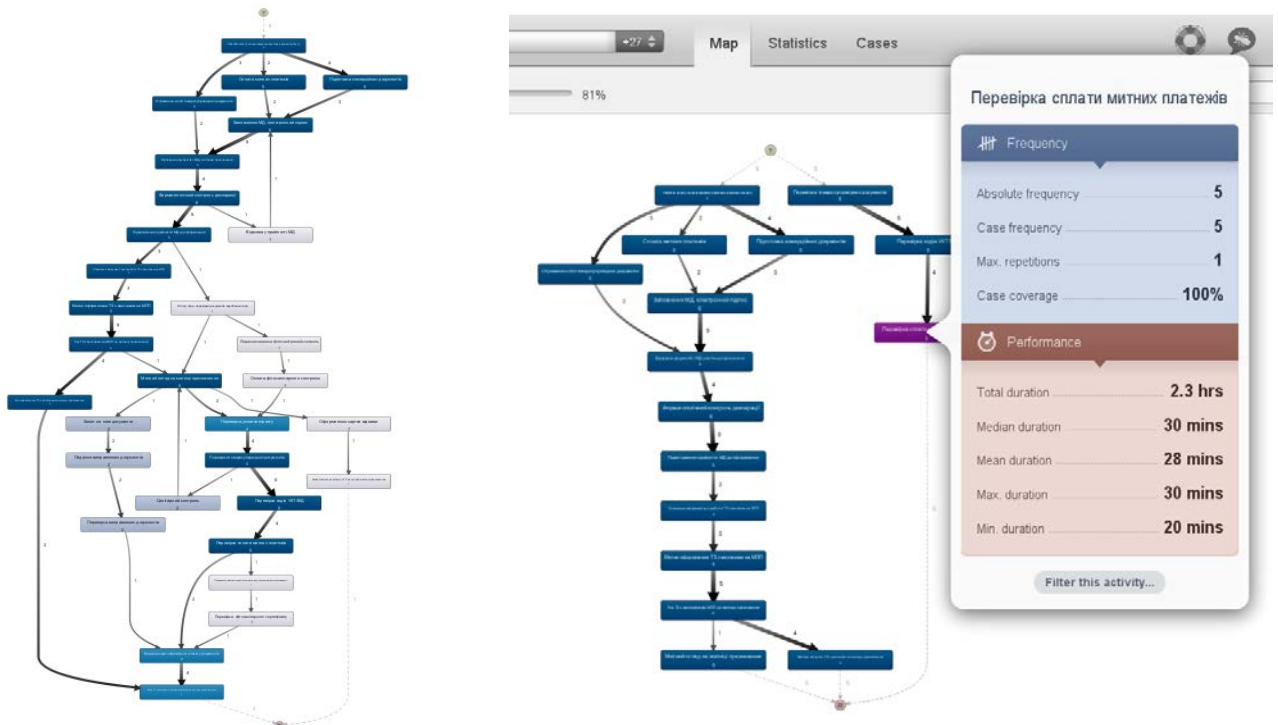


Рисунок 14 – Фільтрування процесу: події, які виконуються найчастіше

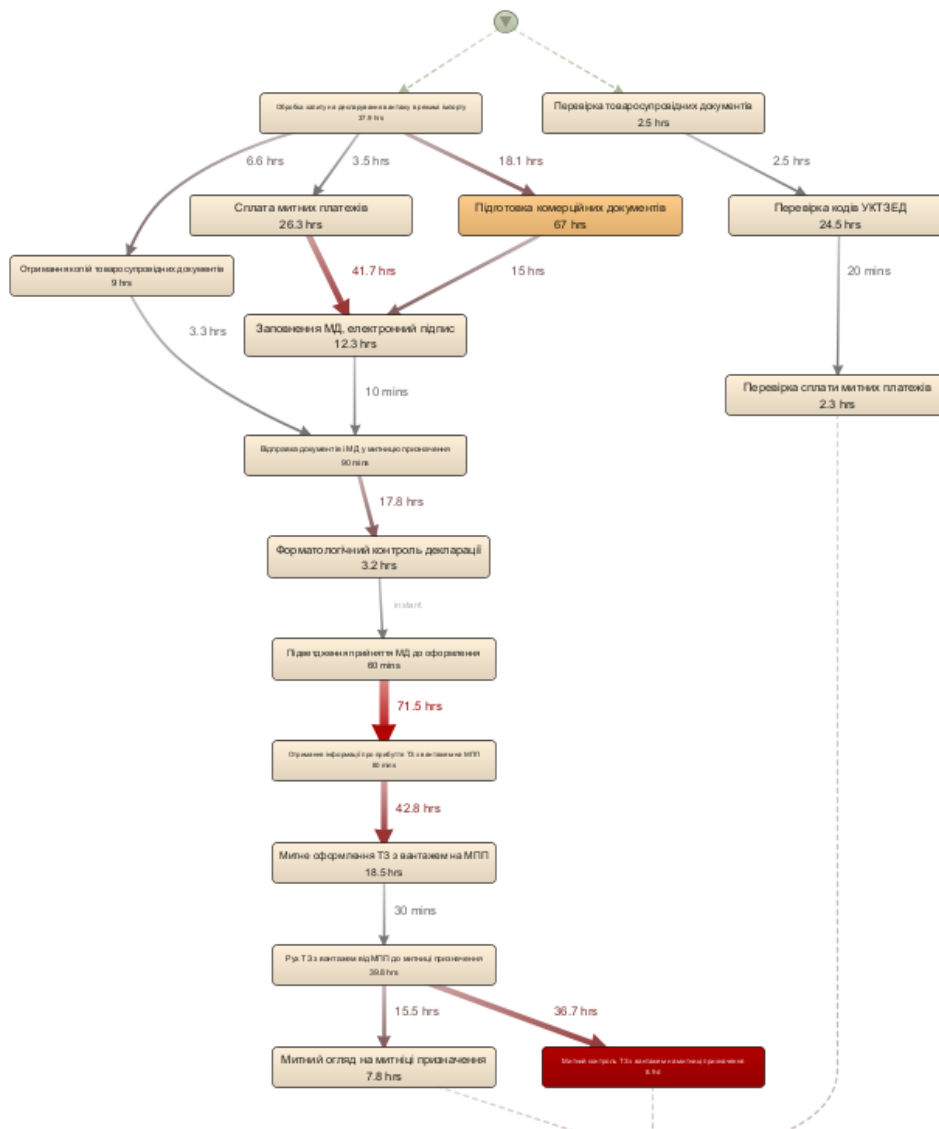


Рисунок 15 – Визначення ефективності процесу контролю за доставкою (тривалість виконання операцій)

Отже, за допомогою програмного забезпечення, яке імплементує моделювання процесів на основі теорії мереж Петрі, для заданого прикладу виділимо ті операції процесу, які тривають найдовше порівняно з іншими (рис. 16). Таким чином, у подальшому аналізуючи конкретні операції, менеджери можуть під час спілкування з виконавцями чи експертами визначити причини затримок у виконанні чи тривалість операції. Крім того, для визначеного процесу визначимо нормативні значення загальної тривалості (наприклад, 4 години на митне оформлення партії товару в митниці призначення) й окремо виділимо ті випадки, які перевищують ці нормативи.

При детальному аналізі визначених кейсів було встановлено, що значні затримки часу відбуваються під час декларування товарів, що потребують проходження фітосанітарного контролю (рис. 16), при чому затримка виникла не через безпосередню тривалість перевірки, а тривалу передачу документів інспектору митного посту.

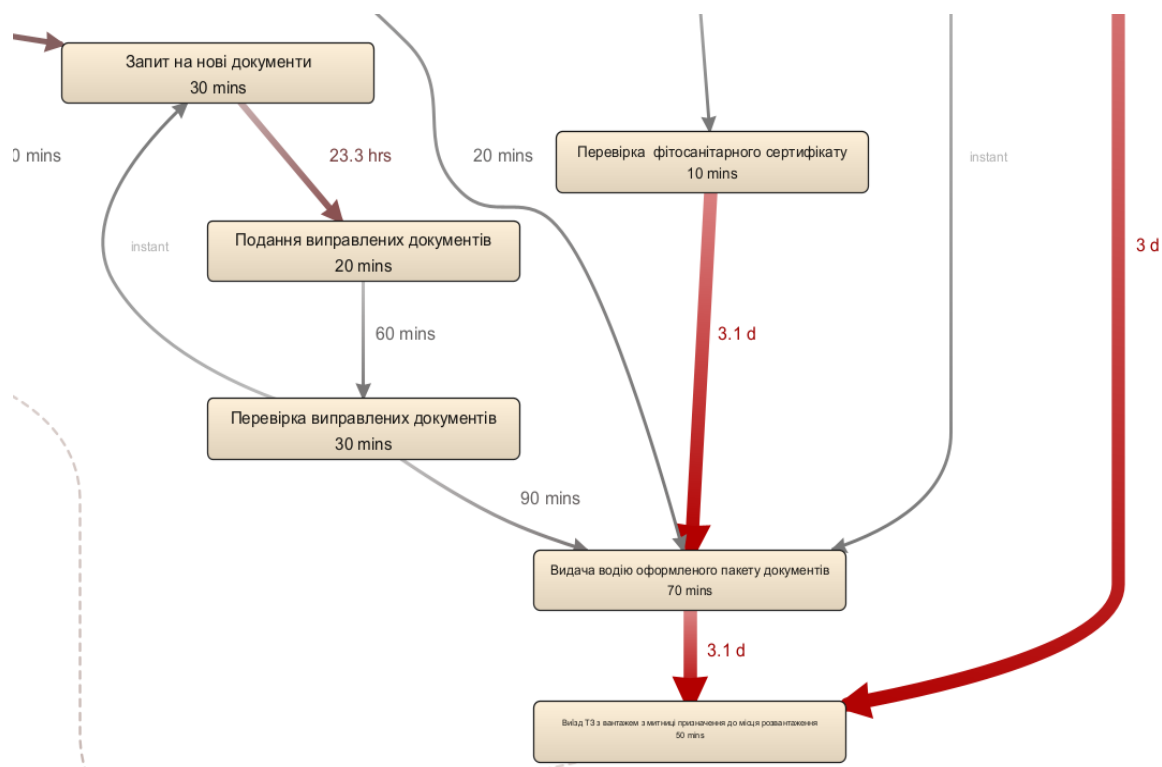


Рисунок 16 – Приклад реалізації функції програмного забезпечення з виділення однієї ділянки мережі з операціями, на які витрачається найбільше часу (витяг)

Окремі можливості програмного забезпечення щодо аналізу процесу продемонстровані лише у невеликій кількості, достатній для учбових цілей (рис.17-21).

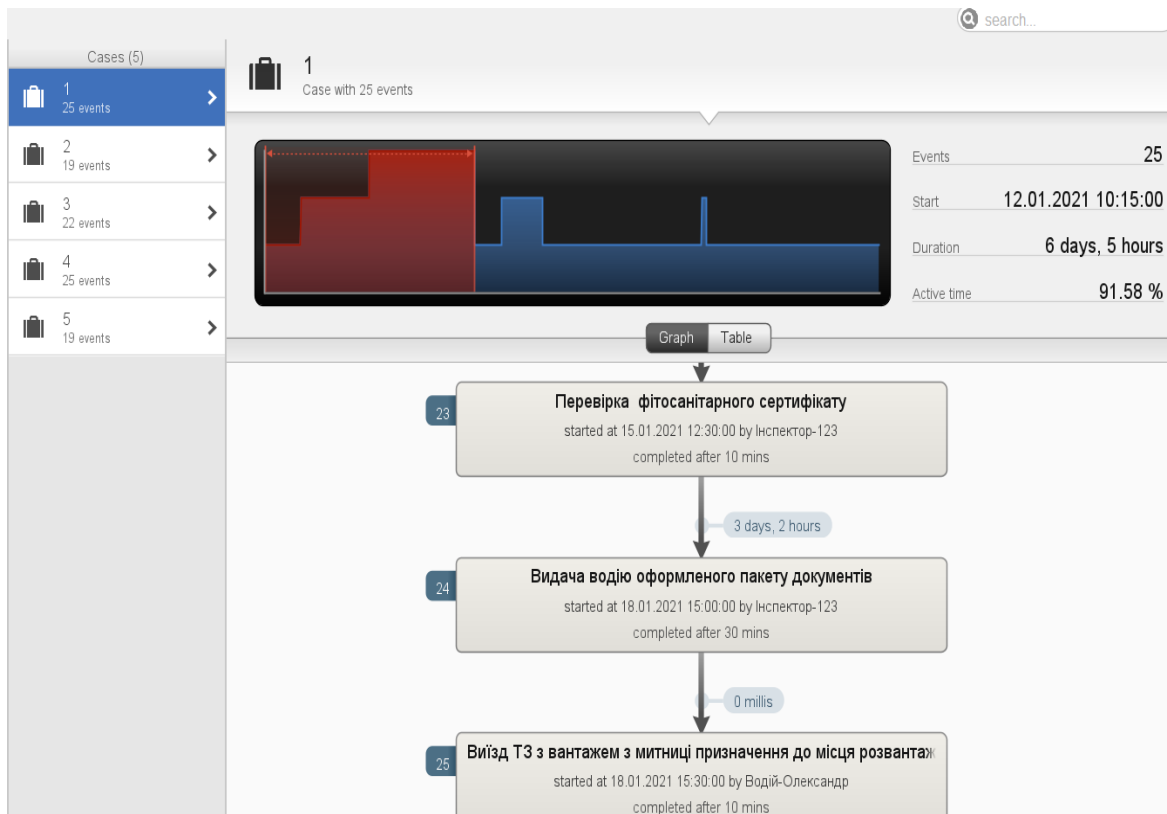


Рисунок 17 – Аналіз кейсу №1 із значною затримкою у виконанні операції

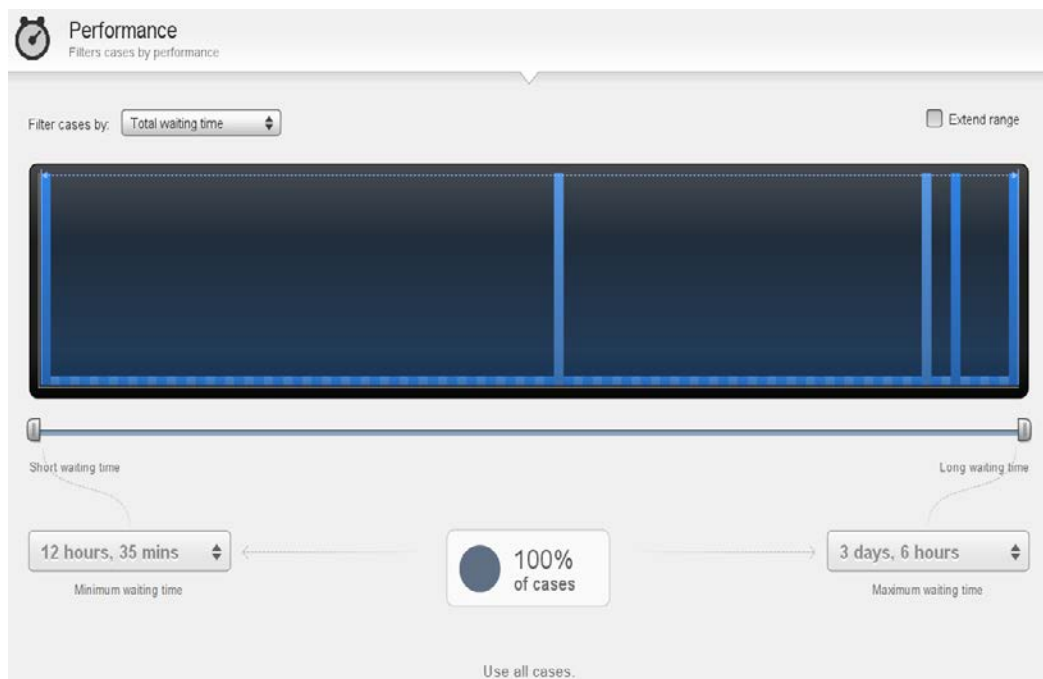


Рисунок 18 – Статистичний аналіз всіх кейсів: час очікування

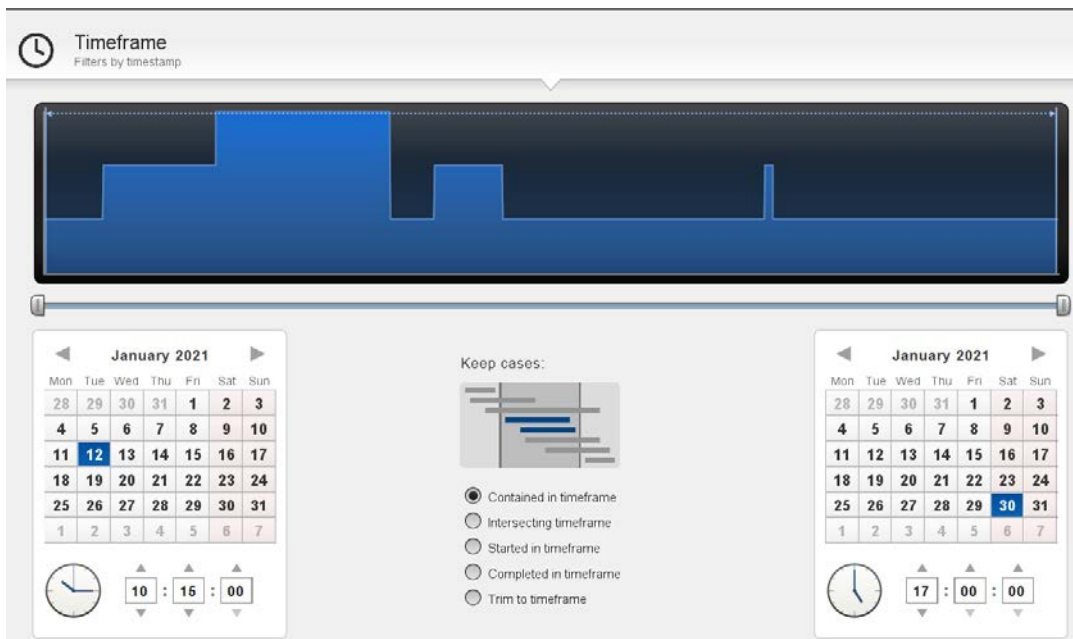


Рисунок 19 – Статистичний аналіз всіх кейсів: календар подій

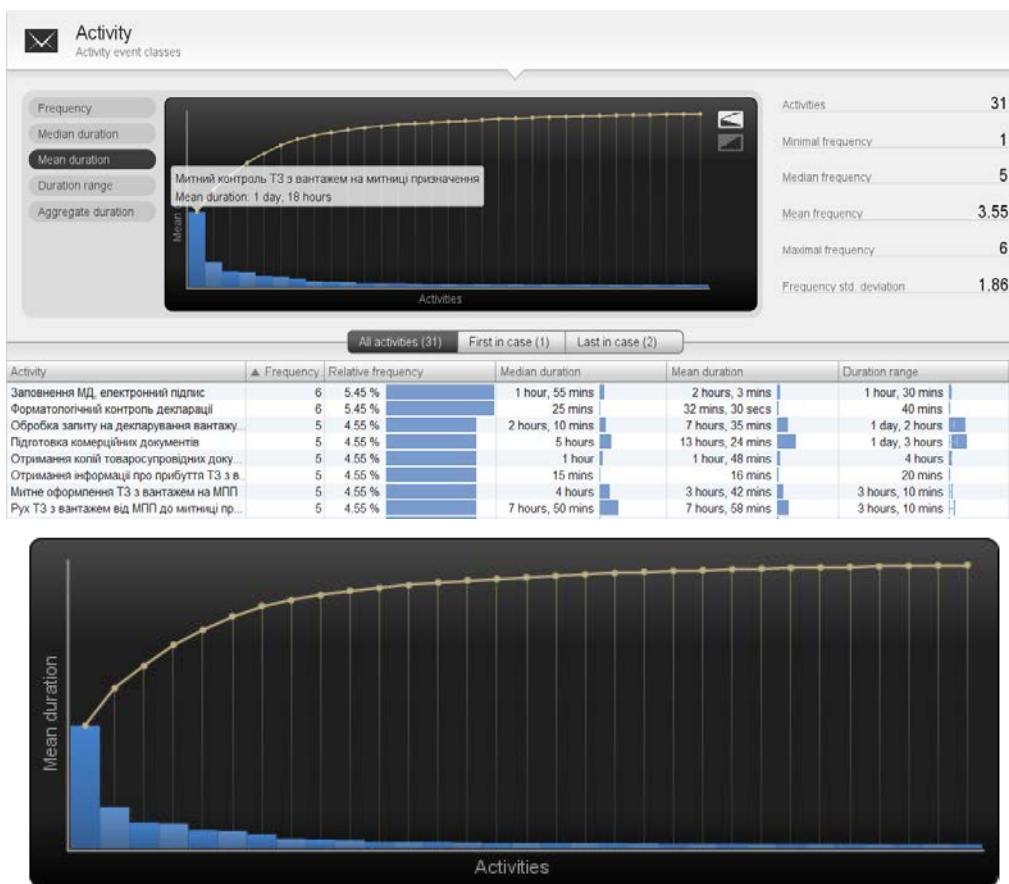


Рисунок 20 – Статистичний аналіз всіх кейсів: середня тривалість виконання кожної з операцій

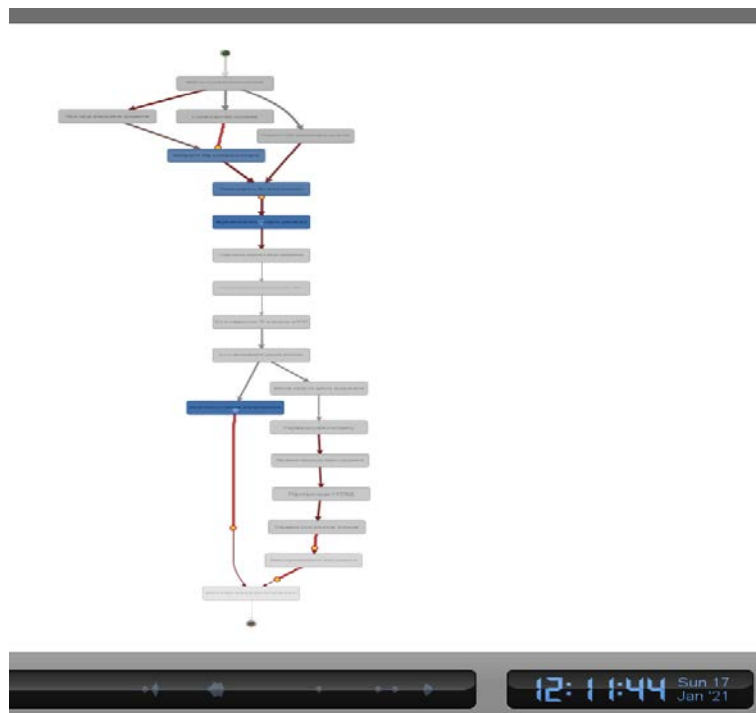


Рисунок 21 – Демонстрація динаміки процесу в режимі реального часу (витяг)

Отже, враховуючі поставлені задачі дослідження, за допомогою спеціалізованих інструментів інтелектуального аналізу процесу визначено:

- як саме виглядає процес: побудова карти подій і виділення певних «проблемних» етапів процесу за критерієм середньої тривалості їх виконання (наприклад, фітосанітарний контроль в структурі митного контролю в митниці призначення);

- чи є відхилення у процесі виконання контролю (співставлення законодавчо визначеної процедури й фактично виконаних етапів процесу митного контролю товарів): у наведеному прикладі невиконання обов'язкових операцій не виявлено;

- чи виконані визначені в законодавчих документах нормативи проведення митного контролю: у прикладі середній час завершення митної операції перевищив нормативи (більше доби замість 4 годин). У подальшому необхідно детальніше дослідити вже не час виконання операцій митними службами, а час очікування перевірки і тривалість подання необхідних для перевірки документів.

ВИСНОВКИ

Для підсумку, в матеріалах розділу спершу формується загальна уява про *інтелектуальний аналіз процесів* (eng. process mining) як засобу подолання розриву між наукою про дані та наукою про процеси, далі наводяться приклади нотацій, за допомогою яких можна побудувати моделі процесів і провести необхідний аналіз результатів моделювання з урахуванням певних ділових аспектів. Мережі Петрі є найстарішою та найкраще розвиненою для цього нотацією. Ця теорія має розвиток й практичне застосування для різних сфер діяльності людини, у тому числі для сфери вантажних автомобільних перевезень, – саме там, де складні процеси потребують постійного моніторингу, ретельного аналізу, прогнозування результативності їх виконання. Запропоновані експертними спільнотами інструменти імплементації теорії мереж Петрі, такі як ProM, надають можливість бізнес-аналітикам і менеджерам широкі можливості щодо підвищення ефективності бізнес процесів в організації за рахунок виявлення неефективних операцій, демонстрації витрат ресурсів під час реалізації процесу, існування відхилень від нормованих значень параметрів процесу, чіткого аналізу деталей процесу, контролю за виконанням поставлених цілей, прогнозування тощо. Для прикладу в матеріалах розділу був схематично наведений спрощений процес митного контролю за переміщенням товарів через митний кордон України, на основі якого була побудована мережа Петрі й виділені ті його операції, які тривали найдовше порівняно з іншими, а також встановлені факти затримок у виконанні окремих операцій. Крім того, після визначення нормованих значень загальної тривалості процесу, окремо виділено ті кейси, які перевищують ці нормативи. Надалі це можна використати для багатостороннього аналізу ситуацій і зниження ризиків ведення бізнесу.

ЛІТЕРАТУРА

- [1] Wil Van der Aalst. Data Scientist: The Engineer of the Future/Proceedings of the I-ESA Conference, volume 7 of Enterprise Interoperability (2014), 14 pages.
- [2] Wil Van Der Aalst M.P., Christian Stahl. Modeling Business Processes: A Petri Net-Oriented Approach (Information Systems)/ The MIT Press; Illustrated edition (2011) - 386 pages.
- [3] Dictionary of Algorithms and Data Structures/ The National Institute of Standards and Technology(NIST). Доступно з: [<https://xlinux.nist.gov/dads/>]
- [4] Lars Reinkemeyer et al. Process Mining in Action: Principles, Use Cases and Outlook. / Springer Nature Switzerland AG, 2020 – 150 pages.
- [5] Інформаційний ресурс ProcessMining. Доступно з: <http://www.processmining.org>[in English]
- [6] Інформаційний ресурс Colored Petri Nets. Доступно з: <http://cpntools.org/> [in English]
- [7] Інструмент моделювання Disco. Доступно з: <https://fluxicon.com> [in English]
- [8] Iuliia Silantieva, Nataliia Katrushenko, Bohdana Kushym. Ensuring effectiveness in handling the movement of goods and passengers by enhancing information and communication technologies/ Current Problems of Transport: Proceedings of the 1st International Scientific Conference (2019), Ternopil, Ukraine. Доступно з: <https://zenodo.org/records/3387287> [in English]