


НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
ІНСТИТУТ ПРОБЛЕМ МОДЕЛЮВАННЯ В ЕНЕРГЕТИЦІ
ІМ. Г.Є. ПУХОВА

КОВАЛЬ ОЛЕКСАНДР ВАСИЛЬОВИЧ



УДК 004.82 : 004.89

МЕТОДИ ТА ЗАСОБИ КОМП'ЮТЕРНОГО МОДЕЛЮВАННЯ СЦЕНАРІЇВ
АНАЛІТИЧНОЇ ДІЯЛЬНОСТІ

Спеціальність 01.05.02 – математичне моделювання та обчислювальні методи

АВТОРЕФЕРАТ

дисертації на здобуття наукового ступеня
доктора технічних наук

Київ – 2021

Дисертацією є рукопис

Робота виконана в Інституті проблем реєстрації інформації Національної академії наук України.

Науковий консультант:

доктор технічних наук, професор,
Додонов Олександр Георгійович
Інститут проблем реєстрації інформації НАН України,
заступник директора з наукової роботи.

Офіційні опоненти:

чл.-кор. НАПН України, доктор технічних наук, професор
Верлань Анатолій Федорович
Інститут проблем моделювання в енергетиці
ім. Г.Є. Пухова НАН України,
головний науковий співробітник;

доктор фізико-математичних наук, професор
Клименко Віталій Петрович
Інститут проблем математичних машин та систем
НАН України,
заступник директора з наукової роботи;

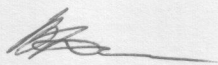
доктор технічних наук, старший науковий співробітник
Чугай Андрій Михайлович
Інститут проблем машинобудування
ім. А.М. Підгорного НАН України,
провідний науковий співробітник.

Захист відбудеться « 30 » вересня 2021 року о 12 годині на засіданні спеціалізованої вченої ради Д 26.185.01 Інституту проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України за адресою: 03164, Київ, вул. Генерала Наумова, 15.

З дисертацією можна ознайомитися у бібліотеці Інституту проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України за адресою: 03164, Київ, вул. Генерала Наумова, 15.

Автореферат розісланий «30» серпня 2021 р.

Вчений секретар
спеціалізованої вченої ради



В.В. Душеба

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Сьогодні ефективність прийняття управлінських рішень практично у всіх областях діяльності спеціалістів, у тому числі науковців та інженерів в процесі проектування складних технічних систем, на пряму залежить від наявності та якості інформації та знань, що аналізується. Але з кожним роком обсяг даних, інформації та знань, які потрібні для прийняття збалансованих рішень, збільшується в геометричній прогресії, причому процес аналізу цього обсягу даних все більш обмежується в часі, що відповідно змушує розробників програмних систем (ПС), які направлені на підтримку аналізу інформації, постійно вдосконалювати методи і технології отримання, структуризації та аналізу різноманітних даних.

З іншої сторони, згідно з дослідженнями всесвітньовідомої консалтингової компанії Gartner від 70% до 80% проектів, пов'язаних з впровадженням складних аналітичних рішень в організаціях та корпораціях, – провальні. Все це спонукає до подальшого пошуку шляхів удосконалення як наукових, так і технологічних підходів, які сприяють раціональній організації аналітичної діяльності (АнД) та підвищують якість інформації, що аналізується, та ефективність ПС в цілому. Починаючи з 2013 та ж компанія Gartner визначила дієву аналітику в десятці найбільш затребуваних інформаційних технологій. При цьому зазначається для того щоб аналітика була дієвою вона повинна мати змістовне наповнення, бути цілеспрямованою, зокрема надавати засоби моделювання та оптимізації сценаріїв аналізу інформації та розвитку ситуації або процесів.

Сучасний розвиток інформаційних технологій, в тому числі й тих, що вирішують задачі АнД, вирізняється рядом тенденцій, які насамперед обумовлені зрушенням від кастомізації, яка на тяжінні довгого часу була золотим стандартом, в сторону персоналізації. До таких тенденцій, які забезпечують цей зсув можна визначити накопичення та використання знань про предметну область (ПрО) та накопичений попит вирішення завдань в ПС, інтелектуалізація як самих ПС так і інтелектуалізації за їх використанням процесів прийняття рішень, структурна алгоритмізація та як наслідок використання сервіс-орієнтованої архітектури при побудові ПС. Ці тенденції вимагають наукового осмислення та опрацювання, розробки нових архітектурних вирішень, включаючи й підходи до проектування, впровадження ПС, спрямованих на підтримку АнД.

Аналіз публікацій у напрямку розробки та використання підходів до моделювання сценаріїв АнД свідчить про існування певних тенденцій, але розв'язання проблеми в цілому не можна вважати завершеним.

Суттєвий внесок в проблему створення методів і засобів комп'ютерного моделювання програмних засобів, зокрема зі застосуванням засобів побудови сценаріїв та моделювання предметної області, внесли результати праць М.З. Згурівського, В.Д. Самойлова, О.Г. Додонова, А.О. Морозова, О.В.Палагіна, А.В.Анісімова, А.Я.Гладуна, А.П. Писаренко, Д.В. Ланде, Н.Г. Загоруйко, Е.А. Рабчевського, В.Ш. Рубашкина, В.М. Одріна, С.С. Картавова, І.П. Норенкова, М.Ю.Уварова, Ю.В.РогушинаЮ Bart Kosko, Lotfi Aliasker Zadeh, T. Gruber, S. Staab, R. Studer, N.F. Noy, A. Gomez-Perez, T.Саати, Ф.Цвікі, Т. Річі, Дж. Х.Лінстона та

М.Турофа, Т. Гордона, О. Хелмера, Н. Долкі, Х. Сакмана, Т. Грубера, Дж. Солтона, А. Гомес-Переса, Н. Гуаріно, П. Фолтса, М. Шамбарда, Д. Челюска, М. Варгас-Вера, Т. Ватсона, П. Льюїса, К. Блашке та ін

Таким чином, є актуальним і необхідним вирішення *науково-прикладної проблеми дослідження, формалізації та розробки теоретичних основ та практичних засад комп'ютерного моделювання сценаріїв аналітичної діяльності в різних предметних областях, зокрема при проектуванні складних технічних систем.*

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконана в межах концепції програми інформатизації Національної академії наук України, визначеної пріоритетним напрямом згідно розпорядження № 146 від 27.02.2004 р., та за тематикою наукових досліджень Інституту проблем реєстрації інформації Національної академії наук України, зокрема в рамках науково-дослідних робіт: «Дослідження методів побудови аналітичної складової корпоративних інформаційно-аналітичних систем» (шифр «КІАС-2008»), № держреєстрації 0108U000260; «Дослідження та розробка технологій комп'ютерного моделювання інформаційно-аналітичних систем» (шифр «АСКІАС-2011»), № держреєстрації 0111U002089; «Дослідити та розробити технології комп'ютерного моделювання систем керування знаннями при проведенні аналітичної діяльності» (шифр «УЗОР-2014»), № держреєстрації 0114U002060; «Методи та засоби комп'ютерного моделювання сценаріїв аналітичної діяльності в системах організаційного управління» (шифр теми «СЦЕНАРІЙ») № держреєстрації 0119U001165.

Окрім того, у роботі використано результати, які отримав автор під час виконання науково-дослідних робіт у Національному технічному університеті України «Київський політехнічний інститут імені Ігоря Сікорського», а саме: «Розробка алгоритмів (документації) та програмного забезпечення виявлення, визначення напрямку та класифікації морських об'єктів для існуючої у Інозамовника вимірювальної системи з дослідження сигналів та перешкод в світовому океані», № держреєстрації 0116U005555; «Дослідження та розробка математичних моделей сигналів та перешкод, розрахунок параметрів розповсюдження звукових хвиль у морському багатопаровому середовищі», № держреєстрації 0116U005598; «Моніторинг фінансових показників виконання бюджетів всіх рівнів», № держреєстрації 0111U004274.

Мета і задачі дослідження. Метою дисертаційної роботи є створення методів і засобів побудови комп'ютерних систем моделювання сценаріїв аналітичної діяльності.

Для досягнення поставленої мети в дисертаційній роботі вирішено такі задачі.

1. Аналіз сучасного стану і особливостей проблеми дослідження, створення та вдосконалення комп'ютерних систем моделювання багаторівневих сценаріїв АнД, вибір та обґрунтування перспективних підходів до її вирішення з метою визначення та зменшення логічно-зв'язаних помилок та семантичної незв'язаності опису багаторівневих сценаріїв чи їх окремих гілок, що призводить до суттєвого зменшення часу моделювання сценаріїв в цілому.

2. Розробка теоретичних засад встановлення співвідношення графічних елементів опису сценарію АнД з описом онтології його предметної області з метою

забезпечення можливості проводити семантичний аналіз коректності як логічно-зв'язаних фрагментів сценарію АнД, так і його верифікацію.

3. Дослідження можливостей та розвиток математичних моделей і методів побудови сценарію АнД з урахуванням можливих наступних кроків сценарію та на основі виведення по прецедентах, що вже накопичені в результаті попереднього моделювання сценаріїв АнД.

4. Розробка та дослідження метода виконання сценаріїв АнД на основі Web-сервісів, вибір яких засновано на механізмі порівняння параметрів метаописів Web-сервісів із параметрами функціональних завдань сценарію.

5. Розробка алгоритмів та програмних засобів комп'ютерного моделювання сценаріїв на основі розроблених та удосконалених моделей (сценарних моделей, моделей опису предметної області в умовах невизначеності, заданих обмежень) та методів побудови та виконання сценаріїв АнД.

6. Створення комп'ютерної моделюючої системи (інтегрованого програмного інструментального середовища) для тестування і оцінки запропонованого теоретичного підходу комп'ютерного моделювання сценаріїв АнД, яка ґрунтується на побудованих моделях, методах та алгоритмах моделювання.

Об'єктом дослідження є процеси функціонування комп'ютерних систем моделювання сценаріїв АнД.

Предметом дослідження є методи і засоби математичного та комп'ютерного моделювання процесів функціонування комп'ютерних систем моделювання сценаріїв АнД.

Методи дослідження. Для розв'язування поставлених задач у дисертації використовуються теоретико-множинні підходи; алгебра систем; теорія графів; методи семантичного моделювання, методи математичного моделювання переходів станів; апарат модельно-орієнтованих підходів до побудови комп'ютерних ПС, методи концептуального моделювання, методи організації засобів комп'ютерного моделювання динамічних процесів і систем, загальні архітектурні принципи побудови ПС, методи побудови сервісно-орієнтованих ПС, методи оцінювання тривалості виконання завдань проекту.

Наукова новизна одержаних результатів. В рамках розв'язання проблеми створення методів і засобів математичного та комп'ютерного моделювання багаторівневих сценаріїв АнД на основі поєднання переваг візуального моделювання і семантичного аналізу та подальшої подальшого автоматичного перетворення метаданих моделі сценарію в виконавче програмне інструментальне середовище отримані наступні наукові результати.

Вперше запропоновано:

– теоретичний підхід до комп'ютерного моделювання багаторівневих сценаріїв АнД, якій відрізняється від відомих тим, що оснований на поєднанні переваг візуального моделювання – графічної нотації BPMN (Business Process Model and Notation) з керованою моделлю і семантичного аналізу на базі онтологічної моделі предметної області задач АнД, яка конвертується з BPMN-моделі, та подальшого автоматичного перетворення метаданих BPMN-моделі в виконавче програмне середовище (серіалізації), що зменшує кількість логічних

розривів в комп'ютерній моделі побудови сценарію АнД, яка будується, та суттєво зменшує час моделювання в цілому;

- метод перетворення BPMN-моделі в модель уявлення про предметну область сценарію АнД мовою OWL (Web Ontology Language) - OWL-модель, якій відрізняється від відомих тим, що за рахунок забезпечення співвідношення категорії основних графічних примітивів нотації BPMN до відповідних класів онтології (сутностей) та зв'язування їх з властивостями та відношеннями між елементами реальної моделі вирішення задачі АнД в нотації BPMN дозволяє проводити семантичний аналіз коректності логічно-зв'язаних фрагментів сценарію вирішення задачі АнД та його верифікацію (особливо багаторівневих сценаріїв);

- інформаційна технологія конвертації BPMN-моделі в OWL-модель сценарію АнД, яка відрізняється від відомих тим, що дозволяє перетворювати метадані опису сценарію та його предметної області в базі знань, накопичуючи нові знання, що дає можливість побудови нових сценаріїв АнД на базі вже існуючих в базі знань. Для спрощення процесу комп'ютерного моделювання сценаріїв залучається графічна нотація BPMN 2, яка дозволяє не тільки моделювати сценарії АнД у вигляді BPMN-діаграм, але й серіалізувати його опис в XML-файл для реалізації сценарію в комп'ютерному середовищі із залученням стандартного програмного забезпечення BPM-систем.

Удосконалено:

- математична модель обчислення розподілу значень можливих подій сценарію та інтелектуальний метод визначення наступних подій сценарію АнД, який відрізняється від відомих тим, що на основі застосування методів машинного навчання, а саме розвитку методу навчання деревами класифікації і регресії (Classification and Regression Trees) з використанням комбінації метрик оцінок ефективності та накопичених знань про семантику предметної області для визначення кореневого вузлу графу, дозволяє виконувати формування дерева рішень для класифікації можливих наступних кроків сценарію. Це дає можливість значно підвищити точність розбиття графу складного сценарію на відповідні гілки при обчисленні ймовірності наступного кроку сценарію;

- метод реалізації сценаріїв АнД на основі Web-сервісів, який відрізняється від відомих тим, що вибір сервісів засновано на механізмі порівняння параметрів їх метаописів із параметрами функціональних завдань сценарію, які задаються його метаописом, що забезпечує можливість виконання завдань кінцевого користувача шляхом динамічного формування послідовності Web-сервісів.

Набули подальшого розвитку:

- процес автоматизованої побудови сценаріїв АнД на основі бібліотеки сценаріїв (прецедентів) для різних предметних областей. Це дало можливість в автоматизованому режимі будувати сценарії на основі експертного опису проблемних ситуацій, здійснювати деталізацію сценаріїв та розгалуження їх. У свою чергу, це суттєво зменшує часові та фінансові витрати на моделювання та реалізацію сценаріїв АнД.

- інтелектуальний програмний засіб, що навчається на діях користувача та на основі отриманих в процесі навчання знань представляє дії користувача у вигляді дерева рішень ймовірного сценарію аналітичного процесу, пропонуючи користувачу найбільш доцільні наступні кроки вирішення завдання АнД;

– тривірнева комп'ютерна моделююча система (інтегроване програмне інструментальне середовище) для тестування і оцінки запропонованого теоретичного підходу (методології) моделювання сценаріїв АнД, яка ґрунтується на запропонованих моделях, методах та алгоритмах моделювання, що дало можливість реалізувати окремі програмні інтелектуальні засоби вирішення задач аналітики для різних предметних областей аналітичної діяльності, ядром баз знань яких є знання та накопичений досвід аналітика (експерта, проектувальника). Програмні інтелектуальні засоби були використані зокрема для вирішення задачі побудови та реалізації сценаріїв моделювання процесів функціонування вимірювальної гідроакустичної системи з метою удосконалення алгоритмів обробки гідроакустичних сигналів, та сценаріїв розподілу енергопостачання з метою оптимізації енерговитрат.

Практичне значення одержаних результатів. Практичне значення отриманих результатів визначається тим, що запропоновані та створені засоби алгоритмічного та програмного забезпечення процесів моделювання сценаріїв АнД забезпечують комп'ютерну реалізацію та взаємодію отриманих видів моделей графічного представлення сценаріїв АнД, семантичних моделей предметної області задач АнД та знань про процес побудови та виконання сценаріїв АнД у вигляді адаптивної об'єктної моделі. За використанням розроблених інтелектуальних програмних засобів здійснюється автоматизована побудова сценарію АнД з урахуванням попередніх його кроків та на основі виведення по прецедентах, що вже накопичені в результаті попереднього моделювання сценаріїв АнД. Такій підхід забезпечує зменшення логічно-зв'язаних помилок та семантичної незв'язаності опису багаторівневих сценаріїв чи їх окремих гілок та тем самим призводить до суттєвого зменшення часу моделювання сценаріїв АнД. Це підтверджують результати проведених імітаційних експериментів на тестових прикладах, які показали, що в залежності від кількості рівнів (кількістю рівнів 4 та 10) графу багаторівневого сценарію та кількості вузлів в рівні (кількість вузлів на рівні 3,5,7,10 та 15) час моделювання мінімально зменшився на 10 % та максимально зменшився на 25 %.

Розроблені комп'ютерні програмні засоби призначені для дослідження та забезпечення процесів моделювання сценаріїв АнД як у лабораторних дослідженнях, так і в умовах застосування в реальних системах із забезпеченням їх інтелектуалізації та адаптивності до змін предметної області завдань, які вирішує аналітична ПС.

Створені методи та засоби математичного та комп'ютерного моделювання сценаріїв АнД використано при розв'язуванні ряду практичних задач, зокрема, для вирішення задачі побудови та реалізації сценаріїв моделювання процесів функціонування вимірювальної гідроакустичної системи з метою удосконалення алгоритмів обробки гідроакустичних сигналів, та сценаріїв розподілу енергопостачання з метою оптимізації енерговитрат.

В роботі вирішено наступні прикладні задачі, що представлені у додатках: «Побудова та реалізація сценаріїв моделювання процесів функціонування вимірювальної гідроакустичної системи» (додаток А); «Комп'ютерне моделювання сценаріїв розподілу енергопостачання» (додаток Б); «Формування сценарію інтерактивної обробки багатовимірних даних» (додаток В); «Побудова та реалізація

сценаріїв аналітичної діяльності в задачах просторового моделювання» (додаток Г); «Моніторинг фінансових показників виконання бюджетів всіх рівнів» (додаток Д). Основні результати роботи та основні положення, які виносяться на захист, пройшли апробацію при розв'язанні модельних і практичних задач.

Результати дисертаційного дослідження прийнято до впровадження в Дочірньому підприємстві «Національний фонд розвитку», ТОВ «Науково-виробниче підприємство «Символ», ТОВ «Інженерна логіка», Національному технічному університеті України «Київський політехнічний інститут імені Ігоря Сікорського».

Особистий внесок здобувача. Усі наукові результати дисертаційної роботи отримані автором особисто. У друкованих працях, опублікованих у співавторстві, автору належать: [1] – формулювання та доведення методології семантичного моделювання побудови систем управління знаннями та алгоритмічних і технологічних засад моделювання аналітичної діяльності; [2] – постановка задачі, класифікація моделей, архітектура та принципи функціонування програмних засобів побудови та реалізації сценаріїв аналітичної діяльності в інформаційно-аналітичних системах; [3] – ідея інтегрованого підходу до визначення якісної оцінки сценарію аналітичної діяльності, якій заснований на виявленні найкоротшого шляху у графі онтології; [4] – постановка задачі та застосування підходу до моделювання змінюваних сценаріїв аналітичної діяльності є залученням технологій BPM-систем з керованої моделлю; [5] – адаптивна мікросервісна архітектура інформаційно-аналітичної системи; [6] – постановка задачі та принципи формування багаторівневих сценаріїв GIS-аналітики за застосуванням BPMN-моделі в поєднанні с онтологічною моделлю ПрО; [7] – постановка та способи вирішення задачі побудови типових сценаріїв на основі аналізу онтології предметної області з використанням представлення її структури у вигляді графу; [8] – інтелектуальний метод побудови сценарію аналітичної діяльності, заснованого на розвитку методу навчання деревами класифікації і регресії з використанням комбінації метрик оцінок ефективності; [9] – застосування підходу до побудови сценаріїв проектування бізнес-процесів виконання інженерних завдань на порталах знань; [10] – застосування підходу використання оцінок ймовірностей для аналізу відповідності джерел інформації запиту пошуку відповідних документів у цих джерелах; [11] – принципи застосування концептуальної моделі сценарно-цільового підходу до аналізу об'єктів дієвої аналітики; [12] – постановка задачі та застосування підходу використання типових сценаріїв аналізу виконання проектів на основі типових задач за створенням репозитарію сценаріїв і репозитарію типових задач; [13] – принципи побудови комп'ютерної моделі системи інформаційного управління з дієвою аналітикою; [14] – принципи застосування адаптивної об'єктної сценарної моделі для побудови сценарію аналітичної діяльності; [15] – постановка задачі та застосування підходу до побудови системи моніторингу виконання держбюджету України, який ґрунтується на інформаційній моделі інформаційно-аналітичної системи, моделі прецедентів предметної області та типових сценаріях моніторингу й аналізу державного бюджету; [16] – постановка задачі та застосування підходу до підтримки управління проектами, заснований на побудові та застосуванні сценаріїв аналізу змін бізнес-процесів та бази знань проекту для визначення оперативних завдань управління проектом; [17] – постановка задачі та застосування підходу до аналізу типових задач проектів, заснований на накопиченні знань про виконання

всіх проектів з часовими параметрами, та схема послідовності обробки даних оцінки своєчасності виконання проектів; [18] – постановка задачі та принципи створення спеціалізованих програмних засобів аналізу проектної діяльності з урахуванням стандартних корпоративних політик безпеки; [20] – принципи застосування розподіленої сервіс-орієнтованої архітектури автоматизованої системи екстреної допомоги населенню за єдиним телефонним номером 112; [21] – принципи побудови комп'ютерної адаптивної моделі корпоративних інформаційно-аналітичних систем, яка базується на концептуальних засадах теорії складних адаптивних систем; [22] – принципи побудови та застосування мікросервісної архітектури інформаційно-аналітичної системи для оцінки рівня міжнародної діяльності, в якій кожен мікросервіс створюється як автономний контейнер; [23] – застосування методу спрощення інженерних завдань за допомогою специфічної алгебраїчної системи інженерних завдань, спрямованих на формування мінімізованого набору веб-сервісів порталу знань; [24] – формування етапів та параметрів розвитку онтологічної моделі предметної області при проектуванні програмної платформи аналізу інформації; [25] – методологія створення онтологічної моделі предметної області інтелектуальної системи дослідження гідроакустичних процесів та принципи побудови її узагальненої архітектури на основі знань, які накопичуються в процесі дослідження; [26] – метод формування сценарію завдань користувача, що дозволяє динамічне визначення впорядкованої послідовності веб-сервісів, необхідних для виконувати завдання користувача.

Апробація результатів дисертації. Основні результати дисертаційних досліджень неодноразово доповідалися на міжнародних наукових конференціях, зокрема на конференціях «3-я Міжнародна конференція з передових інформаційних та комунікаційних технологій» (м. Львів, 2019), «VII, XVIII та XIX Міжнародна науково-практична конференція "Інформаційні технології та безпека"» (м. Київ, 2017-2019), «IEEE Міжнародна Чорноморська конференція з питань комунікацій та мереж» (м. Батумі, Грузія, 2018), «Міжнародна конференція «Радіоелектроніка та інформаційні комунікації»» (м. Київ, 2016, м. Одеса, 2017), «14-18 Міжнародна науково-технічна конференція «Системний аналіз та інформаційні технології»» (м. Київ, 2012–2016), «VII та IX Міжнародна науково-практична конференція аспірантів, магістрів, студентів «Сучасні проблеми наукового забезпечення енергетики»» (м. Київ, 2009, 2011), Науково-практична конференція с міжнародною участю «Системи підтримки прийняття рішень. Теорія і практика» (м. Київ, 2007), «Науково-практична конференція с міжнародною участю «Проблеми моніторингу та забезпечення економічної безпеки держави»» (м. Київ, 2007).

Публікації. За результатами досліджень опубліковано 44 наукових роботи, з них: 2 монографії, 24 статті у наукових фахових виданнях України, що входять до переліку, затвердженого МОН України, 22 статі у виданнях, які включені до міжнародних наукометричних баз, 18 публікацій у працях і матеріалах наукових конференцій, 5 публікацій в закордонних виданнях, 8 публікацій у збірниках наукових праць, які включено до міжнародної наукометричної бази Scopus.

Структура та обсяг роботи. Дисертаційна робота складається з анотації, вступу, шести розділів, висновків, списку використаних джерел з 290 найменувань на 18 сторінках та 4 додатків на 23 сторінках. Загальний обсяг дисертацій 334 сторінки, з них 283 сторінки основного тексту, 75 рисунків, 17 таблиць.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі обґрунтовано актуальність теми, сформульовано мету та основні завдання дослідження, показано зв'язок із науковими програмами, планами, темами, відзначені наукова новизна й практична цінність отриманих результатів, особистий внесок здобувача в роботах у співавторстві, відомості про апробацію результатів дисертаційної роботи.

У першому розділі розглянуто особливості проведення АНД для різних предметних областей як об'єкту моделювання сценаріїв з урахуванням різноманіття факторів, які характеризують діяльність аналітика. Під АНД розуміється сукупність дій користувача (аналітика) ПС для збору, накопичення, обробки та аналізу даних з метою обґрунтування та підготовки прийняття рішень або генерації нових знань. Були визначені типові задачі АНД, в контексті яких під сценарієм АНД як предметом комп'ютерного моделювання розуміється сукупність методів опису та організації АНД (для багатофакторних ПрО при жорстко заданих обмеженнях), застосування яких сприяє генерації різних сценаріїв досягнення цілей аналітичного дослідження з урахуванням багатофакторності та обмежень. Тобто сценарій АНД розглядається як послідовність функціональних задач, начальних, кінцевих та проміжних подій, спрямованих на досягнення цілей аналітичного дослідження. Опис сценарію складається з розгалуженої направленої послідовності елементів, об'єднаних між собою через зв'язки, які утворюють собою структуру у вигляді направленої графа. Кожен елемент є або деякою дією, що реалізується виконавцем сценарію, або деякою подією, що впливає на подальше виконання сценарію. Подія може бути спричинена як змінами зовнішнього середовища, так і зміною стану самого об'єкта дослідження. Кожна дія може бути описана як функція, що повертає деякий результат, або як процедура, що може реалізувати будь-яку послідовність операцій (з заданим рівнем їх деталізації). Графічне відображення (рис.1) сценарію у вигляді можливих варіантів дій, які в залежності настання внутрішніх та зовнішніх факторів дають різні варіанти для кожного кроку (n).

Наведені результати аналізу теоретичних та методологічних засад комп'ютерного моделювання сценаріїв АНД для різних сфер діяльності, включаючи й створення інтелектуальних інформаційних систем, покликаних обробляти значні обсяги інформації (системи підтримки прийняття рішень, портали знань і т. п.).

Проведено аналіз існуючих інструментальних рішень та платформ побудови сценаріїв АНД (RapidMiner, KNIME, Orange, Weka, Alteryx) у контексті їх адаптаційних можливостей до зміни вимог показав, що класичні архітектури, зокрема клієнт-серверні та компонентні (CORBA, DCOM) спрощують адаптацію системи шляхом інкапсуляції функцій та спеціалізації компонент системи, повторного використання коду. Аналіз концепцій реалізації комп'ютерних систем як IBM Cloud service models, Microsoft Azure, Google Cloud, Oracle Cloud, Amazon Web Services (AWS) довів, що спостерігається тенденція надання інтегрованих засобів у виді «хмарних» Cloud-сервісів, в основу яких покладено принципи та технології інтелектуальних систем, повторне використання та накопичення знань про систему та середовище функціонування.

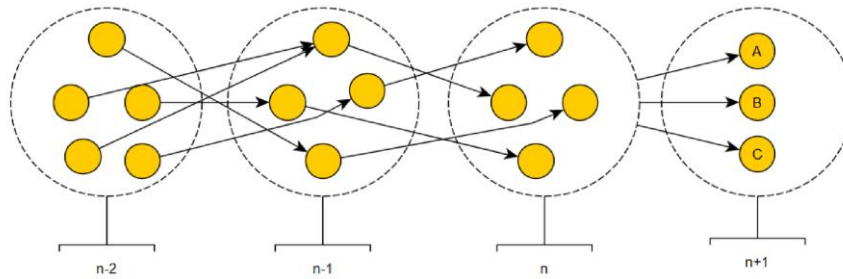


Рисунок 1 – Гіпотетична модель розгалуженого сценарію

Отже, в результаті проведеного аналізу було визначено особливості зазначених ПС, які заважають здійснювати за їх використанням комп'ютерне моделювання сценаріїв АНД особливо при зміні вимог та відповідно опису сценарію, а саме показано:

- в якості засобу опису сценарію процесів обробки даних відомі ПС використовують графічні нотації, що відповідає вимогам критерію функціональної оптимізації та дозволяє формувати підмножину вузлів обробки даних, необхідних для побудови сценаріїв;

- в зв'язку з цим побудова сценарію базується лише на візуальному представленні моделі процесів обробки даних, як наслідок, фіксованому наборі вимог, які, особливо на початковому етапі, визначаються нечітко, без формалізації та урахування семантичного наповнення процесів, що призводить до виникнення на етапі виконання сценаріїв логічно-зв'язаних помилок та необхідності повернення до початку процесу моделювання, що ускладнює процес моделювання та збільшує його час;

- правила логічно-зв'язаних переходів виконання сценаріїв АНД жорстко зашиті у програмному коді аналітичних програмних додатків, що не дозволяє використовувати ці сценарії при зміні цілей, початкових умов та обмежень АНД, тобто не враховується семантична складова АНД;

- ПС більш зосереджені на опису безпосередньо процесів обробки даних та не враховують їх взаємодію в межах багаторівневого сценарію, який виконується аналітиками з різними повноваженнями та компетенціями;

- ПС не використовують сучасні інтелектуальні методи, насамперед машинного навчання, в процесі виконання сценаріїв моделювання процесів обробки даних, тем самим не забезпечують можливості навчання ПС на власному досвіді та адаптування до незапрограмованих змін опису сценаріїв;

- в результаті зміна опису сценарію не може бути оперативно відображена у програмному додатку та, як наслідок, не дає можливості користувачеві, спираючись на свої знання чи на підказки ПС, які формуються з урахування вже накопиченого досвіду моделювання, скорегувати при необхідності сценарій в процесі здійснення аналітичної діяльності, що значно скоротило б час її здійснення.

До того ж, будь-яка зміна умов моделювання сценарію АНД вимагає реінжинірингу, перекодування та тестування програмних додатків моделювання, що займає багато часу та потребує додаткового фінансування.

Таким чином, з урахуванням цих особливостей з метою дослідження та розробки методів та засобів, які б забезпечували можливість адаптуватися до змін вимог до АНД, запропонований у відповідності до модельно-орієнтованого підходу до проектування ПС багаторівневий процес моделювання ПЗ моделювання сценаріїв АНД (рис. 2), якій включає три рівні: концептуальний, об'єктний та виконавчий, або рівень бізнес-процесів, на якому відбувається моделювання процесів та визначається бізнес-логіка ПС, рівень аплікацій та сервісів, що безпосередньо реалізують головні функції ПС, та рівень комп'ютерної моделі, який надає базові інфраструктурні послуги зі збереження, передавання та аналізу даних. До того ж, перспективним для вирішення завдань комп'ютерного моделювання сценаріїв АНД є застосування апарату семантичного моделювання і онтології як цілісної формальної моделі ПрО, яка віддзеркалює семантику сценаріїв АНД та розвиток концепцій, підходів, методів та засобів побудови сценаріїв АНД, що базуються на принципах використання активних моделей середовища функціонування та технологій інтелектуальних систем опрацювання знань.



Рисунок 2 – Багаторівнева модель програмних систем

У другому розділі обґрунтовано теоретичні засади комп'ютерного моделювання сценаріїв АНД на основі онтологічних моделей уявлення про ПрО сценарію АНД мовою OWL (Web Ontology Language) – OWL-моделі та процесного моделювання сценаріїв BPM-систем (Business Process Management). Під комп'ютерним моделюванням сценаріїв АНД в цій роботі розуміється процес відтворення дій користувача аналітичної ПС за допомогою побудови та виконання багаторівневої комп'ютерної моделі сценарію АНД з метою визначення складу та структури сценарію, його налаштування, перевірки і верифікації та забезпечення якості сценарію при використанні його для здійснення АНД. Відмітна особливість

процесно-орієнтованих BPM-систем з керованої моделлю полягає в тому, що на відміну від об'єктно-орієнтованого підходу, побудова сценарію Анд ведеться в термінах ПрО, а не комп'ютерного середовища та технологій для їх виконання.

Тобто, розробник стає захищеним від складнощів безпосереднього програмування виконуваного коду, а зосереджується на вдосконаленні моделі сценарію Анд (з позиції людського фактору), який є графічним описом послідовності робіт або станів об'єкту дослідження. Суть залучення BPM-рішень для моделювання сценаріїв Анд (особливо багаторівневих) полягає в тому, що по аналогії з бізнес-процесами, сценарії Анд можуть бути представлені сукупністю процесних моделей (типових процесів), кожна з яких має початок і кінець та які взаємодіють між собою в межах загального сценарію, використовуючи механізми формування переходів між вершинами спрямованого графа.

Для спрощення самого процесу побудови сценаріїв залучаються графічні методології та відповідні засоби моделювання, зокрема, нотація BPMN (Business Process Model and Notation). Таке рішення дозволяє не тільки моделювати сценарії Анд у вигляді графічних BPMN-діаграм, але й в подальшому автоматично перетворити метадані BPMN-моделі в виконавче програмне середовище (здійснити серіалізацію в формат XML-файлу) для відпрацювання в комп'ютерному середовищі із залученням стандартного програмного забезпечення BPM-систем, що зменшує кількість логічних розривів в комп'ютерній моделі побудови сценарію Анд, яка будується, та суттєво зменшує час моделювання сценарію.

Але будь-які формальні методи, включаючи й методи процесного моделювання, незважаючи на їх безумовні переваги у сенсі формалізації, мають суттєвий недолік – вони не пропонують механізмів змістовного (семантичного) опису та верифікації розгалужених сценаріїв. Для подолання цього недоліку з метою формування коректної моделі складного процесу моделювання сценаріїв Анд запропоновано доповнювати концептуальну модель сценарію Анд, представлену в графічній нотації, його концептуальною онтологічною моделлю, яка включає категорії ПрО, які співвідносяться з основними категоріями нотації BPMN – це підмножини класів, відносин, аксіом, правила логічного висновку, множина екземплярів.

Для реалізації зазначеного пропонується розроблений метод перетворення BPMN-моделі в OWL-модель сценарію, яка додає семантичне пояснення та інтерпретацію процесів. Суть метода в тому, що на першому етапі BPMN-модель має бути конвертована в OWL-файл (у форматі XML), якій включає такий набір сутностей онтології, що співвідноситься з усіма графічними елементами діаграми сценарію в нотації BPMN. Завдяки технології конвертації BPMN-моделі в онтологію, OWL-модель сценарію Анд, фактично, перетворюється в базу знань, яка може розширюватися за рахунок додавання нових сценаріїв Анд.

Нотація BPMN 2.0 використовує специфікацію, що дозволяє визначати складні конструкції моделі процесу у графічному виді. Крім того, BPMN 2.0 визначає, яким чином діаграми, можуть бути трансформовані в виконувани моделі на мові BPEL (Business Process Execution Language). Для здійснення процесного моделювання сценаріїв Анд вживаються певні категорії понять нотації BPMN, які

дозволяють доволі повно визначити сценарії Анд через її графічні елементи (примітиви нотації) у вигляді діаграми.

У математичному сенсі концептуальна модель сценарію – це спрямований граф, який складається з елементів нотації BPMN та описується кортежем

$$Gr(Bp_N) = \{V(Bp_N), E(Gr), L(v), Id(v), Rul^{(Bp)}, p(v)\},$$

де $\{V(Bp_N)\}$ – множина вузлів спрямованого графа $Gr(Bp_N)$, причому множина вузлів графа має вигляд

$$V(Bp_N) = \{St \cup Tk \cup In \cup Gt \cup Fn\},$$

де $\{St\}$ – підмножина вершин, що представляють собою стартові події (Start Event) кожного процесу, що описується графом; $\{Tk\}$ – підмножина вершин, що представляють окремі задачі Task або Sub-Process; $\{In\}$ – підмножина вершин, що представляють собою проміжні події процесу (Intermediate events); $\{Gt\}$ – підмножина вершин, що представляють собою шлюзи (Gateway) – логічні оператори переходів; $\{Fn\}$ – підмножина вершин, що представляють кінцеві події (End Event); $\{E(Gr)\}$ – множина ребр, які зв'язують вершини в напрямлені ребра $(v_i, v_{i+1}) \in E(Gr)$ спрямованого графа; $\{L(v)\}$ – множина міток вершин, які дозволяють однозначно визначати кожну вершину $V(Bp_N)_i$ у спрямованому графі $Gr(Bp_N)$; $\{Id(v)\}$ – функція розмітки вершин для надання унікальних імен кожній вершині, яка має вигляд $Id(v): V(Bp_N) \rightarrow L(v)$; $\{Rul^{(Bp)}\}$ – множина правил зв'язування (функціонального, інформаційного або логічного) графічних елементів нотації, які враховують семантичні обмеження на кожний графічний елемент та умови розгалуження потоків в шлюзах моделі; $p(v): E(Gr) \rightarrow [0,1]$ – булева функція – зіставляє значення ймовірності ребрам, які зв'язують вершини спрямованого графа.

У наслідок суб'єктивного сприйняття аналітиком специфікацій Про завжди виникають помилки моделювання процесів. Ці помилки пов'язані як з компетентністю аналітика, так і з недоліками суто формального опису процесів, які притаманні нотації BPMN. До найбільш вагомих можна віднести такі помилки, які пов'язані з семантичної неоднозначністю елементів нотації BPMN:

- недостатньо повно визначено умови настання кінцевих подій (End Event) Fn окремих підпроцесів складного сценарію;
- недостатньо чітко визначені умови настання проміжних подій сценарію, що призводить до виключення подій з потоку (Process Flow) або неправильного виконання процесу;
- недостатньо пов'язані завдання або підпроцеси сценарію - оскільки не враховані особливості різних об'єктів потоку і правила їх з'єднання;

– нечітко визначені складні умови багатофакторних переходів в логічних операторах (Gateway). Оскільки в нотації BPMN ребра вершини графа описуються булевими змінними $p(v): E(Gr) \rightarrow [0,1]$, а це суттєво ускладнює коректне формування багатофакторних умов із залученням механізму тригерів подій, що також призводить до виключення подій з потоку або до неправильного виконання процесу (Process Flow).

Для усунення цих помилок та формування коректної моделі складного процесу АнД бажано мати додатковий формальний опис – семантичну модель сценарію на мові XML. Формально концептуальну модель онтології ПрО – $Ont(SD)_{BP}$, що описує сценарій АнД та співвідноситься з нотацією BPMN, можна представити у вигляді математичного виразу

$$Ont(SD)_{BP} = \langle C^{(E)}, An^{(C)}, Rel^{(H)}, T^{(A)}, Ax^{(S)}, Rul^{(S)}, Ex^{(C)} \rangle$$

де: $\{C^{(E)}\}$ – підмножина класів (<Classes>) понять ПрО, які співвідносяться з основними категоріями нотації BPMN <Event>, <Task>, <Sub-Process>, <Gateway> та інші; $\{An^{(C)}\}$ – підмножина <Annotation Property>, що використовується для передачі концепту онтології семантики при описі процесів; $\{Rel^{(H)}\}$ – підмножина відносин <Object properties>, які виражають відносини між класами онтології ПрО і співвідносяться з категоріями нотації BPMN <Connecting objects>; $\{T^{(A)}\}$ – підмножина атрибутів <Data properties> для опису властивостей класів $\{C^{(E)}\}$ (назви, їх типи, області значень даних тощо); $\{Ax^{(S)}\}$ – підмножина аксіом <Axioms>, які задають правила виведення для встановлення співвідношень між <Classes> онтології і основними нотації BPMN. Аксіоми записуються предикатами першого порядку і задають правила виводу для встановлення співвідношення між підмножиною класів (<Classes>) онтології ПрО та основними класами нотації BPMN; $\{Rul^{(S)}\}$ – правила логічного висновку для встановлення дійсного співвідношення між <Classes> онтології ПрО і графічними категоріями нотації BPMN; $\{Ex^{(C)}\}$ – множина екземплярів (<Instances>) класів $\{C^{(E)}\}$, що використовується для встановлення співвідношення між класами онтології конкретного сценарію і екземплярами графічних елементів, які описують структурну діаграму в нотації BPMN.

Взаємодія між конструкціями графічної нотації BPMN та OWL онтологією ПрО показана на рисунку 3. Напрям відображення інтерпретації моделі сценарію АнД описує відображення (Mapping) від нотації BPMN до онтології ПрО, а відображення уявлення – зворотне відображення.

Для ефективної конвертації отриманої BPMN-моделі в OWL-модель сценарію АнД, яка здатна відповідати на семантичні DL-запити, графічні елементи BPMN-моделі мають бути пов'язані з текстовою документацією відповідно до формату системи моделювання BPMN. Умовою взаємодії моделей теоретично є підтримка співвідношення 1:1 між концептами OWL онтології та примітивами BPMN нотації,

які входять до складу спрямованого графу моделі сценарію, але на практиці співвідношення можуть бути неоднозначними. Постає задача визначення співвідношень 1:1.

Це вирішується на другому етапі методу за рахунок побудови тривірневої моделі онтології отриманих BPMN-моделей для кожного концепту (рис. 3), що є розвитком положення теорії побудови моделі багаторівневої онтології (Bunge-Wand-Weber, BWW). Другий рівень (найбільш складний) складає мета-модель BPMN 2.0, яка і визначає співвідношення 1:1 класів онтології до графічних примітивів (Core Element Set), дозволяє відобразити графічні елементи BPMN в поняттях онтології.

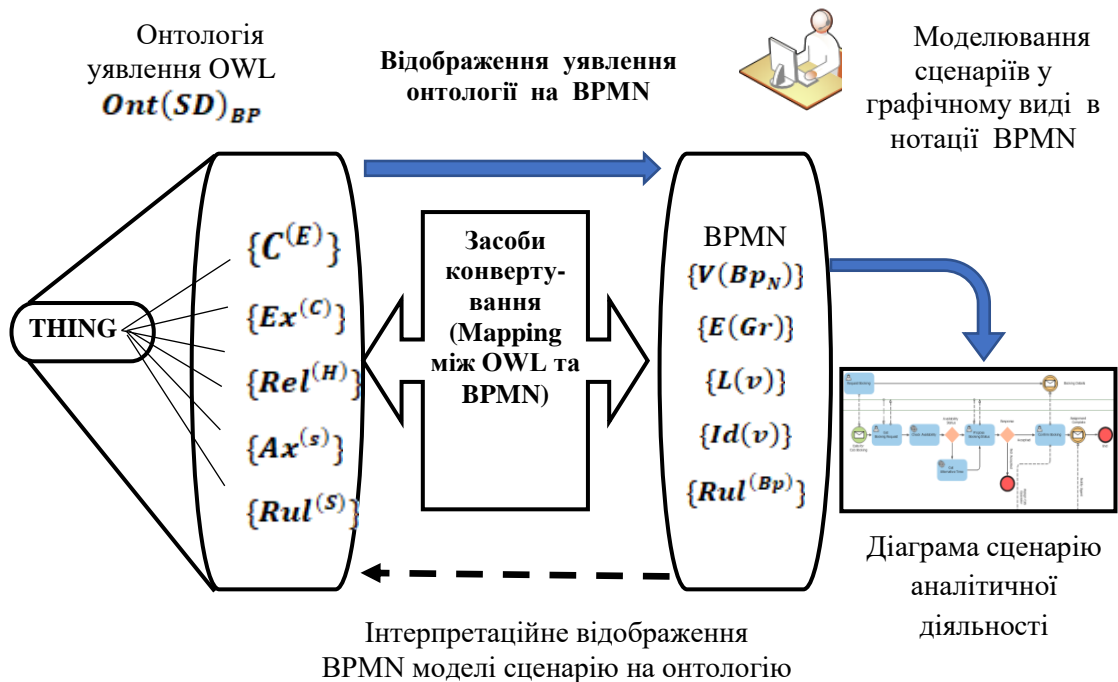


Рисунок 3 – Взаємодія між графічними конструкціями нотатції BPMN та онтологією Про

В теорії конвертація BPMN діаграм в OWL-моделі призводить до створення БЗ, що розкриває можливість здійснення семантичного аналізу із застосуванням відомих алгоритмів аналізу за допомогою семантичних запитів Web-мовами: DL Query, SPARQL, Semantic Web Rule Language та іншими. Отже, онтологічна мета-модель BPMN 2.0 являє собою деталізований опис класів (графічних примітивів), процесів, учасників, інформаційних ресурсів, умов розгалуження потоків з можливістю співвідношення елементів діаграми з конкретними екземплярами діаграми для дослідження якості OWL-моделі на рівні знань. Мета-модель BPMN 2.0 є основою для формування реальної моделі сценарію за рахунок механізмів зв'язування класів мета-моделі з сукупністю екземплярів, які саме й описують реальну онтологію сценарію, надаючи унікальні імена для кожного екземпляра класів (рис. 4).

Процес створення онтології виконується для кожного окремого сценарію та містить наступні кроки: а) кожний сценарій аналізують та визначають усі необхідні для її створення сутності та відношення, обмеження та операції; б) будують концептуальну модель сценарію з використанням виявлених компонент, причому ця модель повинна у першу чергу використовувати компоненти з існуючої мета-моделі

онтології BPMN; в) онтологію сценарію валідують – визначають її відповідну повноту та вирішують можливі протиріччя, особливо умови логічно-незв'язаних переходів. Побудова онтології на основі BPMN моделі сценарію дозволяє провести семантичну *валідацію* сценарію, що в традиційних підходах здійснюється експертами на основі суб'єктивних оцінок.

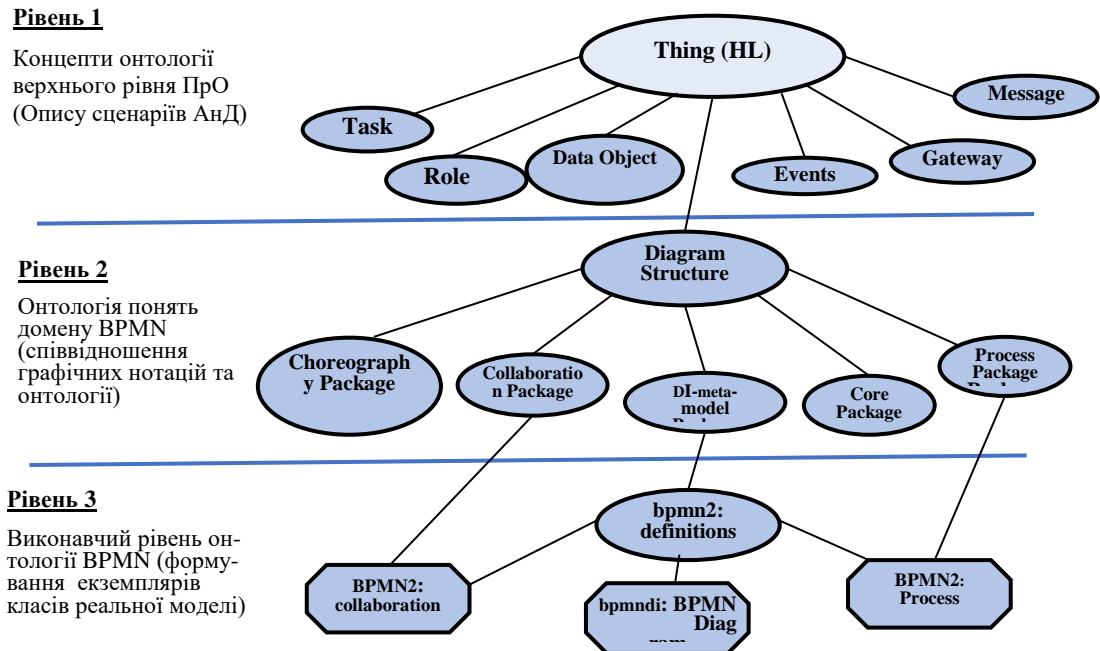


Рисунок 4 – Базова структура OWL мета-моделі нотатції BPMN 2.0

Накопичування формалізованих знань про побудову сценаріїв дає можливість полегшити процес моделювання складних сценаріїв АнД та зменшити його час за рахунок його інтелектуалізації. Одним із проявів інтелектуалізації є накопичення знань о діях користувача-аналітика. В цьому контексті сенс накопичення знань полягає в тому, щоб програмний засіб був в змозі самостійно класифікувати нові дані, одержувані при аналізі, і пропонувати користувачу найбільш доцільні кроки сценарію, виходячи з накопичених знань та ситуації що склалася на попередніх кроках. Але при цьому у контексті виконання багаторівневих сценаріїв АнД, одним із ключових питань залишається вирішення задачі забезпечення семантичної сумісності процесів, з яких складається будь-який складний сценарій АнД.

У третьому розділі пропонуються підходи для вирішення задачі забезпечення семантичної сумісності процесів моделювання кроків сценарію АнД (особливо при здійсненні переходів в багаторівневому сценарію) за рахунок застосування онтології ПрО для вирішення конфліктів, пов'язаних з гетерогенністю середовища, в якому ці сценарії реалізуються, та обґрунтовано інтелектуалізацію ПС процесу побудови сценарію АнД на основі онтологічних моделей задач та опису і організації знань для побудови сценаріїв.

Для визначення того, чи сумісні процеси при визначенні умов переходів між кроками виконання складного сценарію АнД, викладено та доведено набір аксіом та теорема, які теоретично обґрунтовують умови забезпечення семантичної сумісності переходів. Запропоновано підхід для розв'язання розбіжностей, які виникають у

переходах. Цей підхід засновано на базі онтологічної моделі семантичного медіатора, якій застосовується аналогічно підходу у технологіях Semantic Web Service.

З теоретичної точки зору виконання кроків складного сценарію АнД описується двома принципово різними компонентами: даними та процесами. Модель семантичного медіатора базується на кількох функціях забезпечення сумісності, які включають головним чином:

- семантичну анотацію даних для перевірки коректності логічних переходів між різними кроками сценарію;
- семантичне вирівнювання гетерогенних виконавчих комп'ютерних моделей при взаємодії різних рівнів складного сценарію;
- загальну модель даних для єдиного представлення ПрО сценаріїв АнД;
- онтологічні моделі ПрО сценаріїв для підтримки процесів семантичного посередництва при вирішенні розбіжностей, які виникають у переходах між кроками виконання складного сценарію АнД.

Під час виконання сценарію АнД семантичний медіатор здійснює автоматичний аналіз умов переходів з $S_n(Pr_i)$ кроку сценарію на наступний $S_{n+1}(Pr_j)$ крок. При цьому внутрішні рішення, прийняті в середині $S_n(Pr_i)$ кроку, у цьому випадку не мають значення, оскільки медіатор під час роботи діє на рівні повідомлень (умови переходів), надісланих та отриманих під час взаємодії кроків сценарію. На попередньому кроці можуть виникнути процеси, які формують умови переходів $Cond_{S_i}$. На рис. 5 показано роль та функції медіатора при розв'язанні розбіжностей в складному переході між двома кроками сценарію.

Для представлення зв'язку між двома процесами переходу між кроками сценарію $S_n(Pr_1)$ та $S_{n+1}(Pr_2)$ використовуються такі позначення:

$$\frac{Cond_{S_i} \text{ sequence of } S_n(Pr_1)}{Cond_{S_i} \text{ sequence of } S_{n+1}(Pr_2)}$$

Якщо перехід між двома кроками має більше двох процесів, як це показано на рисунку, для представлення зв'язку між кроками можна використовувати кілька пар за кількістю можливих станів переходів:

$$\frac{Cond_{S_i} \text{ sequence of } S_n(Pr_1) \text{ in interaction with } S_{n+1}(Pr_1)}{Cond_{S_i} \text{ sequence of } S_{n+1}(Pr_1) \text{ in interaction with } S_n(Pr_1)}$$

$$\frac{Cond_{S_i} \text{ sequence of } S_n(Pr_1) \text{ in interaction with } S_{n+1}(Pr_2)}{Cond_{S_i} \text{ sequence of } S_{n+1}(Pr_2) \text{ in interaction with } S_n(Pr_1)}$$

$$\frac{Cond_{S_i} \text{ sequence of } S_n(Pr_1) \text{ in interaction with } S_{n+1}(Pr_j)}{Cond_{S_i} \text{ sequence of } S_{n+1}(Pr_j) \text{ in interaction with } S_n(Pr_1)}$$

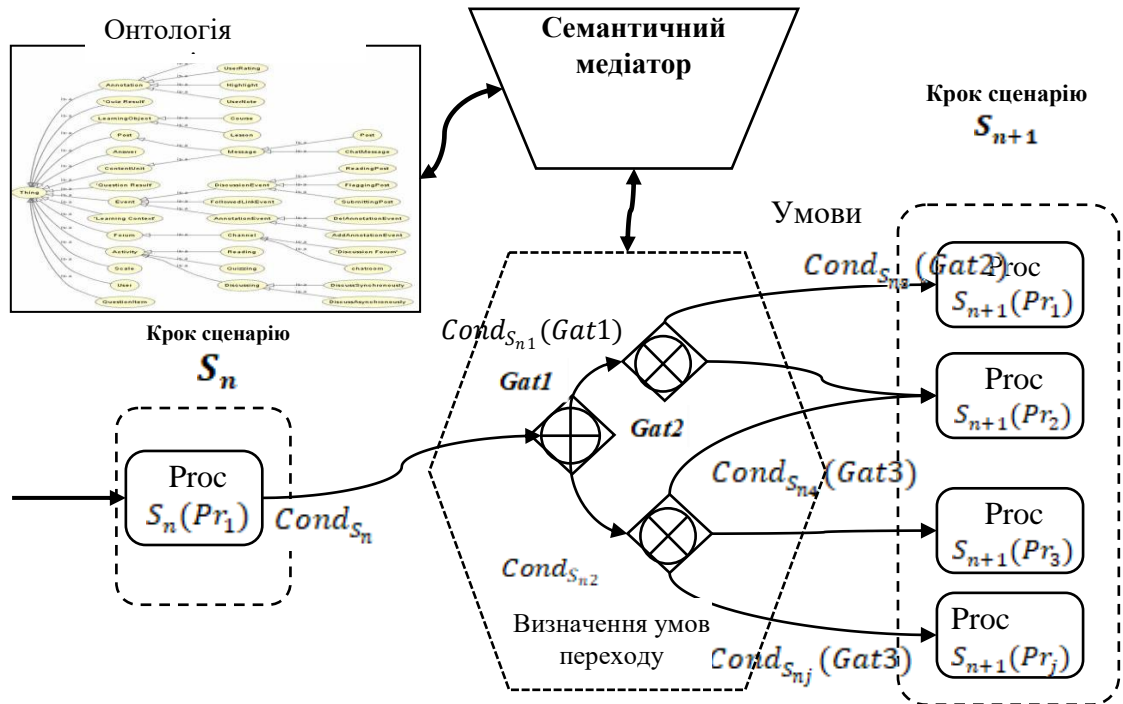


Рисунок 5 – Місце посередника при розв'язанні розбіжностей в складному переході

Порядок умов переходів у послідовності зображено символом \rightarrow . Наприклад, процес кроку $S_n(Pr_1)$ взаємодіє з процесом $S_{n+1}(Pr_2)$, якщо виконуються наступні умови:
ЯКЩО

$$S_n(Pr_1) = S_{n+1}(Pr_2) ,$$

де

$$S_{n+1}(Pr_1) = Cond_{S_i} \rightarrow Cond_{S_{n1}}(Gat1)^1 \rightarrow Cond_{S_{n1}}(Gat2)^2,$$

та

$$S_{n+1}(Pr_1) = Cond_{S_i} \rightarrow Cond_{S_{n1}}(Gat1)^3 \rightarrow Cond_{S_{n1}}(Gat3)^1.$$

Перехід між двома кроками складного сценарію зазвичай складається з багатьох умов, тому пошук еквівалентності у шаблонах умов переходів до наступних кроків – це зовсім не тривіальне завдання. Інтуїтивно найпростіший спосіб це спочатку визначити невідповідності, а потім шукати спосіб їх усунення. В процесі обробки умов переходів ідентифікуються три можливі випадки: повна відповідність, вирішені невідповідності та невирішені невідповідності. У випадку «вирішені невідповідності» можна класифікувати типові ситуації, які можуть бути автоматично вирішені медіатором з використанням онтологічного опису переходів сценарію:

1. Процес $S_n(Pr_1)$ надсилає повідомлення $Cond_{S_i}$, тоді як процес $S_{n+1}(Pr_1)$ не готовий його отримувати.

2. Процес $S_n(Pr_1)$ надсилає одразу кілька екземплярів повідомлень $Cond_{S_i}^{e1}$, $Cond_{S_i}^{e2}$, $Cond_{S_i}^{en}$, тоді як процес $S_{n+1}(Pr_1)$ хоче отримати цю інформацію як частину кількох повідомлень.

3. Процес $S_n(Pr_1)$ надсилає повідомлення в порядку $Cond_{S_i}^{e1}$, $Cond_{S_i}^{e2}$, $Cond_{S_i}^{en}$ іншому порядку, ніж їх очікує процес $S_{n+1}(Pr_1)$. Порядок надісланих повідомлень може бути автоматично інвертований до проєкції, визначеної у відповідних аксіомах та лемах, які виведено у дисертаційній роботі.

4. Процес $S_n(Pr_1)$ надсилає повідомлення $Cond_{S_i}$ яке отримує процес $S_{n+1}(Pr_1)$, і очікує підтвердження того, що повідомлення було отримано цільовим процесом.

Існує набір невідповідностей, які посередник процесу SCM не може автоматично вирішити. Ці розбіжності стосуються неіснуючих повідомлень, які не відповідають семантичним умовам переходу. Наприклад, процес $S_{n+1}(Pr_1)$ очікує повідомлення $Cond_{S_i}$, яке не було надіслано процесом $S_n(Pr_1)$ до певного часу. Інший випадок розглядає ситуацію, коли процес $S_{n+1}(Pr_1)$ отримав повідомлення, але процес $S_n(Pr_1)$ очікує підтвердження його отримання.

В реальності взаємодія між кроками сценарію складається не лише з елементарних невідповідностей, представлених вище, але може включати й проєкції процесів, які при певних умовах розглядаються як семантичні еквіваленти процесів – проєкцій перетворення, метою отримання яких є формалізація умов переходів для більш коректного застосування онтології сценарію при вирішенні невідповідностей з залученням медіатора. Для цього досліджується як семантична сумісність між двома прогнозами впливає на сумісність прогнозованих процесів.

Теорема. *Будь-які два процеси сценарію $S_n(Pr_i)$ та $S_{n+1}(Pr_j)$ вважаються сумісними (семантично зв'язаними) тоді і тільки тоді, коли їх існуючі проєкції $S_n(Pr_1)^k$ та $S_{n+1}(Pr_1)^f$ відповідних процесів сценарію, теж повністю відповідні*

$$\forall S_n(Pr_i), S_{n+1}(Pr_j), \text{compatible}(S_n(Pr_i), S_{n+1}(Pr_j)) \Leftrightarrow$$

$$\exists k, f \text{ match}(S_n(Pr_1)^k, S_{n+1}(Pr_1)^f)$$

$$1 \leq i \leq k \text{ та } 1 \leq j \leq f,$$

де k і f - кількість можливих проєкцій перетворення для процесів $S_n(Pr_i)$ та $S_{n+1}(Pr_j)$, відповідно; $S_n(Pr_1)^k$ та $S_{n+1}(Pr_1)^f$ - позначення проєкцій, які вважається семантичними еквівалентами відповідних процесів та можуть використовуватися для аналізу умов переходів між кроками сценарію із застосуванням онтології ПрО.

Для отримання проєкцій використовується набір аксіом перетворення, який наведено в дисертації.

З теореми слідує, що процеси сценарію $S_n(Pr_i)$ та $S_{n+1}(Pr_j)$ вважаються сумісними, якщо хоча би одна пара проєкцій $S_n(Pr_1)^k$ та $S_{n+1}(Pr_1)^f$ є відповідною – *match*.

Демонстрація застосування теореми при вирішенні невідповідностей при взаємодії кроків сценарію передбачає доведення положення теореми в обох напрямках, тобто:

$$\begin{aligned} & \forall S_n(Pr_i), S_{n+1}(Pr_j), \text{compatible} (S_n(Pr_i), S_{n+1}(Pr_j)) \\ & \Rightarrow \exists k, f \text{ match} (S_n(Pr_1)^k, S_{n+1}(Pr_1)^f) \end{aligned}$$

та

$$\begin{aligned} & \exists k, f \text{ match} (S_n(Pr_1)^k, S_{n+1}(Pr_1)^f) \\ & \Rightarrow \forall S_n(Pr_i), S_{n+1}(Pr_j), \text{compatible} (S_n(Pr_i), S_{n+1}(Pr_j)) \end{aligned}$$

Рівноважність цього твердження доводиться шляхом застосування принципу суперечності та індукції. З припущення, що $\exists k, f, \text{match} (S_n(Pr_1)^k, S_{n+1}(Pr_1)^f)$, впливає наступне:

$$\begin{aligned} & \nexists k, f, \text{match} (S_n(Pr_i)^k, S_{n+1}(Pr_j)^f) \Leftrightarrow \\ & \forall S_n(Pr_i), S_{n+1}(Pr_j), \neg \text{match} (S_n(Pr_i), S_{n+1}(Pr_j)) \end{aligned}$$

та

$$\begin{aligned} & \forall S_n(Pr_i), S_{n+1}(Pr_j), \neg \text{match} (S_n(Pr_i), S_{n+1}(Pr_j)) \\ & \Rightarrow \neg \text{compatible} (S_n(Pr_i)^k, S_{n+1}(Pr_j)^f) \end{aligned}$$

Для доведення цих положень використовується метод вичерпання, тобто розглядаються всі можливі випадки, кожний з яких аналізується на предмет відповідності положень теореми.

Теорема представляє визначення сумісності, яке використовуються медіатором. Концептуально алгоритм, який реалізується медіатором, складається з таких кроків:

- розклад будь-якого повідомлення до рівня атомарної інформації, яку воно несе (тобто до рівня екземпляра);
- якщо процеси використовують різні базові онтології, для кожного отриманого екземпляра створюється відповідний екземпляр з точки зору онтології другого процесу;
- здійснюється семантичний аналіз умов переходів та вирішення проблем невідповідності процесів для визначення того, яка інформація очікується та у якому форматі;

– зберігання результатів вирішення проблем в БЗ у вигляді окремих екземплярів.

Вирішення задачі забезпечення семантичної сумісності процесів моделювання кроків сценарію Анд дає можливість перейти до вирішення завдання інтелектуалізації процесу побудови сценарію Анд, Загальна архітектура інтелектуальних ПЗ, що підвищують адаптивні можливості ПС шляхом накопичення та повторного використання знань у формі онтологічних моделей сценаріїв Анд, показано на рисунку 6. Отже, архітектура містить необхідні інструменти для здійснення моніторингу за діями користувачів, засоби консолідації інформації в сховище даних, витягу, перетворення, трансформації даних з метою їх адаптації до алгоритмів data mining (зокрема, методів машинного навчання деревами класифікації і регресії – Classification and Regression Trees), а також засоби візуалізації результатів та можливості корегування сценаріїв на підставі аналізу попередніх кроків.

Одною з головних вимог до сучасних аналітичних ПС є комфортність самого процесу спілкування з системою за рахунок її інтелектуалізації. Тобто аналітик сподівається на дружній інтерфейс ПС, який може пропонувати йому найбільш ймовірний крок сценарію, виходячи з аналізу попередніх дії та накопичених знань. Для підвищення рівня ефективності Анд та враховуючи зростаючу складність самих процесів та технологій аналізу даних, постає задача автоматизації процесу формування сценаріїв та його коригування виходячи з парадигми інтелектуалізації ПС.

Одним з підходів вирішення цієї задачі є використання методів машинного навчання. Як витікає з рис. 6, побудова потрібного для вирішення задач Анд сценарію, – це складний процес, який включає формування початкових умов для його побудови, наявність інтегрованої моделі опису знань та основних типових сценаріїв.

Інтегрована модель опису знань - $InM(KB)$, заснована на парадигмі Semantic Web, на підставі якої можуть поєднуватися різні моделі взаємодії в певній Про. Інтегрована модель описується математичним виразом:

$$InM(KB) = \langle Ont(SD), SNet, FScen, PCase, PRes \rangle,$$

де $Ont(SD)$ – прикладна онтологія, що задає опис основних сутностей сценарію у вигляді класів цілей, задач та результатів аналітичного дослідження та опису IP, необхідних для здійснення Анд в певній Про

$$Ont(SD) = \langle C^{(SD)}, R^{(In)}, Rb^{(C)}, T^{(SD)}, Ac^{(C)}, Cntr, Rul^{(S)} \rangle;$$

$SNet$ – об'єктно-орієнтована семантична мережа (у вигляді графової моделі), яка описує взаємодію (верхнього рівня) об'єктів прикладної онтології та IP, баз даних, повідомлень, які складають невід'ємну частину сценарію виконання Анд; $FScen$ – модель виконавчих сценаріїв Анд (як окремих процесів обробки даних $\langle Pool \rangle$, так і дій учасників $\langle Participant \rangle$) у вигляді сукупності окремих процесів $\langle Pool \rangle$, яка

описує задачі, типові процеси обробки даних, аналітичні функції та процедури, притаманні для дій аналітика у цієї сфері. На базі компонентів семантичної мережі формуються виконавчі сценарії Анд (у вигляді загального графу або сукупності процесних графів) для здійснення операцій над екземплярами класів прикладної онтології.

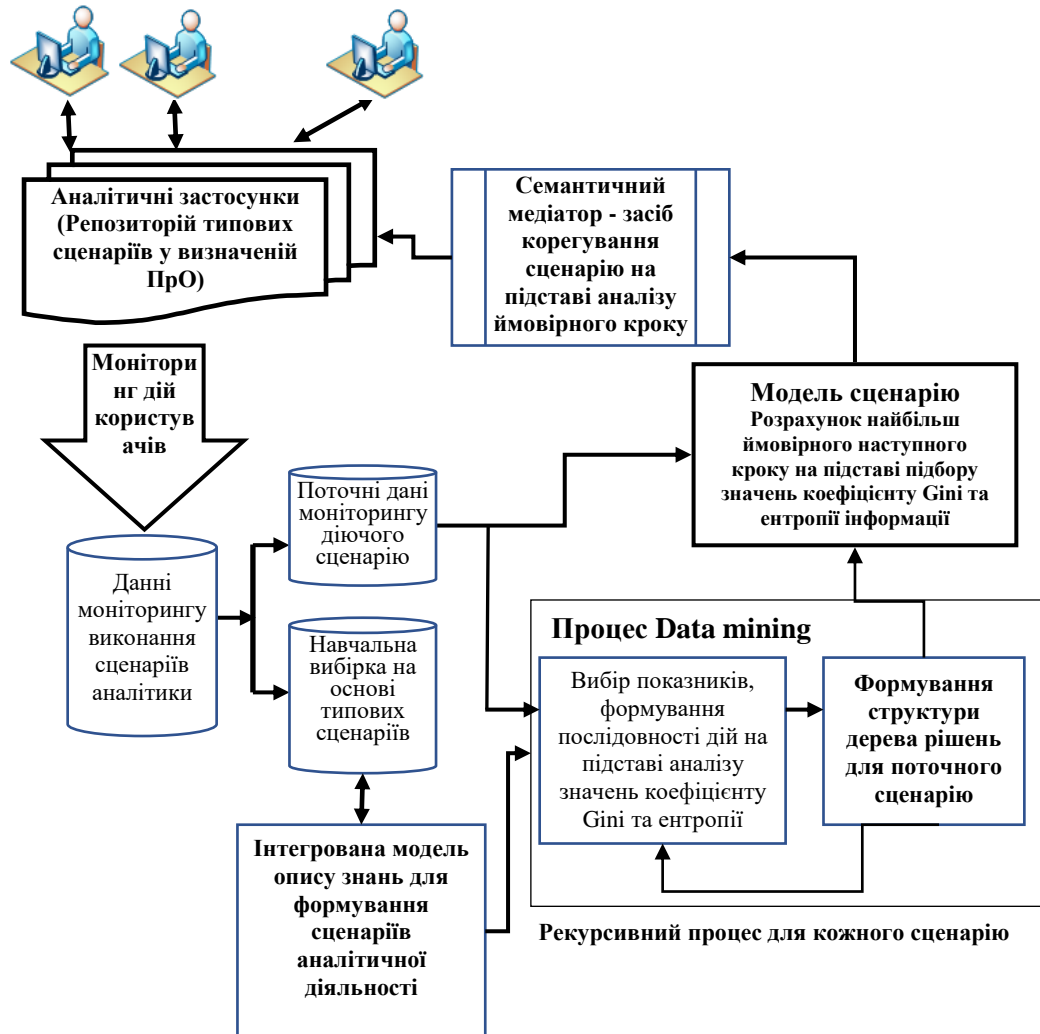


Рисунок 6 – Роль та місце інтелектуальних програмних засобів при виконанні сценаріїв аналітичної діяльності

Модель сценаріїв Анд враховує реальні обмеження на значення атрибутів об'єктів семантичної мережі *SNet*; *Pres* – модель механізмів виведення нових знань та обробки даних з метою отримання звітних матеріалів та візуалізації результатів Анд. Для сучасних ПС ця модель повинна містити відповідні механізми як чіткого виведення (на основі правил та аксіом), так і механізми виведення на основі нечітких правил, які маніпулюють з лінгвістичними змінними; *PCase* – виконавча модель ПЗ Анд, яка складається з певної сукупності окремих програмних модулів, призначенням яких є реалізації сценаріїв, множині відповідних аналітичних функцій, процедур обробки даних та виводу нових знань.

Для універсализації опису знань, які використовуються для формування навчальних вибірок та найбільш ймовірних сценаріїв аналізу, пропонується дотримуватися наступних формальних правил:

- поняття та відносини, які представляються класами ПрО та ПЗ, мають інкапсулювати в собі семантичні властивості та обмеження – $\{Cntr\}$, що задаються на атрибутах класів $Ac^{(c)}$ та записуються логічними виразами виду: $Cntr(Cn_{i1}, \dots, Cn_{iw})$, де $Cn_{ik} \in Ac^{(ci)}$ або $Cn_{ik} \in T^{(SDi)}$, та може бути або ім'ям атрибута, або константою, для завдання детермінованих умов виводу знань за допомогою відповідних програм.

- об'єкти та відносини семантичній мережі – $SNet$ обов'язково мають бути визначені в моделі прикладної онтології - $Ont(SD)$;

- обмеження на значення атрибутів об'єктів представляються у функціональній моделі – $FScen$ у вигляді обмежень, які накладаються на сценарії виконання функції та процедур, що реалізують завдання системи;

- все процеси виведення і обробки інформації мають здійснюватися системою продукційних правил – $PRul$, яка пов'язана з семантичною мережею $SNet$ та виконавчими сценаріями $FScen$, що реалізують завдання Анд.

Прикладна онтологія $Ont(SD)$ задає комплексний опис понять і відносин як ПрО, так і ПЗ, що його реалізує. Такий опис фактично визначає структуру семантичної і функціональної мережі, за допомогою якої представляються і знання про потрібний сценарій, а також знання про функції (операції), які доцільно виконувати при вирішенні конкретних аналітичних завдань. Слід також підкреслити, що за допомогою семантичної мережі задаються сутності ПрО, в термінах яких описуються продукційні правила логічного виводу нових знань. Завдяки інтегрованої моделі з бази знань ПС формується модель подання знань, яка найбільш співвідноситься з цілями та завданнями аналітичного

Процес інтелектуалізації програмних засобів запропонований за використанням інтелектуального методу визначення наступних подій сценарію Анд, якій будується на базі удосконалення методу Classification and Regression Trees (CART) та відрізняється від нього наявністю і можливістю використання інтегрованої моделі опису знань $InM(KB)$ та можливістю застосування різних метрик при аналізі якості розбиття та через нього вибору наступного кроку ймовірної дії аналітика. На відміну від існуючих підходів при виконанні машинного навчання, пропонується можливість вибору найбільш оптимальної метрики оцінки якості наближення до бажаного результату навчання – коефіцієнту Gini або методу розрахунку ентропії корисності інформації.

Формально аналітичний процес представляється орієнтованим графом – $Gr(Sc)$, який відображує послідовності можливих дій користувача по досягненню мети аналітичного дослідження. Для більшої реалістичності вважаємо, що деякі дані про дії можуть бути або втрачені, або спотворені (зображені у вигляді порожнього кола). Крім того, одна з дій на графі позначена не символами, а цілим числом, яке розглядається як граничне значення (рис. 7). Завдання полягає в тому, щоб з накопичених даних (послідовності кроків при вирішенні певного кола завдань)

отримати знання, на основі яких інтелектуальний засіб (алгоритм CART) в змозі запропонувати найбільш вірогідну дію користувача (процедуру обробки даних), виходячи з аналізу попередніх кроків та спираючись на знання про семантику предметної області.

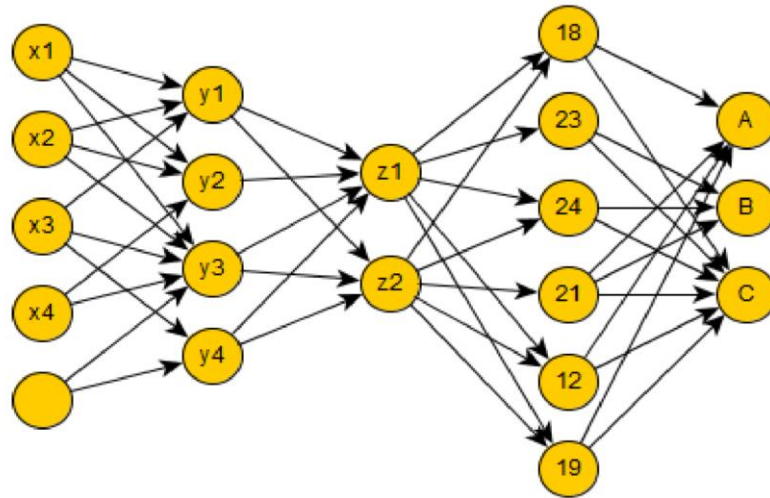


Рисунок 7 – Граф всіх варіантів можливих дій користувача

Першим кроком запропонованого методу є опис матриці всіх можливих станів орієнтованого графа – $Gr(Sc)$ (включаючи й втрачені або спотворені дані). В останньому елементі кожного рядку матриці вказується кінцева дія користувача ("А", або, "В", або "С"), виконання якої потрібно спрогнозувати – передбачити.

Передбачення ймовірних дій користувача потребує, по-перше, навчання інтелектуальної системи на підставі контрольної вибірки, а, по-друге, формування обґрунтованої гіпотези щодо можливості досягнення останнього елемента кожного рядку матриці станів, тобто забезпечення працездатності усіх вузлів графу – передачі даних для аналізу від одної дії користувача до другої.

Умова забезпечення працездатності усіх вузлів графу:

$$\forall v_i \in V \Leftrightarrow \omega_i(\tau) = 1, \quad \tau \in 0, t),$$

де $V = \{v_i\}$ – множина вузлів графу; $\omega(\tau)$ – булева функція, що набуває значення 1, якщо вузол дій приймає участь в діях користувача і 0 – у іншому випадку; τ, t – поточний час дій у вузлів обробки аналітичних даних.

Умова забезпечення передачі даних між вузлами графу по основних або резервних маршрутах дій користувача для досягнення мети аналітичного дослідження:

$$\forall v_i, v_j \Leftrightarrow \exists v_j \in \Gamma_i, \quad i, j = 1, 2, \dots, n,$$

де Γ_i – множина досяжних вузлів графу з вузла v_i ; n – число вузлів дій, що приймають участь в аналітичному процесі.

Із сформульованих умов витікає необхідна умова стійкості аналітичного процесу, тобто можливість виконання аналітичного процесу для досягнення мети

дослідження в усіх вузлах орієнтованого графа. Дана умова полягає в наявності знань про процес Анд у вузлі та наявність маршрутів, які відображають послідовність дій аналітика для досягнення мети дослідження:

$$\forall G(V, L), v_i \in V, l_{ij} \in L \Leftrightarrow \begin{cases} \forall v_i \in V \Leftarrow \omega_i(\tau) = 1, \tau \in [0, t); \\ \forall v_i, \forall v_j \Leftarrow \exists v_j \in \Gamma_i, i, j = \overline{1, n}. \end{cases}$$

Необхідна і достатня умова стійкості аналітичного процесу, який відображує послідовності дій аналітика щодо досягненню мети дослідження полягає в наявності знань про процес аналітичної діяльності і наявності альтернативних маршрутів передачі інформації:

$$\forall G(V, L), v_i \in V, l_{ij} \in L \Leftrightarrow \begin{cases} \forall v_i \in V \Leftarrow \omega_i(\tau) = 1, \tau \in [0, t); \\ \forall v_i, \forall v_j \Leftarrow \exists v_j \in \Gamma_i^{\geq 2}, i, j = \overline{1, n}. \end{cases}$$

Отже, поточна структура знаходиться на межі стійкості аналітичного процесу, якщо граф отриманої структури зв'язний, та має у своєму складі мости ($N_L \geq 1$) - ребро графа, що сполучає два підграфи, або вузли з'єднання (кореневі вузли) ($N_V \geq 1$):

$$\{K = 1\} \wedge [\{N_V \geq 1\} \vee \{N_L \geq 1\}],$$

де K – це число компонентів графа, а умова $K = 1$ означає, що граф зв'язний; N_V, N_L – число вузлів зв'язності та мостів графа відповідно.

Наявність в структурі моста або вузла зв'язності (кореневого вузлу), які сполучають два підграфи, означає, що усі маршрути дій користувача для досягненню мети дослідження з вершин одного підграфу у вершини іншого включатимуть цей міст або вузол зв'язності. Ця подія істотним чином знижує стійкість аналітичного процесу. Аналіз структур показує, що якщо граф, що описує аналітичний процес, знаходиться на межі стійкості, то система працездатна та виконує встановлений об'єм функцій. Однак, у випадку хоча б однієї відмови моста або вузла зв'язності система переходить у нестійкий стан.

Тому другим кроком запропонованого методу є створення кореневого вузлу графу $Gr(Sc)$ з метою вибору найкращій змінної, виходячи з посилу, що досягнення кінцевих дій можливе лише за стовпцем матриці всіх можливих станів орієнтованого графа (кроком n), який містить значення **Z1** або **Z2** (Рис. 7). Для цього спочатку необхідно виконати обчислення неоднорідності вхідної множини значень, які містить матриця можливих дій. Мірою неоднорідності множини є ентропія. Для підвищення якості розбиття скористаємося коефіцієнтом Gini (Gini coefficient) – $K_{Gini}(D)$, який є мірою нерівності розподілу деякої величини (D), що приймає значення між 0 і 1, де 0 означає абсолютну рівність, а 1 позначає повну

нерівність. Цей коефіцієнт $K_{Gini}(D)$ використовується у деревах рішень при виборі розподілу деякої величини (D). За допомогою коефіцієнта Gini вимірюється двійковий розділ для кожного атрибута – вірогідність появи різних подій P_1 в листі дерева. Коефіцієнт Gini розраховується по формулі

$$K_{Gini}(D) = 1 - \sum_{i=1}^n P_i^2$$

Обчислення розподілу значень матриці можливих дій дерева рішень складається з наступних кроків:

1. Обчислення $K_{Gini}(D)$ для кожного значення рядку – тобто вірогідність досягнення кожного з можливих результатів приналежності до певного розбиття. Вірогідність обчислюється шляхом ділення лічильника появ цього результату на загальне число рядків в множині.

2. Розрахування вірогідності для кожного рядку методом підсумування. Результатом цього кроку є сумарна вірогідність $\sum_{i=1}^n P_i^2$, яка показує, що для випадково обраного рядка може бути прогнозовано не той результат, який насправді має місце. Слід нагадати, що чим вище вірогідність, тим гірше розбиття. Оскільки вірогідність 0 асоціюється з найкращим результатом, це означає що всі рядки вже розподілені правильно.

3. Оскільки розділення для визначеного атрибута (наприклад – $Z1$) розпадається на дві частини $D1$ і $D2$, то $K_{Gini}(D)$ для цього розподілу обчислюється за формулою:

$$K_{Gini_{Z1}}(D) = \frac{|D1|}{|D|} K_{Gini}(D1) + \frac{|D2|}{|D|} K_{Gini}(D2)$$

Аналогічно виконується розрахунок $K_{Gini_{Z1}}(D)$ для атрибута $Z2$.

4. Наступним кроком є розрахунок ентропії інформації по Шенону – $H_{gr}(D)$, яка обчислюється по формулі:

$$H_{gr}(D) = \sum_{i=1}^I P_i \log_2 \frac{1}{P_i} = - \sum_{i=1}^I P_i \log_2 (P_i)$$

Для отримання оцінки того наскільки хороша обрана змінна, алгоритм спочатку обчислює ентропію всієї групи $H_{gr}(D)$, а потім намагається розбити групу по можливих значеннях кожної змінної та обчислює ентропію двох нових груп. Найкраща розбивка визначається обчисленням інформаційного виграшу – різницею між поточною ентропією і середньозваженої ентропією двох нових груп, після чого вибирається та, для якої інформаційний виграш максимальний.

5. Далі алгоритм створює дві гілки: якщо умова істинна, та якщо умова помилкова. Обчислюючи для кожного вузла найкращий атрибут і розщеплюючи гілки, алгоритм будує дерево рішень.

6. Побудова дерева рішень полягає в обчисленні найкращого критерію розщеплення «True» або «False», в якості якого використовується середньозважена ентропія, яка розраховується для кожної пари. Зростання гілки припиняється, якщо інформаційний виграш, отриманий від розщеплення в даному вузлу, виявляється менше або дорівнює нулю.

В процесі побудови дерева рішень може виникнути ситуація коли деякі дані можуть бути втрачені або спотворені, то з метою усунення такої ситуації запропоновано внесення змін до функції прогнозування наступного кроку, які враховують відсутність даних попереднього етапу – застосовуючи метод оцінки ентропії $H_{gr}(D)$.

За результатом роботи запропонованого методу формується дерево рішень, в якому можливі переходи до різних станів позначено як «True» та «False». Для спрощення безпосередньо процесу вибору наступного кроку дій при формування сценарію Анд («True» або «False»), алгоритм методу доповнено більш зручним механізмом формування семантичних умов переходу у вигляді послідовності операторів «if – then». Семантичною умовою переходу по дереву рішень є вибір коефіцієнту Gini $K_{Gini}(D)$ або ентропії по Шенонну – ймовірності досягнення кожного з можливих результатів приналежності до певного розбиття (рис. 8).



Рисунок 8 – Приклад формування дерева рішень для дослідження гідроакустичної області

Отже, використання інтелектуального методу визначення наступних подій сценарію Анд дозволяє суттєво зменшити кількість помилкових дій користувачів (особливо нових користувачів) при формуванні складних сценаріїв з множиною різноманітних умов використання операторів аналізу даних. На базі методу розроблений інтелектуальний програмний засіб, що навчається на діях користувача та на основі отриманих в процесі навчання знань представляє дії користувача у вигляді дерева рішень ймовірного сценарію аналітичного процесу, пропонуючи

користувачу найбільш доцільні наступні кроки вирішення завдання АнД. При функціонуванні цього засобу також задіяний процес автоматизованої побудови сценаріїв АнД на основі бібліотеки сценаріїв (прецедентів) для різних предметних областей, що дало можливість в автоматизованому режимі будувати сценарії на основі експертного опису проблемних ситуацій, здійснювати деталізацію сценаріїв та розгалуження їх, та, у свою чергу, суттєво зменшити часові витрати на моделювання та реалізацію сценаріїв АнД.

У четвертому розділі показано як формалізовані знання, що накопичуються на рівнях концептуального та об'єктного моделювання сценаріїв АнД, використовуються на рівні розробки комп'ютерної моделі виконання сценарію. Як правило, АнД здійснюється в Web-середовищі, характерною ознакою якого є наявність множини Web-сервісів, включаючи хмарні сервіси, значна частина яких має вільне розповсюдження і пропонує вирішення певного кола стандартних функцій, таких як: пошук, витяг даних та знань, верифікація, візуалізація, різні методи обробки даних та може бути задіяна.

Якщо розглядати АнД у Web-середовищі як послідовність виконання типових в сенсі використання стандартних функцій обробки даних завдань, то виникає завдання використання множини вже існуючих Web-сервісів та рішень для прискорення розробки програмного забезпечення системи моделювання сценаріїв. В цьому контексті сценарій АнД можна розглядати як формування послідовності виконання певних функцій обробки даних із залученням вже існуючих у Web-середовищі сервісів і включенням їх до аналітичного програмного додатку. Безумовно, Web-сервіси, які залучаються мають корелюватися з функціональним призначенням сценарію, яка описується його семантикою.

Для вирішення цього завдання запропоновано метод виконання сценаріїв АнД на основі різноманітних Web-сервісів, якій складається з наступних кроків:

1. Декомпозиція загального сценарію вирішення завдання аналітика Sc на послідовність елементарних завдань (функцій, процедур, операцій) - $UsTask_i$ та оцінка можливості залучення Web-сервісу $WebS_i$ для виконання бажаної функції або процедури.

2. Відбір потрібних Web-сервісів із наявних служб для включення їх до загального сценарію АнД.

3. Планування взаємодії (включаючи умови виклику сервісу) відібраних Web-сервісів для реалізації загального сценарію АнД.

4. Безпосереднє виконання послідовності елементарних завдань із залученням відібраних Web-сервісів з метою реалізації загального сценарію.

Після проведення декомпозиції загальний сценарій складається з послідовності виконання функціональних завдань (наприклад, обробки даних, обчислювальних та аналітичних функцій, процедур тощо)

$$Sc = \{ UsTask_1, UsTask_2, \dots, UsTask_i \}$$

Кожне функціональне завдання характеризується певним набором ознак (параметрів), що описують особливості їх виконання (джерела, формати даних, обмеження тощо). Кожен Web-сервіс теж описується деяким набором параметрів

(наприклад, назва Web-сервісу, виконувана функція, формати даних, взаємозв'язки з іншими Web –сервісами тощо). Такі набори параметрів формують свій метаопис – MD_{SD} (метадані), як для опису функціональних завдань, так і Web-сервісів, які залучаються для їх виконання. З урахуванням особливостей ПрО. Специфікація метаописів MD_{SD} як функціональних завдань, так і Web-сервісів заснована на стандарті Дублінського ядра і визначає метадані у вигляді

$$MD_{SD} = \langle At^{Func}, D^{Name} \rangle ,$$

де At^{Func} – являє атрибут опису функції, який визначає завдання або Web-сервіс, що може бути залучений для її виконання; D^{Name} – дані, що визначають ім'я та інші характеристики, необхідні для визначення взаємодії між функціональними процедурами та Web-сервісами, що їх реалізують.

Таким чином, кожне елементарне завдання $UsTask_i$ може бути описане певним набором метаданих з конкретними значеннями At^{Func} та D^{Name} саме для цієї задачі, включаючи й Web-сервіси, що залучаються.

Формування метаданих для опису компонентів загального сценарію АнД може бути представлена наступним математичним виразом

$$Sc_{MD} \ni (UsTask_1, \dots, UsTask_i) \equiv \langle Task_i, MD_{SD_i} \rangle$$

де Sc – загальний сценарій вирішення завдання аналітика – визначає набір функцій та процедур, які потрібні для вирішення задачі; $UsTask_j$ – і-те елементарне завдання у складі загального сценарію; $Task_i$ – назва і-ого елементарного завдання у складі загального сценарію; MD_{SD_i} – метаопис і-ого елементарного завдання, потрібний для пошуку відповідних Web-сервісів у Web-середовищі.

Кореляція між завданням користувача і сценарієм (що відповідає завданню цього користувача) описуються відношення еквівалентності:

$$UsTask_k \sim Sc_k$$

Набір Web-сервісів описується як:

$$WebS_{Sc_k} \ni WebS_1, WebS_i = \langle UsTask_i, MD_{SD_i} \rangle,$$

де $WebS$ – множина Web-сервісів, які можуть залучатися для вирішення задач користувача; $WebS_i$ – і-ий Web-сервіс; $UsTask_i$ – назва елементарного завдання, яке визиває і – ий Web-сервіс; MD_{SD_i} – метаопис і – того Web-сервісу.

Кореляція між сценарієм і Web-сервісами описується відношенням включення ($WebS_1$ включені в сценарій):

$$Sc_k \supset (WebS_1, \dots, WebS_d)$$

Взаємодія компонентів загального сценарію АнД вирішення завдання користувача та Web-сервісів, які залучаються для виконання сценарію, показано на рис. 9. Рисунок також відображає кореляцію взаємовідносин між компонентами сценарію та множиною Web-сервісів.

Головну увагу приділено процесу формування динамічного сценарію використання Web-сервісів за допомогою методу впорядкованого дерева, що спростить процедуру його модифікації. Враховуючи, що впорядковане дерево – це дерево з коренем, в якому визначено порядок проходження дочірніх вузлів, застосування методу впорядкованого дерева дозволяє визначити послідовність виконання сценаріїв.

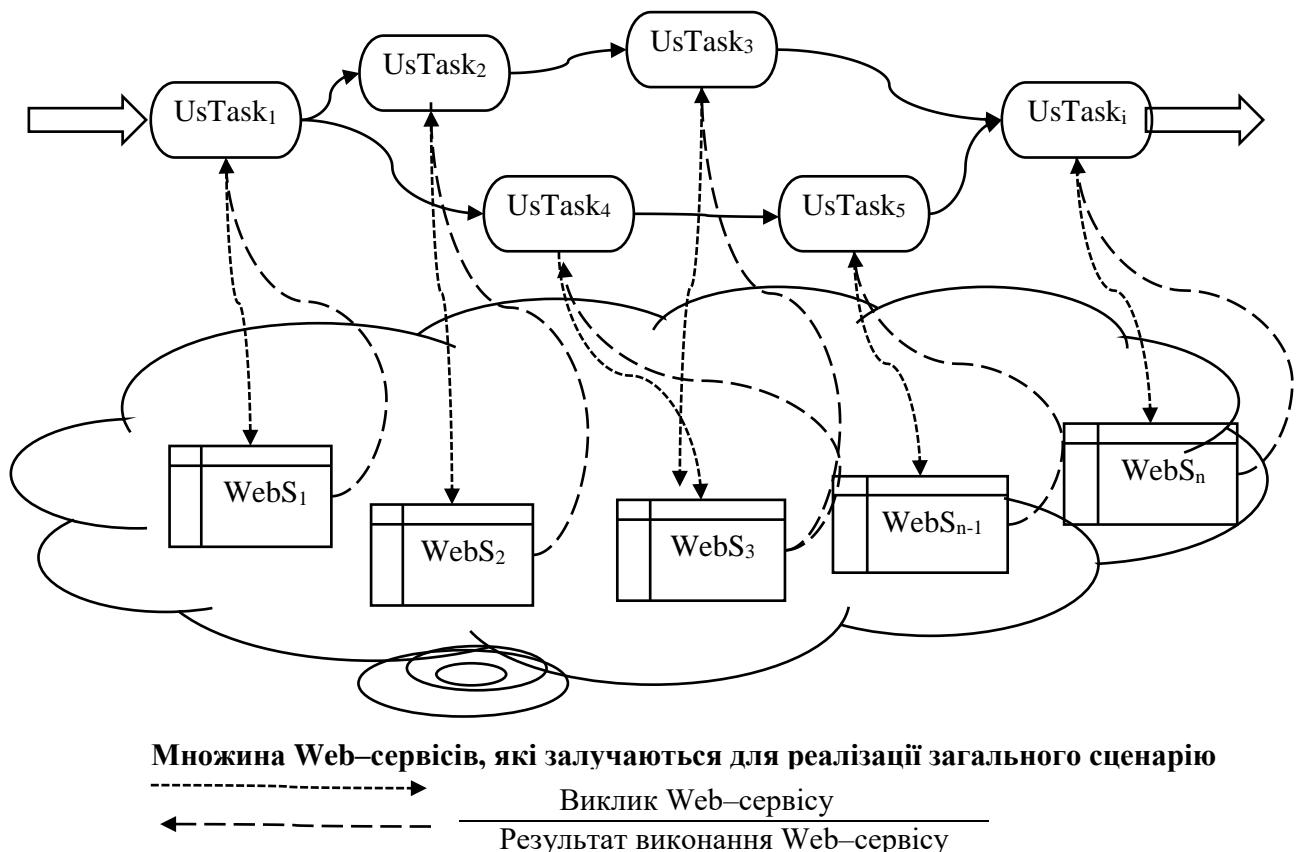


Рисунок 9 – Схема взаємодії компонентів загального сценарію вирішення завдання користувача та Web-сервісів

Отже, метод реалізації сценаріїв АнД на основі Web-сервісів, вибір яких засновано на механізмі порівняння параметрів метаописів Web-сервісів із параметрами функціональних завдань сценарію, які задаються його метаописом, що забезпечує можливість виконання завдань кінцевого користувача шляхом динамічного формування послідовності Web-сервісів, що дозволяє мінімізувати час на динамічне виконання сценаріїв для АнД.

У п'ятому розділі наведено архітектуру та функціональні можливості розробленого прототипу комп'ютерної моделюючої системи (інтегрованого програмного інструментального середовища) для тестування і оцінки запропонованого теоретичного підходу комп'ютерного моделювання сценаріїв АнД, яка ґрунтується на запропонованих та побудованих моделях, методах та алгоритмах моделювання.

За результатами досліджень запропоновано створення системи у вигляді Web-порталу для моделювання сценаріїв АнД, якій інтегрує різноманітні інструменти моделювання АнД складається з наступних програмних платформ та мов високого рівня (рис. 10):

1) Засоби формування та адміністрування Web-portal комп'ютерного моделювання сценаріїв АнД.

2) Графічний редактор для моделювання сценаріїв інформаційної взаємодії верхнього рівня – BizAgi Process Modeler Version 3.6., який підтримує стандарт ISO/IEC 19510. Окрім візуального відображення, сценарії може бути експортовано в БД у форматі XML або XPDL та конвертовано у OWL-модель онтології для подальшого аналізу.

3) RapidMiner Studio – студія та графічний редактор для моделювання сценаріїв та процедур обробки даних аналітики (структурованих, слабо структурованих, неструктурованих, DataMining, TextMining та WebMinin) Результатом дослідження є виконувана модель сценарію в форматі XML-коду.

4) Редактор та фреймворк для моделювання онтологій – Protégé 5 - Free, open-source ontology editor and framework for building intelligent systems.

5) Машина логічного виводу на прецедентах для сфер діяльності, заснованих на накопиченому досвіді – jCOLIBRI 1.0 in a nutshell. A software tool for designing CBR systems.

6) Мова програмування для опису та виводу знань в Web-просторі – OWL2 Web Ontology Language Document Overview.

7) Мова високого рівня Python 3.6.4. для формування інтерфейсів та роботи з бібліотеками NumPy.

8) Фреймворк та система управління середовищем Anaconda для широкомасштабній обробки даних, аналітики, застосування методів машинного навчання. різних програмних продуктів. Anaconda підтримує мови програмування Python та R.

Основними акторами у системі моделювання є:

- експерт предметної галузі він створює або модифікує сценарій онтологічну модель, валідує її, здійснює нагляд над її використанням;
- агент моделювання виконує моделі у середовищі моделювання.

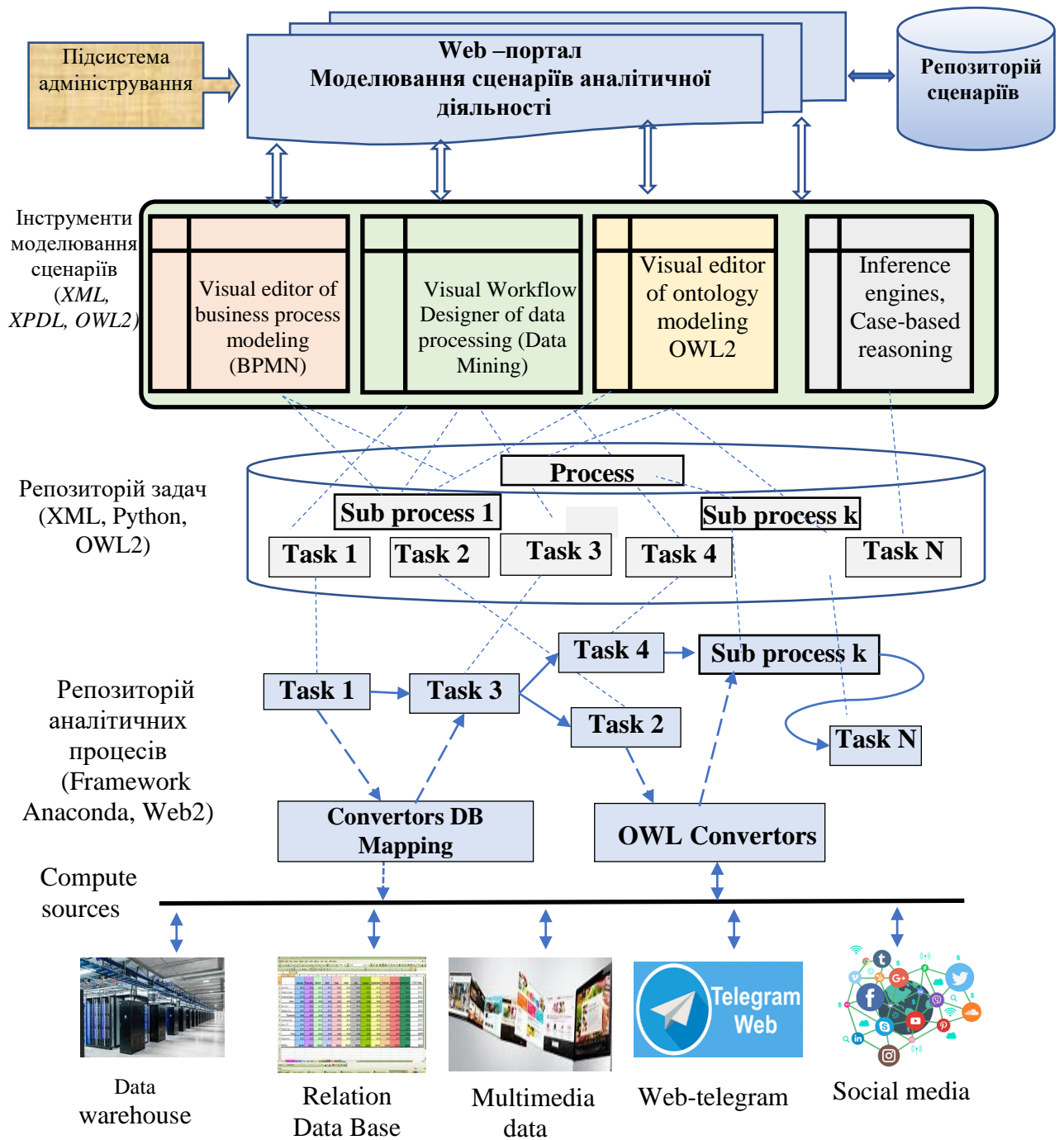


Рисунок 10 – Структурна схема комп'ютерної моделюючої системи

Зведену інформацію про акторів, сценарії та відповідні інструментальні програмні засоби або сервіси системи моделювання відображено у табл. 1.

Моделі сценарію АнД створюються та валідуються експертом у даній предметній галузі і відображають знання цієї людини щодо способу вирішення певної задачі. У розділі розроблено xml-орієнтовані мовні засоби для подання та опрацювання моделей, визначено протокольні засоби взаємодії з сервісами середовища виконання моделей.

Основні функції (варіанти використання) цих інструментальних засобів такі:

- створення та оптимізація моделі сценарію з використанням класів та фактів онтології; визначення додаткових обмежень;

Таблиця 1 – Сценарії використання комплексу моделювання

Актор	Сценарій використання	Програмний засіб (сервіс)
Експерт предметної галузі (Data scientist)	Створення моделі сценарію. Під час створення моделі відбувається її валідація та тестування, задаються правила для автоматизованої верифікації.	BizAgi Process Modeler, RapidMiner Studio, Редактор онтологій та моделей
	Оцінка результатів експлуатації моделей. Модифікація моделей.	
	Тестування моделі	
	Опрацювання онтології	
Агент моделювання	Виконання моделей	Інтерпретатор моделей
Середовище моделювання	Організація взаємодії моделей.	Брокер взаємодії моделей
	Пошук інформації	Провайдер інформаційних послуг

– робота з онтологією: пошук класів онтології та фактів; модифікація класів та відношень; визначення залежностей між певними класами онтології та існуючими моделями; визначення впливу змін в онтології на існуючі моделі;

– робота з базою фактів: пошук потрібних фактів, додавання або вилучення фактів з бази;

– визначення або налаштування існуючого інтерпретатора моделей – компонента, що виконує модель;

– тестування моделей та мереж (агрегатів) моделей: запуск та виконання моделей на тестовій базі фактів, перевірка відповідності поведінки моделей сценаріїв АнД очікуваним результатам.

Функціональні можливості прототипу середовища моделювання сценаріїв АнД проілюстровано на прикладах вирішення ряду практичних задач, зокрема при вирішенні завдань моделювання процесів функціонування вимірювальної гідроакустичної системи, комп'ютерного моделювання сценаріїв розподілу енергопостачання та побудови і реалізації сценаріїв АнД в задачах просторового моделювання.

У шостому розділі проаналізовано фактори, що впливають на ефективність застосування запропонованих моделей та методів моделювання сценаріїв аналітичної діяльності для побудови аналітичних програмних систем, та розроблено оціночні формули для визначення ступеня покращення окремих складових характеристик якості програмних систем.

По-перше, використання онтологічних моделей впливає на якісні аспекти процесу створення та експлуатації програмних систем, які їх використовують. Задача аналізу переваг використання онтологічних моделей вирішується як задача оцінки ступеня підвищення якості ПС в результаті їх використання на всіх етапах її життєвого циклу.

Запропоновані вимоги, яким повинна відповідати онтологічна модель, що забезпечить логічність висновків, однакову відповідь на однакові, але по-різному сформульовані пошукові запити, насамперед вимоги щодо цілісності онтологічної моделі. Визначені умови, при яких оцінки онтологічної моделі свідчать про її відповідність запропонованим вимогам.

Досліджені наявні метрики оцінки онтологічної моделі з урахуванням методологічних підходів щодо оцінки якості програмного продукту Стандарту ISO/IEC 9126 (Software engineering — Product quality), включаючи й якість моделі (Quality model) та визначені критерії оцінки якості моделі онтології. В разі, якщо онтологічна модель не відповідає критеріям оцінки її якості, здійснюють її структурну перевірку та оптимізацію змістової частини послідовною редукцією її графу до виконання вимог вибраних критеріїв, причому розв'язування цих задач розноситься у часі, щоб зберегти цілісність онтології.

По друге, ефективність застосування запропонованих моделей та методів моделювання сценаріїв АнД визначається оцінкою ефективності самого процесу моделювання сценарію. Визначення такої оцінки з використанням онтологічної моделі для багатофакторного аналізу на відміну від однофакторного потребує побудови алгоритмів, що зможуть надати можливість однозначного вибору сценарію на основі ієрархічної за структурою моделі онтології ПрО, на основі якої будуються сценарії.

Найбільш відповідним до поставленої задачі є структурний підхід до розгляду онтології, оскільки представлення структури онтології у вигляді багаторівневого ієрархічного графа дає можливість виміряти її властивості за допомогою метрики, по якій можна буде визначити її якість та сформулювати рекомендації щодо її поліпшення. Важливим аспектом структурного підходу є необхідність оцінки ефективності побудови онтологічної моделі на основі оцінки скорочення витрат пошуку по графу, що описує онтологію, у порівнянні з неспрямованим пошуком шляху. Для таких графів, як правило, використовується оцінки затрат, як на обробку пов'язаною з діями інформації в окремих вузлах графу, так і оцінки затрат на самі дії (по ребрах графу) при переході з одного вузла до іншого. Таким чином можна сумарно оцінити затрати на кожний можливий шлях відповідно до сценарію чи окремих фрагментів цього шляху для порівняння та визначення найбільш ефективних шляхів, що включаються у опис обраного сценарію.

Послідовність кроків алгоритму оцінки ефективності моделювання сценарію АнД наступна:

Крок 1. Вибір на нижньому шарі вершин графа відповідної оцінки ефективності пошуку оптимального шляху на графі за відповідними критеріями.

Крок 2. Аналіз зв'язків виділених вершин з вершинами наступного рівня, який розташований вище по ієрархії моделі.

Крок 3. Збереження ребр графа, які відповідають знайденим зв'язкам, для подальшого аналізу і використання в пошукових операціях.

Крок 4. Повторення ітерацій до виходу на джерело пошуку на верхньому рівні.

Крок 5. Консолідація усіх зібраних ребр і вершин графа, який побудований на підставі відповідної оцінки ефективності, тобто запиту на пошук інформації.

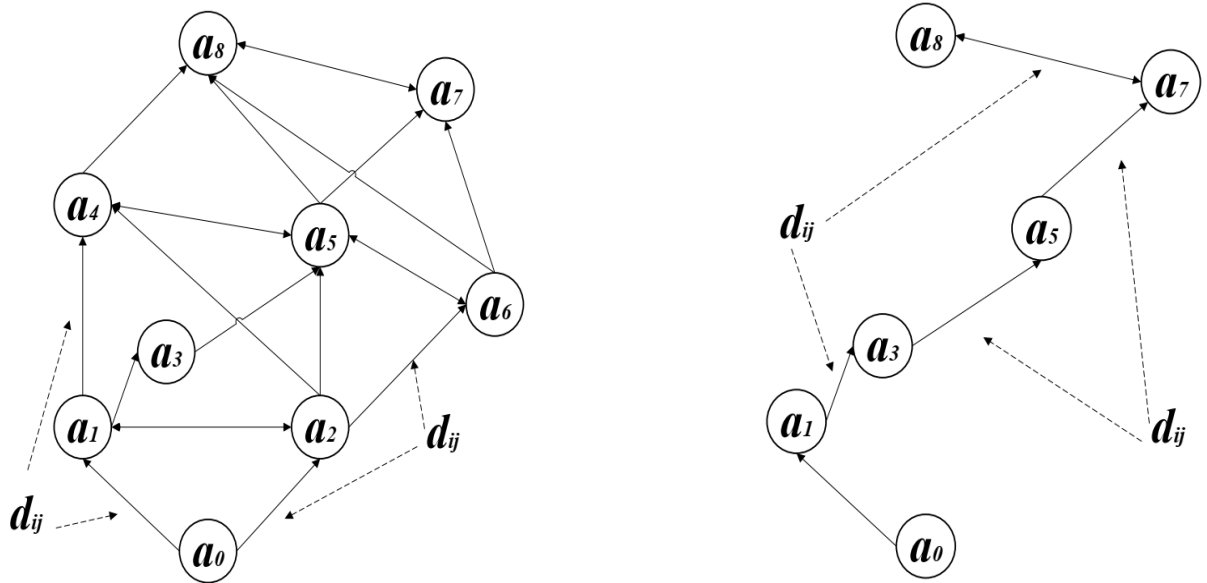
Крок 6. Упорядкування всіх шляхів по постійному графу відповідно до функції оцінки.

Крок 7. Аналіз графічної моделі сценарію, яка побудована на основі вибору найкоротших з мінімальною кількістю ребр на графі (рис. 11б).

Оцінка ефективності процесу моделювання сценарію аналітичної діяльності як ефективності пошуку оптимального шляху на графі може бути визначена як сукупна оцінка затрат на дії по кожному з вузлів та оцінки затрат на перехід до наступного вузлу у процесі переходу від начального вузлу до кінцевого з визначенням можливих шляхів переходу:

$$P_a = \sum_k (t_k + r_k),$$

де t_k – затрати часу на кожному з k вузлів, що відповідають обраному шляху на графі; r_k – затрати ресурсів на кожному з k вузлів, що відповідають обраному шляху на графі; $k \in K$ – номери вузлів обраного шляху на графі.



а) Можливі шляхи

б) Обраний оптимальний шлях

Рисунок 11 – Приклад моделі для побудови оптимального шляху від початково a_0 до кінцевого a_8 вузлу по можливим шляхам (дугам) d_{ij}

Відповідно можна визначити критерій оцінки як суму затрат на обробку в ребрах переходу між вузлами як

$$P_d = \sum_{i,j} (T_{ij} + R_{ij}),$$

де T_{ij} – затрати часу на кожному з ij -тому ребрі графу, що відповідають обраному шляху на графі; R_{ij} – затрати ресурсів на кожному з ij -тому ребрі графу, що відповідає обраному шляху на графі; $i, j \in K$ – номери вузлів обраного шляху на графі.

Тоді загальний критерій може бути визначений як сума P_a та P_d .
Сумарний критерій

$$P = P_a + P_d.$$

чи

$$P = \sum_k (t_k + r_k) + \sum_{i,j} (T_{ij} + R_{ij}), \quad k, i, j \in K,$$

K – сукупність індексів, що відповідають обраному шляху.

При цьому як критерій ефективності так і оцінка отриманих результатів в конкретному вузлу, що визначається відповідно до конкретної задачі наявністю необхідних результатів за оцінками аналітиків-експертів:

$$P = \sum_k (c_t t_k + c_r r_k) + \sum_{i,j} (c_T T_{ij} + c_R R_{ij}), \quad k, i, j \in K,$$

де c_t – коефіцієнт впливу затрат часу кожному з k вузлів, c_r – коефіцієнт впливу затрат ресурсів на кожному з k вузлів, c_T – коефіцієнт впливу затрат часу на кожному з ij -тому ребрі графу, c_R – коефіцієнт впливу затрат ресурсів на кожному з ij -тому ребрі графу. Крім того $c_t, c_r, c_T, c_R \geq 0$.

В залежності від відзначень цих коефіцієнтів аналітиком визначається той чи інший тип задачі пошуку оптимального шляху: за критерієм часу, за мінімальністю затрат ресурсів чи з використанням комбінованого критерію. Обрання загального сумарного напрямку шляху може визначатися наявністю найбільш ефективного напрямку з точки зору значення критерій V на кожному послідовному кроці вибору напрямку на графі.

Тригер оцінки результатів на кожному кроці визначає: чи є покращення результатів, чи ні.

$Q_i > 0$ – вузол розглядається, як можливий для порівняння та вибору.

$Q_i = 0$ – вузол a_i не розглядається, як можливий для порівняння та вибору.

Цю оцінку має виконувати аналітик на основі отриманих результатів.

Для кожного рівня у порівнянні розглядаються ті напрямки(ребра) та вузли, що відповідають $Q_i > 0$. Інші напрямки (ребра) та вузли з подальшого розгляду вилучаються.

Значення $Q_i > 0$ може визначатися за певними обраними мірами аналітиком, як оцінка збільшення результатів рішення аналітичної задачі, що вирішується.

При відсутності наявних та обґрунтованих методів оцінки $Q_i = 1$ – вузол a_i розглядається, як такий, що покращує рішення задач аналітиком. $Q_i = 0$ – вузол a_i не розглядається, як можливий для порівняння та вибору, як такий, що не надає нових знань чи нових результатів для аналітика.

Без зменшення ступеню загальності, модель, де розглядаються затрати пов'язані як з сукупною оцінкою затрат на роботи по кожному з вузлів, так і оцінки затрат на перехід до наступного вузлу у процесі переходу від начального вузлу до кінцевого, може розглядатися, як модель з урахування тільки затрат на ребрах графу.

При цьому затрати в вузлах будуть включені до затрат ребр графу

$$P_{ij} = t_j + r_j + T_{ij} + R_{ij}, \quad i, j \in K.$$

Були проведені імітаційні експерименти моделювання сценаріїв за використанням запропонованих методів та розроблених програмних засобів на двохстах тестових прикладах вирішення задач аналітики з усередненням результатів тестових прикладах для різних ПрО (онтологічних моделей). Оцінка затрат визначена як оцінка скорочення часу виконання аналітичних процедур. Результати тестування показують, що використання запропонованих методів та програмних засобів, які створені на їх базі, дає можливість значно скоротити час здійснення АНД.

Таким чином, проаналізовано фактори, що впливають на ефективність застосування запропонованих моделей та методів моделювання сценаріїв аналітичної діяльності для побудови аналітичних програмних систем, та розроблено оціночні формули для визначення ступеня покращення окремих складових характеристик якості програмних систем.

ВИСНОВКИ

У дисертаційній роботі вирішено важливу науково-прикладну проблему дослідження, формалізації та розробки теоретичних основ та практичних засад комп'ютерного моделювання сценаріїв аналітичної діяльності в різних предметних областях, зокрема при проектуванні складних технічних систем. В тому числі отримані наступні результати.

1. Проаналізовано проблему побудови та виконання складних сценаріїв АнД (формальні методи побудови сценаріїв, включаючи графові моделі, визначення їх переваг та недоліків). Обґрунтовано актуальність вирішення проблеми підвищення ефективності моделювання сценаріїв розвитку процесів та події завдяки інтеграції онтологічних та сценарних моделей, та розробки та впровадження інтелектуальних методів та програмних засобів для зменшення часу здійснення АнД.

2. На підставі результатів аналізу вперше запропоновано теоретичний підхід (методологія) до комп'ютерного моделювання багаторівневих сценаріїв АнД на основі поєднання переваг візуального моделювання – графічної нотації BPMN с керованою моделлю і семантичного аналізу на базі онтологічної моделі предметної області задач АнД, яка конвертується з BPMN-моделі, та подальшої серіалізації BPMN-моделі в виконавче програмне середовище, що зменшує кількість логічних розривів в комп'ютерній моделі побудови сценарію АнД, яка будується, та суттєво зменшує час моделювання в цілому.

3. Запропоновано метод перетворення BPMN-моделі в OWL-модель сценарію АнД за рахунок забезпечення співвідношення категорії основних графічних примітивів нотації BPMN до відповідних класів онтології (сутностей) та зв'язування їх з властивостями та відношеннями між елементами реальної моделі вирішення задачі АнД в нотації BPMN, що дозволяє проводити семантичний аналіз коректності логічно-зв'язаних фрагментів сценарію вирішення задачі АнД та його верифікацію (особливо багаторівневих сценаріїв).

4. Запропоновано інформаційну технологію конвертації BPMN-моделі в OWL-модель сценарію АнД, яка дозволяє перетворювати метадані опису сценарію та його предметної області в базі знань, накопичуючи нові знання, що дає можливість побудови нових сценаріїв АнД на базі вже існуючих в базі знань. Для спрощення процесу комп'ютерного моделювання сценаріїв залучається графічна нотація BPMN 2, яка дозволяє не тільки моделювати сценарії АнД у вигляді BPMN-діаграм, але й серіалізувати його опис в XML-файл для реалізації сценарію в комп'ютерному середовищі із залученням стандартного програмного забезпечення BPM-систем.

5. Удосконалено математичну модель обчислення розподілу значень можливих подій сценарію та інтелектуальний метод визначення наступних подій

сценарію АнД на основі методів машинного навчання, якій дозволяє виконувати формування дерева рішень для класифікації можливих наступних кроків сценарію. Це дає можливість значно підвищити точність розбиття складного сценарію при обчисленні ймовірності наступного його кроку.

6. Удосконалено метод реалізації сценаріїв АнД на основі Web-сервісів, вибір яких засновано на механізмі порівняння параметрів метаописів Web-сервісів із параметрами функціональних завдань сценарію, які задаються його метаописом, що забезпечує можливість виконання завдань кінцевого користувача шляхом динамічного формування послідовності Web-сервісів.

7. Набув подальшого розвитку процес автоматизованої побудови сценаріїв АнД на основі бібліотеки сценаріїв (прецедентів) для різних предметних областей. Це дало можливість в автоматизованому режимі будувати сценарії на основі експертного опису проблемних ситуацій, здійснювати деталізацію сценаріїв та розгалуження їх. У свою чергу, це суттєво зменшує часові та фінансові витрати на моделювання та реалізацію сценаріїв АнД.

8. Запропоновано та створено інтелектуальний програмний засіб, що навчається на діях користувача та на основі отриманих в процесі навчання знань представляє дії користувача у вигляді дерева рішень ймовірного сценарію аналітичного процесу, пропонуючи користувачу найбільш доцільні наступні кроки вирішення завдання АнД.

9. Запропонована та створена трирівнева комп'ютерна моделююча система (інтегроване програмне інструментальне середовище) для тестування і оцінки запропонованого теоретичного підходу (методології) моделювання сценаріїв АнД, яка ґрунтується на запропонованих моделях, методах та алгоритмах моделювання, що дало можливість реалізувати окремі програмні інтелектуальні засоби вирішення задач аналітики для різних сфер, ядром баз знань яких є знання та накопичений досвід аналітика (експерта, проектувальника), зокрема для вирішення задачі моделювання функціонування вимірювальної гідроакустичної системи з метою удосконалення алгоритмів обробки гідроакустичних сигналів та інтелектуальної системи розподілу енергопостачання з метою оптимізації енерговитрат.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці, в яких опубліковані основні результати дисертації:

1. Додонов А.Г., Сенченко В.Р., Коваль А.В. Аналітика и знания в компьютерных системах. Киев: ИПРИ НАН Украины, «КПИ имени Игоря Сикорского», 2020. 315 с.

2. Додонов О.Г., Коваль О.В., Глоба Л.С., Бойко Ю.Д. Комп'ютерне моделювання інформаційно-аналітичних систем, Київ: ІПРІ НАН України, 2017. 239 с.

3. Kuzminykh V., Koval O., Melnyk U., Otroh S. Evaluating the Quality of Modeling the Scenario of Information Analysis on a Branched Network. *Сучасний захист інформації*. 2019. № 3(39). С. 70-76.

4. Додонов О.Г., Сенченко В.Р., Коваль О.В., Бойченко А.В. Моделювання сценаріїв аналітичної діяльності на основі нотації BPMN OWL. *Реєстрація, зберігання і обробка даних*. 2020. Т. 22, № 1. С. 31-48.
5. Koval O.V., Kuzminykh V.O., Svistunov S.Y., Xu Beibei, Zhu Shiwei. Data collection for analytical activities using adaptive microservice architecture. *Реєстрація, зберігання і обробка даних*. 2021. Т. 23, № 1. С. 11-31.
6. Додонов О.Г., Коваль О.В., Сенченко В.Р., Швайко В.Г. Формування та реалізація сценаріїв аналітики в задачах просторового моделювання. *Реєстрація, зберігання і обробка даних*. 2020. Т. 22, № 3. С. 39-57.
7. Kuzminykh V., Koval O., Otrokh S. Refining the typical scenarios by additional factors. *Математичне та комп'ютерне моделювання. Серія: Технічні науки*. 2019. Вип. 20. С. 68-78.
8. Додонов О.Г., Коваль О.В., Сенченко В.Р., Шпурик В.В. Автоматизована система формування сценарію аналітичної діяльності. *Реєстрація, зберігання і обробка даних*. 2019. Т. 21, № 1. С. 11-22.
9. Novogrudska R., Globa L., Koval O. The Simplification Method of Engineering Task Sequences used for Engineering Knowledge Portals. *Вісник Харківського Національного університету. Серія: Математичне моделювання. Інформаційні технології. Автоматизовані системи управління*. 2018. Вип. 37. С. 37-44.
10. Koval O.V., Kuzminykh V.A., Khaustov D.V. Using stochastic automaton for data consolidation. *Наукові вісті НТУУ «КПІ»*. 2017. № 2. С. 29-36.
11. Коваль А.В., Бойко Ю. Д., Волкова Е. А. Особенности сценарно-целевого подхода к анализу объектов действенной аналитики. *Системні дослідження та інформаційні технології*. 2015. № 1. С. 57-67.
12. Коваль О.В., Кузьмініх В.О. Реалізація сценарного підходу в управлінні проектами на основі типових задач. *Реєстрація, зберігання і обробка даних*. 2015. Т. 17, № 1. С. 77-87.
13. Коваль О.В., Зайцева А. Верифікація комп'ютерної моделі системи інформаційного управління. *Вісник НТУУ «КПІ»*. Серія: Інформатика, управління та обчислювальна техніка. 2014. № 61. С. 43-48.
14. Koval O.V., Zaytseva K.A., Boyko Yu.D. Formation of Analytical activity Scenarios. *Системні дослідження та інформаційні технології*. 2014. № 1. С. 20–25.
15. Діденко О.О., Сенченко В.Р., Коваль О.В. Система моніторингу державного бюджету України. *Наукові вісті НТУУ «КПІ»*. 2013. № 6. С. 37-48.
16. Koval A.V., Kuzminykh V.A., Voronko M.P., Khaustov D.V. Development of a Scenario-Based Project Management System Construction in Enterprises with the Functional Organizational Structure. *Informatyka Automatyka Pomiaru w Gospodarce i Ochronie Środowiska*. 2013. No. 4. P. 26-30.
17. Кузьмініх В.О., Коваль О.В., Воронько М.П. Оцінка часу виконання типових задач проектів на підприємствах з функціональною організаційною структурою, *Реєстрація, зберігання і обробка даних*. 2012. Т. 14, № 3, С. 70-80.

18. Кузьмініх В.О., Коваль О.В., Хаустов Д.В., Коростельова Є.Ю. Управління агрегованими групами проектів, *Реєстрація, зберігання і обробка даних*. 2011. Т. 13, № 3. С. 109-116.

19. Коваль О.В. Узагальнена архітектура аналітичної складової корпоративних інформаційно-аналітичних систем. *Реєстрація, зберігання і обробка даних*. 2011. Т. 13, № 2. С. 53-73.

20. Додонов О.Г., Коваль О.В., Дзюбаненко Р.І., Цепков П.А., Маюров М.О., Жидовленко Ю.О. Концептуальні рішення створення автоматизованої системи екстреної допомоги населенню за єдиним телефонним номером 112. *Реєстрація, зберігання і обробка даних*. 2010. Т. 12, № 2. С. 165-180.

21. Додонов О.Г., Коваль О.В., Сенченко В.Р. Методологія побудови корпоративних інформаційно-аналітичних систем. *Реєстрація, зберігання і обробка даних*. 2007. Т. 9, № 4. С. 60-75.

22. Kuzminykh V., Koval O., Xu B., Zhu Sh. Microservice architecture of the system for assessing the level of international activity. *International scientific journal «Internauka»*. 2021. No. 6. С. 67-79.

23. Globa L., Novogradska R., Koval O. The Approach to Users Tasks Simplification on Engineering Knowledge Portals. *Advances in Intelligent Systems and Computing*. 2018. Vol. 889. P. 150-158. (Scopus).

24. Globa Larysa, Novogradska Rina, Koval Alexander, and Senchenko Vyacheslav. Ontology for Applications Development. *IntechOpen*. 2018. Ch. 2. P. 29–53.

25. Globa Larysa, Novogradska Rina, Koval Alexander and Senchenko Vyacheslav. Examples of Ontology Model Usage in Engineering Fields. *IntechOpen*. 2018. Ch. 3. P. 55-81.

26. Koval O., Globa L., Novogradska R. The approach to web services composition. *Advances in Intelligent Systems and Computing*. 2017. Vol. 534. P. 293-304. (Scopus).

Праці апробаційного характеру:

27. O. Koval, V. Kuzminykh, S. Otrokh and V. Kravchenko. “Optimization of Scenarios for Collecting Information Streaming Wide-Area Network”, 3rd International Conference on Advanced Information and Communications Technologies (AICT), 2-6 July, 2019, Lviv. 2019, pp. 213-215. (Scopus)

28. Oleksandr Koval, Valeriy Kuzminykh and Maksym Voronko. “Standard Analytic Activity Scenarios Optimization based on Subject Area Analysis”, in Selected Papers of the XIX International Scientific and Practical Conference "Information Technologies and Security" (ITS 2019) CEUR Workshop Proceedings, vol. 2577, pp. 37-46, 2019. (Scopus)

29. Aleksandr Koval, Safwan Al Salameh and Oleh Andriichuk. “Usage of Expert Classification in Diagnostic Expert Systems’ Knowledge Bases Construction”, in Selected Papers of the XVIII International Scientific and Practical Conference "Information Technologies and Security" (ITS 2018), CEUR Workshop Proceedings, vol. 2318, pp. 139-149, 2018. (Scopus)

30. V. O. Kuzminykh, A. V. Koval and M. V. Osypenko. “Methods of Machine Training on the Basis of Stochastic Automatic Devices in the Tasks of Consolidation of Data from Unsealed Sources”, in Selected Papers of the XVII International Scientific and Practical Conference on Information Technologies and Security (ITS 2017) CEUR Workshop Proceedings, vol. 2067, pp. 63-68, 2017. (Scopus)

31. L. S. Globa, R. L. Novogradska, and A. V. Koval. “Ontology Model of Telecom Operator Big Data”, in 2018 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom), 4-7 June, 2018 Batumi, Georgia. IEEE, 2018, pp. 1-5.

32. R. L. Novogradska, L. S. Globa, O. V. Koval, and V. R. Senchenko. “Ontology model of intelligent modeling system for marine facilities identification”, in International Conference Radio Electronics & Info Communications (UkrMiCo’2017), 11-15 September, 2017, Kyiv. Odesa, 2017, pp. 536-541. (Scopus)

33. V. R. Senchenko, and A. V. Koval. “The technology of semantic modeling for knowledge management systems in environment Protégé”. Матеріали XVII Міжнародної науково-практичної конференції “Інформаційні технології і безпека (ITS-2017)”, 30 листопада, 2017, Київ. Київ: ООО «Інжиніринг», 2017, вип.17. С. 211-235.

34. Khazanovych Y., Gerayimchuk M., Koval A., Senchenko V. «Ontological Design of an Intelligent System for Studying Hydroacoustical Processes», IEEE International Conference on Power, Control, Signals & Instrumentation Engineering (ICPCSI-2017), Saveetha Engineering College, Chennai, India, 2017, pp. 135-138.

35. V. R. Senchenko, O. V. Koval, L. S. Globa, and R. L. Novogradska. «Intelligent modeling system based on cloud-technology», International Conference Radio Electronics & Info Communications UkrMiCo 2016. IEEE, 2016, pp. 449-452. (Scopus)

36. Глоба Л.С., Коваль О.В., Новогрудська Р.Л., Сенченко В.Р. Створення сценаріїв обробки даних на основі онтології. Матеріали 18-ї Міжнародної науково-технічної конференції SAIT–2016 «Системний аналіз та інформаційні технології», 30 травня, 2016 р., Київ. Київ: ІПСА КПІ, 2016. С. 262-265.

37. O.V. Koval, and K.A. Volkova. Computer modelling of management information system with actionable analytics. Матеріали 17-ї міжнародної науково-технічної конференції SAIT–2015 «Системний аналіз та інформаційні технології», 25 червня, 2015 р., Київ. Київ: ІПСА КПІ, 2015. С. 32.

38. Коваль А.В., Бойко Ю.Д., Зайцева Е.А. Модель сценарно-целевого подхода при построении информационно-аналитической системы. Системный анализ и информационные технологии: сб. тезисов 16-ой междунар. науч.-техн. конф. (Київ, 26 мая 2014 р.). Київ: НК «ІПСА», 2014. С.105-106.

39. Зайцева Е.А., Коваль А.В. Комп’ютерне моделювання соціально-економічного розвитку регіону. Системный анализ и информационные технологии: сб. тезисов 15-ой междунар. науч.-техн. конф. (Київ, 31 мая 2013 р.). Київ: НК «ІПСА», НТУУ «КПІ», 2013. С. 103.

40. Коваль А.В., Сенченко В.Р. Построение системы анализа выполнения госбюджета на основе сценарного подхода. Системный анализ и информационные технологии: сб. тезисов 14-ой междунар. науч.-техн. конф. (Киев, 24 апр. 2012 р.). Киев: ИПСА НТУУ «КПИ», 2012. С. 202-203.

41. Коваль А.В., Діденко О.О. Імітаційне моделювання засобів підтримки діяльності в корпоративних інформаційно-аналітичних системах. Сучасні проблеми наукового забезпечення енергетики: зб. матеріалів ІХ міжн. наук.-практ. конф. аспірантів, магістрів, студентів. Київ. 2011. С. 239.

42. Калішук А.А., Коваль О.В. Автоматизація генерації сценаріїв аналітичної діяльності. Сучасні проблеми наукового забезпечення енергетики: збірник доповідей VII між нар. наук.-практ. конф. аспірантів, магістрів, студентів. Київ, 2009. С. 135.

43. Додонов О.Г., Коваль О.В. Урядова інформаційно-аналітична система з питань надзвичайних ситуацій. Системи підтримки прийняття рішень. Теорія і практика: збірник доповідей наук.-практ. конф. з міжнародною участю. Київ: ІПММС НАНУ, 2007. С. 29-32.

44. Коваль А.В., Сенченко В.Р. Система моніторингу фінансово-економічних та макроекономічних показників як ефективний інструмент прийняття управлінських рішень. Проблеми моніторингу та забезпечення економічної безпеки держави: збірник наукових праць наук.-практ. семінару з міжнародною участю. Київ, 2007. С. 58-60.

АНОТАЦІЯ

Коваль О.В. Методи та засоби комп'ютерного моделювання сценаріїв аналітичної діяльності. – На правах рукопису.

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 01.05.02 «Математичне моделювання та обчислювальні методи» – Інститут проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України, Київ, 2021.

В дисертаційній роботі вирішена важлива науково-прикладна проблема дослідження, формалізації та розробки теоретичних основ та практичних засад комп'ютерного моделювання сценаріїв аналітичної діяльності (АнД) в різних предметних областях, зокрема при проектуванні складних технічних систем.

Запропоновано теоретичний підхід до комп'ютерного моделювання багаторівневих сценаріїв АнД, який відрізняється від відомих тим, що оснований на поєднанні переваг візуального моделювання – графічної нотації BPMN з керованою моделлю і семантичного аналізу на базі онтологічної моделі предметної області задач АнД та подальшої серіалізації BPMN-моделі в програмне середовище, що зменшує кількість логічних розривів в комп'ютерній моделі сценарію АнД та суттєво зменшує час моделювання в цілому. В основі запропонованого підходу лежить метод перетворення BPMN-моделі в модель представлення предметної області сценарію АнД мовою OWL - OWL-модель, який відрізняється від відомих тим, що дозволяє проводити семантичний аналіз коректності логічно зв'язаних

фрагментів сценарію вирішення задачі АнД та його верифікацію (особливо багаторівневих сценаріїв). Запропонована інформаційна технологія конвертації BPMN-моделі в OWL-модель сценарію аналітичної діяльності, що дозволяє перетворювати метадані опису сценарію та його предметної області в базі знань, накопичуючи нові знання, що дає можливість побудови нових сценаріїв АнД, використовуючи вже існуючі в базі знань. Набули подальшого розвитку математична модель обчислення розподілу значень можливих подій сценарію та інтелектуальний метод визначення наступних подій сценарію АнД, який відрізняється від відомих тим, що дозволяє виконувати формування дерева рішень для класифікації можливих наступних кроків сценарію, що дає можливість значно підвищити точність розбиття графу складного сценарію на відповідні гілки при обчисленні імовірності наступного кроку сценарію. Удосконалено метод реалізації сценаріїв АнД на основі Web-сервісів, який забезпечує можливість виконання завдань кінцевого користувача шляхом динамічного формування послідовності Web-сервісів. Набули подальшого розвитку інтелектуальний програмний засіб, що навчається на діях користувача та пропонує користувачу найбільш доцільні наступні кроки вирішення завдання АнД, та процес автоматизованої побудови сценаріїв АнД на основі бібліотеки сценаріїв (прецедентів) для різних предметних областей.

Реалізація запропонованих моделей, методів та комп'ютерних засобів в рамках трирівневої комп'ютерної моделюючої системи, призначеній для тестування і оцінювання запропонованого теоретичного підходу моделювання сценаріїв АнД, забезпечила можливість вирішення задачі побудови та реалізації сценаріїв як в лабораторних дослідженнях, так і в умовах застосування в реальних моделюючих системах.

Ключові слова: аналітична діяльність, сценарій аналітичної діяльності, онтологічна модель предметної області, комп'ютерне моделювання багаторівневих сценаріїв аналітичної діяльності, інтелектуальний метод визначення наступних подій сценарію, трирівнева комп'ютерна моделююча система моделювання сценаріїв аналітичної діяльності.

АННОТАЦІЯ

Коваль А.В. Методы и средства компьютерного моделирования сценариев аналитической деятельности. – На правах рукописи.

Диссертация на соискание ученой степени доктора технических наук по специальности 01.05.02 «Математическое моделирование и вычислительные методы» – Институт проблем моделирования в энергетике им. Г.Е. Пухова НАН Украины, Киев, 2021.

В диссертационной работе решена важная научно-прикладная проблема исследования, формализации и разработки теоретических и практических основ компьютерного моделирования сценариев аналитической деятельности (АнД) в различных предметных областях, в частности при проектировании сложных технических систем.

Предложен теоретический подход к компьютерному моделированию многоуровневых сценариев АнД, который отличается от известных тем, что основан на сочетании преимуществ визуального моделирования – графической нотации BPMN с управляемой моделью и семантического анализа на базе онтологической модели предметной области задач АнД и последующей сериализации BPMN-модели в программной среде, что уменьшает количество логических разрывов в компьютерной модели сценария аналитической деятельности и существенно уменьшает время моделирования в целом. В основе предложенного подхода лежит метод преобразования BPMN-модели в модель представления предметной области сценария АнД на языке OWL - OWL-модель, который отличается от известных тем, что позволяет проводить семантический анализ корректности логически связанных фрагментов сценария решения задачи аналитической деятельности и его верификацию (особенно многоуровневых сценариев). Предложенная информационная технология конвертации BPMN-модели в OWL-модель сценария аналитической деятельности, позволяет превращать метаданные описания сценария и его предметной области в базе знаний, накапливая новые знания, дает возможность построения новых сценариев АнД, используя существующие в базе знаний. Получили дальнейшее развитие математическая модель расчета распределения значений возможных событий сценария и интеллектуальный метод определения последующих событий сценария АнД, который отличается от известных тем, что позволяет выполнять формирование дерева решений для классификации возможных следующих шагов сценария, что дает возможность значительно повысить точность разбивки графа сложного сценария на соответствующие ветви при вычислении вероятности следующего шага сценария. Усовершенствован метод реализации сценариев АнД на основе Web-сервисов, который обеспечивает возможность выполнения задач конечного пользователя путем динамического формирования последовательности Web-сервисов. Получило дальнейшее развитие интеллектуальное программное обеспечение, которое обучается на действиях пользователя и предлагает пользователю наиболее целесообразные следующие шаги решения задачи АнД, и процесс автоматизированного построения сценариев АнД на основе библиотеки сценариев (прецедентов) для различных предметных областей.

Реализация предложенных моделей, методов и компьютерных средств в рамках трехуровневой компьютерной моделирующей системы, предназначенной для тестирования и оценки предложенного теоретического подхода моделирования сценариев АнД, обеспечила возможность решения задачи построения и реализации сценариев как в лабораторных исследованиях, так и в условиях применения в реальных моделирующих системах.

Ключевые слова: аналитическая деятельность, сценарий аналитической деятельности, онтологическая модель предметной области, компьютерное моделирование многоуровневых сценариев аналитической деятельности, интеллектуальный метод определения последующих событий сценария,

трехуровневая компьютерная система моделирования сценариев аналитической деятельности.

ABSTRACT

Koval O.V. Methods and tools for computer modeling of analytical scenarios. – As the manuscript.

The thesis for obtaining a scientific degree of Doctor of Technical Sciences in specialty 01.05.02 – Mathematical modeling and computational methods, Pukhov Institute for Modelling in Energy Engineering of the NAS of Ukraine, Kyiv, 2021.

The thesis solves an important scientific and technical problem of research, formalization and development of theoretical and practical bases of computer modeling of analytical activity scenarios in various subject domains, at designing of complex technical systems.

A theoretical approach to computer modeling of multilevel analytical activity scenarios is proposed, which unlike the known ones is based on combining the advantages of visual modeling - graphic BPMN notation with controlled model and semantic analysis based on ontological model of subject domain of analytical activity problems and further serialization of the BPMN model into a software environment that reduces the number of logical gaps in the computer model of the analytical activity scenario and significantly reduces the overall simulation time. The proposed approach is based on the method of transforming the BPMN-model into a model of representation of the subject domain of the analytical activity scenario with OWL language - OWL-model, which unlike the known ones allows to conduct a semantic analysis of the accuracy of logically related fragments of the analytical activity solution and its verification (especially in multilevel scenarios). The proposed information technology for converting BPMN-model into OWL-model of analytical activity scenario, which allows to convert metadata description of the scenario and its subject domain in the knowledge base, accumulating new knowledge, which allows to build new analytical activity scenario using already existing in the knowledge base.

An approach is proposed to solve the problem of ensuring the semantic compatibility of the processes of modeling the steps of a multilevel scenario of analytical activity through the use of the ontology of the subject area to resolve conflicts. A set of axioms and theorems that theoretically substantiate the conditions for ensuring the semantic compatibility of transitions is presented and proved. An approach to resolving differences that arise in transitions is proposed. This approach is based on the ontological model of the semantic mediator, which is applied similarly to the approach in Semantic Web Service technologies.

The mathematical model for calculating the distribution of values of possible scenario events and the intelligent method for determining subsequent events of the analytical activity scenario are further developed, which unlike the known ones allows to perform the formation of a decision tree to classify the possible next steps of the scenario and significantly increases the accuracy of splitting the graph of a complex scenario into appropriate branches when calculating the probability of the next step of the scenario. The method of implementing the analytical activity scenario based on Web-services has been improved, which provides the ability to perform end-user tasks by dynamically forming a sequence of Web-services. Further developed are an intelligent software tool, which is taught by the user's actions and offers the user the most appropriate next steps to solve the

analytical activity problem, and the process of automated construction of analytical activity scenarios based on a library of scenarios (precedents) for different subject domains.

The implementation of the proposed models, methods and computer tools in a three-tier computer modeling system designed to test and evaluate the proposed theoretical approach to modeling the analytical activity scenarios, provided an opportunity to solve the problem of building and implementing scenarios in laboratory research and in terms of application in real modeling systems.

The offered and created in work algorithms and software for processes of modeling the scenarios of analytical activity provide computer realization and interaction of the received models for graphic representation of analytical activity scenarios, semantic models of subject domain of analytical activity problems and knowledge of process for construction and execution of analytical activity scenarios as adaptive object model. The developed intelligent software tools allow to automate the construction of the analytical activity scenario taking into account its previous steps and on the basis of derivation of precedents that have already been accumulated as a result of preliminary modeling of analytical activity scenarios. This approach reduces the logical errors and semantic incoherence of the description of multi-level scenarios or their individual branches and thus leads to a significant reduction of modeling time of analytical scenarios.

The created methods and tools of mathematical and computer modeling of analytical activity scenarios are used to solve the problem of construction and realization of modeling scenarios for functioning of measuring hydroacoustic system to improve the algorithms of processing the hydroacoustic signals; scenarios of energy supply distribution for optimization the energy consumption; scenarios of interactive processing of multidimensional data; construction and implementation of scenarios of analytical activities in the problems of spatial modeling; monitoring of financial indicators of budget execution at all levels.

The results of the thesis research were implemented in the Subsidiary Enterprise “National Development Fund”, LLC “Research and Production Enterprise “Symbol”, LLC “Engineering Logic”, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

Keywords: analytical activity, scenario of analytical activity, ontological model of the domain, computer modeling of multilevel scenarios of analytical activity, intelligent method for determining subsequent events of the scenario, three-level computer modeling system for modeling scenarios of analytical activity.