МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЗАХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Романець Ігор Євгенович

УДК 004.056:004.94:654.18

ІНТЕЛЕКТУАЛІЗОВАНА КОМП'ЮТЕРНА СИСТЕМА ВИЯВЛЕННЯ ЗЛОВМИСНИХ ПОВІДОМЛЕНЬ IP-ТЕЛЕФОНІЇ В КОРПОРАТИВНІЙ МЕРЕЖІ

05.13.05 – комп'ютерні системи та компоненти

Автореферат

дисертації на здобуття наукового ступеня
кандидата технічних наук

Тернопіль – 2025

Дисертацією є рукопис.

Робота виконана у Західноукраїнському національному університеті Міністерства освіти і науки України.

Науковий керівник: доктор технічних наук, професор
САЧЕНКО Анатолій Олексійович,
Західноукраїнський національний університет,
(м. Тернопіль),
професор кафедри інформаційно-обчислювальних систем і управління.

Офіційні опоненти: доктор технічних наук, професор,
ОПАНАСЕНКО Володимир Миколайович
Інститут кібернетики ім. В.М. Глушкова НАН України,
(м. Київ),
Завідувач відділу мікропроцесорної техніки №205

кандидат технічних наук, доцент,
КЛЬОЦ Юрій Павлович
Хмельницький національний університет (м. Хмельницький),
завідувач кафедри кібербезпеки

Захист відбудеться **28 серпня 2025 р. о $14^{\underline{00}}$** годині на засіданні спеціалізованої вченої ради Д 58.082.02 у Західноукраїнському національному університеті за адресою: 46009, м. Тернопіль, вул. Львівська, 11А (корпус 11, зал засідань вченої ради).

З дисертацією можна ознайомитись у бібліотеці Західноукраїнського національного університету за адресою: 46009, м. Тернопіль, вул. Сергія. Короля, 4.

Вчений секретар
спеціалізованої вченої ради Д 58.082.02
кандидат технічних наук, доцент Леся ДУБЧАК

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Сучасні корпоративні мережі все частіше використовують IP-телефонію (VoIP) як ефективний засіб комунікації через її гнучкість, масштабованість та економічні переваги. Однак зростання популярності VoIP супроводжується збільшенням кількості кіберзагроз, зокрема зловмисних повідомлень (спам, фішинг, атаки на відмову в обслуговуванні тощо), що становлять серйозну небезпеку для конфіденційності, цілісності та доступності комунікаційних сервісів.

IP-телефонія забезпечує гнучкість, масштабованість та економічність корпоративного зв'язку, дозволяє інтегрувати голосові сервіси з іншими інформаційними системами, включаючи CRM, ERP та інструменти управління знаннями.

Однак, разом із перевагами, IP-телефонія стає привабливою ціллю для зловмисників, що обумовлює виникнення нових загроз, зокрема – розсилання зловмисних SIP-повідомлень, фішингу, спуфінгу, DoS-атак, перехоплення конфіденційної інформації.

Значний внесок у сферу дослідження безпеки IP-телефонії (VoIP) та виявлення зловмисних повідомлень внесли зарубіжні та вітчизняні вчені, список яких наведено у дисертаційні роботі.

У контексті сучасних викликів, зумовлених гібридними загрозами та цифровими трансформаціями, питання підвищення ефективності виявлення зловмисних повідомлень в IP-телефонії набуває особливої актуальності. Це особливо важливо для корпоративного середовища, де забезпечення інформаційної безпеки прямо впливає на стійкість бізнесу, збереження критичних даних та довіру клієнтів.

На сьогодні існуючі методи, технічні засоби та програмні рішення не дозволяють своєчасно виявити та локалізувати зловмисну активність у VoIP-мережах. Більшість з них не враховують динаміку змін у поведінкових шаблонах трафіку, мають низьку адаптивність до нових типів загроз і не здатні до самонавчання. Це вимагає впровадження нових підходів, що базуються на використанні технологій штучного інтелекту, машинного навчання, експертних систем та аналізу аномалій.

У зв'язку з цим актуальним є науково обґрунтоване розв'язання задачі підвищення ефективності захисту IP-телефонії із врахуванням своєчасності виявлення зловмисних VoIP-повідомлень шляхом створення інтелектуалізованої комп'ютерної системи, що відповідає сучасним тенденціям у сфері кібербезпеки та цифрової трансформації.

Розв'язання цієї задачі сприятиме підвищенню рівня захищеності корпоративних інформаційних систем, забезпеченню безперервності бізнес-процесів, зниженню ризиків фінансових та репутаційних втрат.

Зв'язок роботи з науковими програмами, планами, темами. Науково-дослідна робота за темою дисертації проводилася у відповідності з координаційним планом науково-дослідних робіт і науково-технічних програм Міністерства освіти і науки України, виконанні фундаментальної держбюджетної роботи «Математичне та комп'ютерне моделювання об'єктів з розподіленими параметрами на основі поєднання онтологічного та інтервального аналізу» (держреєстраційний номер 0122U001497).

Мета і задачі дослідження. Метою роботи є підвищення ефективності захисту IP-телефонії у корпоративній мережі шляхом створення інтелектуалізованої комп'ютерної системи виявлення зловмисних повідомлень, що базується на онтологічній моделі її функціонування та методах інтелектуального аналізу даних.

Для досягнення цієї мети у дисертаційній роботі поставлено наступні завдання:

1) аналіз сучасних методів та засобів захисту IP-телефонії у корпоративній мережі, які включають компоненти інтелектуального аналізу даних;
2) розробити онтологічну модель функціонування IP-телефонії, яка включатиме формалізацію основних понять та зв’язків між ними;
3) розробити метод автоматизованого наповнення онтології тематичних повідомлень в корпоративній мережі;
4) розробити метод виявлення аномалій в трафіку IP-телефонії на основі групування повідомлень;
5) розробити засіб захисту від термінації трафіку в IP-телефонії на основі використані нечіткої логіки, що сприятиме підвищенню рівня безпеки та якості зв'язку.
6) розробити архітектуру та структурно-функціональні схеми базових підсистем інтелектуалізованої системи;
7) здійснити експериментальні дослідження розроблених методів та засоби виявлення аномальних повідомлень у IP-телефонії корпоративної мережі.

Об’єкт досліджень: процеси захисту IP-телефонії з використанням інтелектуальних методів аналізу даних.

Предмет досліджень: методи та засоби захисту IP-телефонії в корпоративній мережі.

Методи дослідження. Для вирішення поставлених задач у дисертаційній роботі використовувалися такі методи: методи системного аналізу та огляду літератури – методи штучного інтелекту, зокрема онтологічного інжинірингу та представлення знань; методи машинного навчання, зокрема кластерного аналізу (групування) та виявлення аномалій; методи об'єктно-орієнтованого проектування та програмування; методи експериментальних досліджень та статистичної обробки даних.

Наукова новизна отриманих результатів.

Основні наукові результати, висунуті на захист.

Вперше розроблено:

– онтологічну модель опису VoIP повідомлень в системі IP-телефонії, яка включає формалізацію основних понять у формі окремих концептів та опис зв’язків між цими поняттями, що у сукупності створило можливість розробки нових, більш ефективних методів виявлення аномалій в трафіку IP-телефонії;
– метод виявлення аномалій в трафіку IP-телефонії, який, ґрунтується на поєднані методів групування VoIP-повідомлень та контекстно-частотного аналізу, що у сукупності дозволило підвищити ефективність виявлення аномальних повідомлень.

Удосконалено:

- підсистему захисту від термінації трафіку в IP-телефонії, яка відрізняється від відомих тим, що здатна гнучко адаптуватися до змінних умов і приймати рішення на основі нечітких даних, що дозволяє враховувати численні критерії, такі як IP-адреса клієнта, час здійснення дзвінка, часовий пояс і якість зв'язку, а також сприяє підвищенню рівня безпеки та якості зв'язку.
- архітектуру системи виявлення зловмисних повідомлень IP-телефонії в корпоративній мережі, яка відрізняється від відомих наявністю блоків і процедур виявлення аномальних повідомлень в трафіку IP-телефонії на основі онтологічного підходу та розподілу доступу в режимі реального часу з використанням нечіткої логіки, що дозволило спростити функції адміністратора системи та зменшити час оброблення повідомлень.

Практичне значення отриманих результатів полягає в тому, що за результатами проведених автором теоретичних досліджень:

– розроблено систему, яка демонструє підвищення рівня виявлення зловмисних повідомлень при зниженні кількості хибних спрацьовувань в корпоративній мережі, що забезпечує значне зменшення фінансових втрат від VoIP-шахрайства та підвищення надійності зв'язку;

– створено прототип програмного рішення для Asterisk із відкритим інтерфейсом налаштування правил реагування, що спрощує його впровадження та адаптацію під потреби підприємств різного масштабу;

– запропоновані методи нечіткого виведення та онтологічного аналізу можуть бути використані як окремі модулі в існуючих системах безпеки мережі, що підвищує масштабованість та гнучкість рішення;

– результати дослідження можуть лягти в основу технічних рекомендацій із захисту IP-телефонії для IT-відділів корпоративних замовників, а також можуть бути враховані при формуванні політик інформаційної безпеки в галузевих стандартах.

Теоретичні та практичні результати роботи використані у ПП «Колумбус», компанії «Infineon Technologies AG» та у Тернопільській обласній військовій адміністрації в управління цифрової трансформації для захисту від несанкціонованого доступу до інформації, а також у навчальному процесі при викладанні дисциплін "Комп'ютерні системи", «Комп'ютерні мережі», «Програмно-апаратні засоби захисту інформації».

Особистий внесок здобувача. Усі наукові результати, викладені у дисертаційній роботі, отримані автором особисто і повністю розкрито у публікаціях. У друкованих працях, опублікованих у співавторстві, автору належить: метод виявлення аномалій в трафіку IP-телефонії за допомогою групування VoIP повідомлень на основі контекстно-частотного аналізу; метод автоматизованого наповнення онтології тематичних повідомлень у корпоративній системі IP-телефонії, який ґрунтується на формалізованому представленні повідомлень за допомогою деревовидних структур та на описі операцій взаємодії засобами алгебри кортежів; запропоновано онтологію для опису повідомлень у системах IP-телефонії, яка формалізує основні концепції як окремі сутності та деталізує зв'язки між ними.

Апробація результатів дисертації. Основні положення та результати дисертаційної роботи доповідались та обговорювались на 7 міжнародних науково-практичних конференціях, зокрема:

1. IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications – IDAACS (м. Бухарест, Румунія, 2017 р.);
2. 10th International Conference on Electronics, Computers and Artificial Intelligence – ECAI (м. Яси, Румунія, 2018 р.);
3. 11th International Conference on Advanced Computer Information Technologies – ACIT (м. Деггендорф, Німеччина, 2021 р.);
4. 12th International Conference on Advanced Computer Information Technologies – ACIT (м. Руженберок, Словаччина, 2022 р.);
5. IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications – IDAACS (м. Дортмунд, Німеччина, 2023 р.);
6. 13th International Conference on Advanced Computer Information Technologies – ACIT (м. Вроцлав, Польща, 2023 р.);
7. 14th International Conference on Advanced Computer Information Technologies – ACIT (м. Чеське Будейовіце, Чехія, 2024 р.).

Публікації. За темою дисертації опубліковано 14 наукових праць, з них: 3 одноосібних статті у наукових фахових виданнях України, рекомендованих МОН України, 4 – у виданнях, що індексуються в міжнародних наукометричних базах Scopus/Web of Science, 7 тез доповідей на міжнародних конференціях, що індексуються в міжнародних наукометричних базах Scopus/Web of Science.

Структура та обсяг роботи. Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Робота містить 138 сторінок основного тексту. Загальний об’єм дисертації – 170 сторінок, 1 таблиця, 68 рисунків, 4 додатки, 112 найменування використаних джерел.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** наведено загальну характеристику роботи, обґрунтовано актуальність теми, визначено об’єкт та предмет дослідження, сформульовані мета і задачі розкрито наукову та практичну цінність отриманих результатів, а також подані відомості про апробацію дисертаційної роботи.

У **першому розділі** досліджено структуру IP-телефонії, а також її особливості, проведено аналіз існуючих методів і засобів виявлення зловмисних повідомлень IP-телефонії, досліджено методи штучного інтелекту для виявлення зловмисних повідомлень, методи виявлення нетипової поведінки при передачі голосу. Проведено всебічний аналіз сучасного стану та перспектив розвитку систем виявлення зловмисних повідомлень в IP-телефонії. Розглянуто ключові аспекти, що стосуються структури IP-телефонії та її особливостей, які створюють як переваги, так і нові виклики у сфері кібербезпеки. Визначено, що попри значні економічні та функціональні вигоди, притаманні IP-телефонії, її інтеграція з мережею Інтернет робить її вразливою до широкого спектру кібератак, включаючи DoS/DDoS, спуфінг, фішинг

та інші зловмисні дії, спрямовані на порушення конфіденційності, цілісності та доступності голосових комунікацій.

Проведено аналіз існуючих методів і засобів виявлення зловмисних повідомлень IP-телефонії. Детально розглянуто застосування методів штучного інтелекту, таких як машинне навчання (зокрема, методи класифікації та кластеризації) та глибоке навчання, які демонструють значний потенціал у розпізнаванні складних патернів аномальної поведінки та зловмисних повідомлень. Окрему увагу приділено методам виявлення нетипової поведінки при передачі голосу, що дозволяє ідентифікувати загрози, які не завжди підпадають під відомі сигнатури атак, базуючись на відхиленнях від нормального трафіку та поведінкових шаблонів користувачів. Виявлено, що, попри значні досягнення, існуючі підходи часто стикаються з проблемами масштабованості, високого рівня хибно позитивних спрацьовувань та складності адаптації до нових типів атак.

Визначено перспективні шляхи розвитку предметної області та здійснено постановку задачі дослідження. Встановлено, що подальші дослідження повинні бути зосереджені на розробці та вдосконаленні гібридних методів, які комбінують сигнатурний аналіз з методами штучного інтелекту та аналізом поведінки для підвищення ефективності виявлення зловмисних повідомлень. Особлива увага має бути приділена розробці адаптивних систем, здатних до самонавчання та оперативного реагування на динамічно змінювані загрози. Це дозволить підвищити рівень безпеки та стійкості систем IP-телефонії до сучасних кіберзагроз.

У **другому розділі** обґрунтовано доцільність використання методів аналізу мовлення та відповідної обробки природної мови, які дозволяють створювати більш точні та ефективні системи виявлення аномального трафіку та потенційно небезпечних комунікацій в системі VoIP, виділено напрямок використання інтелектуальних засобів для аналізу контенту в системі VoIP. Запропоновано онтологію опису VoIP повідомлень в системі IP-телефонії, здійснено формалізацію основних понять у формі окремих концептів та описано зв’язки між цими поняттями, запропоновано метод виявлення аномалій в трафіку IP-телефонії на основі групування VoIP повідомлень, який, на відміну від існуючих, ґрунтується на поєднанні знання-орієнтованого підходу з використанням онтології та контекстно-частотного аналізу, запропоновано метод автоматизованого наповнення онтології тематичних повідомлень в корпоративній системі IP-телефонії, який ґрунтується на формалізованому представленні повідомлень за допомогою деревовидних структур та на описі операцій взаємодії засобами алгебри кортежів.

Завдяки постійному розвитку мережевих технологій та протоколів передачі даних, які використовуються в VoIP, якість голосу та зв'язку постійно покращується, що безпосередньо впливає на якість обслуговування.

Сучасні VoIP-системи надають широкий спектр додаткових функцій, таких як конференц-зв'язок, зміна голосу, відеодзвінки, інтеграція з CRM-системами.

Зростання мобільності відбувається завдяки використанню мобільних додатків та можливості робити виклики через Інтернет з будь-якого місця. Це дозволяє мобільним користувачам мати доступ до VoIP-зв'язку незалежно від їхнього місця перебування.

Інтеграція з іншими технологіями – VoIP стає складовою частиною інтегрованих комунікаційних рішень, які включають в себе електронну пошту, чат, відеоконференції та інші засоби спілкування.

Особливо необхідно відзначити такий напрям, як забезпечення безпеки. Розвиток технологій шифрування, аутентифікації та контролю доступу допомагає забезпечити безпеку VoIP-зв'язку та захистити дані користувачів від несанкціонованого доступу та перехоплення. Підсумовуючи, необхідно відзначити, що розвиток VoIP відбувається швидкими темпами, відкриваючи нові можливості для спілкування та покращення бізнес-комунікацій, а напрямок забезпечення безпеки VoIP є особливо актуальним в сучасних умовах.

Оскільки забезпечення безпеки функціонування VoIP є одним із ключових аспектів, які безпосередньо впливають на якість IP-телефонії та розширення сфер її використання, то проаналізуємо більш детально відомі методи та засоби, які використовуються для захисту інформації у VoIP (рисунок 1).

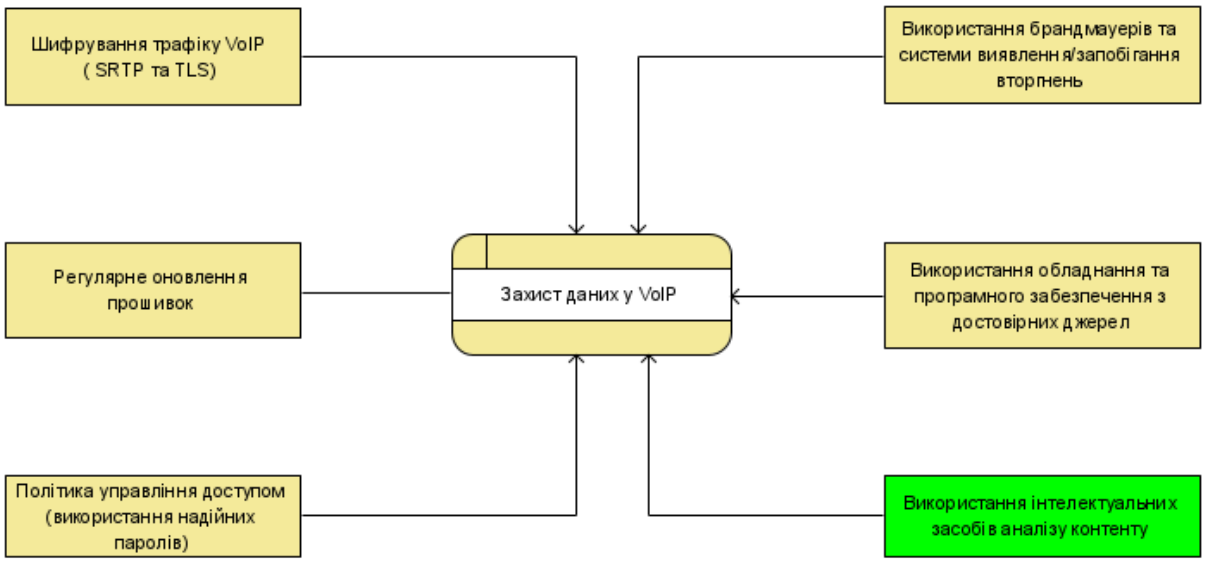

Рис. 1. Підходи до захисту інформації в системі VoIP

Системи VoIP вразливі до різних загроз безпеки, включаючи підслуховування, фішинг і атаки на відмову в обслуговуванні (DoS). Відомими підходами до зміцнення безпеки в IP-телефонії є наступні:

- шифрування трафіку VoIP за допомогою різних протоколів, таких як безпечний транспортний протокол реального часу (SRTP) і безпечний протокол транспортного рівня (TLS);
- ефективна політика управління доступом, що включає використання надійних паролей та контроль доступу, щоб запобігти несанкціонованому проникненню;
- регулярне оновлення прошивок та програмного забезпечення для виправлення вразливостей безпеки;
- використання брандмауерів та систем виявлення/запобігання вторгненням для захисту від кіберзагроз;
- використання обладнання та програмного забезпечення VoIP від перевірених постачальників, які дотримуються відповідних стандартів та протоколів;
- використання інтелектуальних засобів аналізу контенту.

З постійним розвитком технологій штучного інтелекту, особливо цікавим є напрямок використання інтелектуальних засобів для аналізу контенту в системі VoIP. Розглянемо основні сучасні підходи, які можна використовувати для аналізу VoIP контенту (рисунок 2).

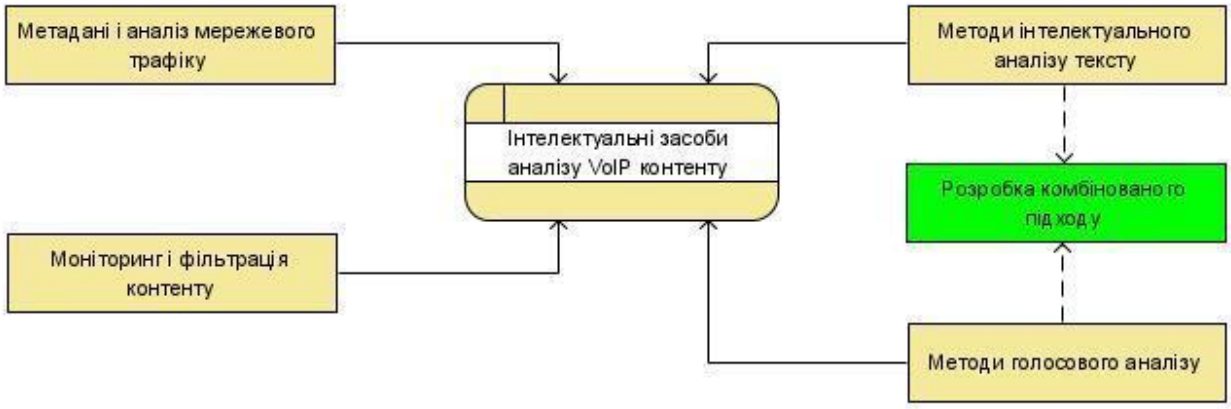

Рис. 2. Підходи до інтелектуального аналізу контенту в системі VoIP

Використання алгоритмів машинного навчання та обробки природної мови (NLP) для сканування тексту VoIP-повідомлень на ключові слова, фрази або семантичні зв'язки дозволяють знаходити інформацію, яка носить зловмисний характер. Використання методів та спеціалізованого програмного забезпечення для голосового аналізу допомагає визначати специфічні звуки, мовні ознаки або відтінки, що характеризують пропаганду чи загрози. Дослідження метаданих VoIP-повідомлень, таких як IP-адреси, час і тривалість дзвінка, здійснюється для виявлення аномалій або звичайних патернів, що можуть вказувати на можливу злочинну діяльність.

Важливо відзначити, що ефективність аналізу VoIP-повідомлень (як голосових, так і текстових) на вміст злочинної інформації може залежати від комбінації цих методів та технологій, а також від постійного вдосконалення і адаптації до нових загроз.

Створення онтології функціонування VoIP вимагає співпраці між експертами в предметній області, інженерами та дослідниками, які знайомі з технологіями, стандартами та найкращими підходами до організації VoIP. Така онтологія сприятиме полегшенню взаємодії, семантичній інтеграції та допомагатиме в системному проектуванні, управлінні, виправленні несправностей, а також сприятиме покращенню системи захисту інформації в VoIP.

Онтологія даної предметної області формується на основі структурованого представлення концептів, сутностей, зв’язків у сфері технології VoIP. Вона дозволить отримати формалізовану та семантично насичену структуру для розуміння та представлення знань про протоколи, компоненти та їх взаємодію.

Онтологія VoIP повинна включати:

- поняття та сутності: стек протоколів VoIP, представлення багаторівневої архітектури протоколів VoIP, включаючи такі протоколи, як SIP (протокол ініціації сеансу), RTP (транспортний протокол реального часу), SDP (протокол опису сеансу) та інші; компоненти VoIP: об’єкти, залучені до зв’язку VoIP, такі як користувацькі додатки (мобільні телефони з додатками, IP-телефони), шлюзи VoIP, проксі-

сервери, реєстратори та медіа-сервери; опис життєвого циклу VoIP виклику: представлення різних етапів VoIP виклику, включаючи встановлення виклику, узгодження, розрив виклику та систему управління викликом;

- відношення: комунікаційні зв'язки між об'єктами VoIP, такі як зв'язок користувач-користувач, зв'язок користувач-сервер, зв'язок сервер-сервер і зв'язок медіа-потоку; відношення залежностей між компонентами та протоколами VoIP, такі як залежність SIP від базових транспортних протоколів, таких як UDP або TCP, і залежність RTP від RTCP для зворотного зв'язку та управління;
- властивості та атрибути: властивості протоколу VoIP, такі як формати повідомлень, заголовки, методи, параметри та коди стану, визначені SIP, RTP та іншими відповідними протоколами; атрибути компонентів VoIP, наприклад IP-адреси, порти, кодеки, підтримувані функції, можливості та конфігурації; події, пов'язані зі зв'язком VoIP, такі як ініціювання виклику, завершення виклику, утримання/відновлення виклику, переадресація виклику, переадресація виклику та події медіапотоку; дії, які виконують об'єкти VoIP у відповідь на події, такі як надсилання SIP-повідомлень, узгодження параметрів медіа, встановлення та розрив медіа-сеансів і обробка сигналів управління викликами;
- ієрархічна таксономія: класифікація об'єктів і протоколів VoIP в ієрархічні таксономії на основі їхніх ролей, функцій і зв'язків; категоризація технологій VoIP на основі моделей розгортання (локальні, хмарні), мережевих архітектур (однорангові, клієнт-сервер) і сценаріїв використання (корпоративні VoIP, мобільний VoIP);
- семантичні обмеження: обмеження на зв'язки та взаємодію між об'єктами VoIP, що забезпечують узгодженість в онтології; правила, що регулюють поведінку та використання протоколів VoIP, включаючи правила обробки повідомлень, встановлення сеансу та узгодження медіа для окремих протоколів.

При побудові онтології опису VoIP повідомлень в системі IP-телефонії необхідно здійснити формалізацію основних понять у формі окремих концептів та описати зв'язки між цими поняттями. Розробка онтології для повідомлень VoIP передбачає створення структурованого представлення концептів, сутностей та зв'язків у межах відповідного домену. На рисунку 3 представлено онтологічний граф для формалізованого опису VoIP повідомлень.

На першому етапі здійснюємо ідентифікацію домену, яка включає визначення області та меж описуваної онтології. У нашому випадку це буде зв'язок за протоколом передачі голосових повідомлень через Інтернет , що охоплює концепції, пов'язані з передачею голосових даних через IP-мережі.

На другому етапі визначаємо основні поняття та сутності, які пов'язані з VoIP повідомленнями:

- протоколи VoIP (SIP, RTP, RTCP);
- компоненти VoIP (клієнт VoIP, сервер VoIP, шлюз VoIP);
- типи повідомлень VoIP (пакети INVITE, ACK, BYE, RTP);
- параметри VoIP (IP-адреса клієнта, IP-адреса призначення, номери портів, тип корисного навантаження);
- повідомлення (голосові та текстові), текстові повідомлення характеризуються основної частиною, автором, переліком ключових слів, оцінкою та комбінаційним представленням.

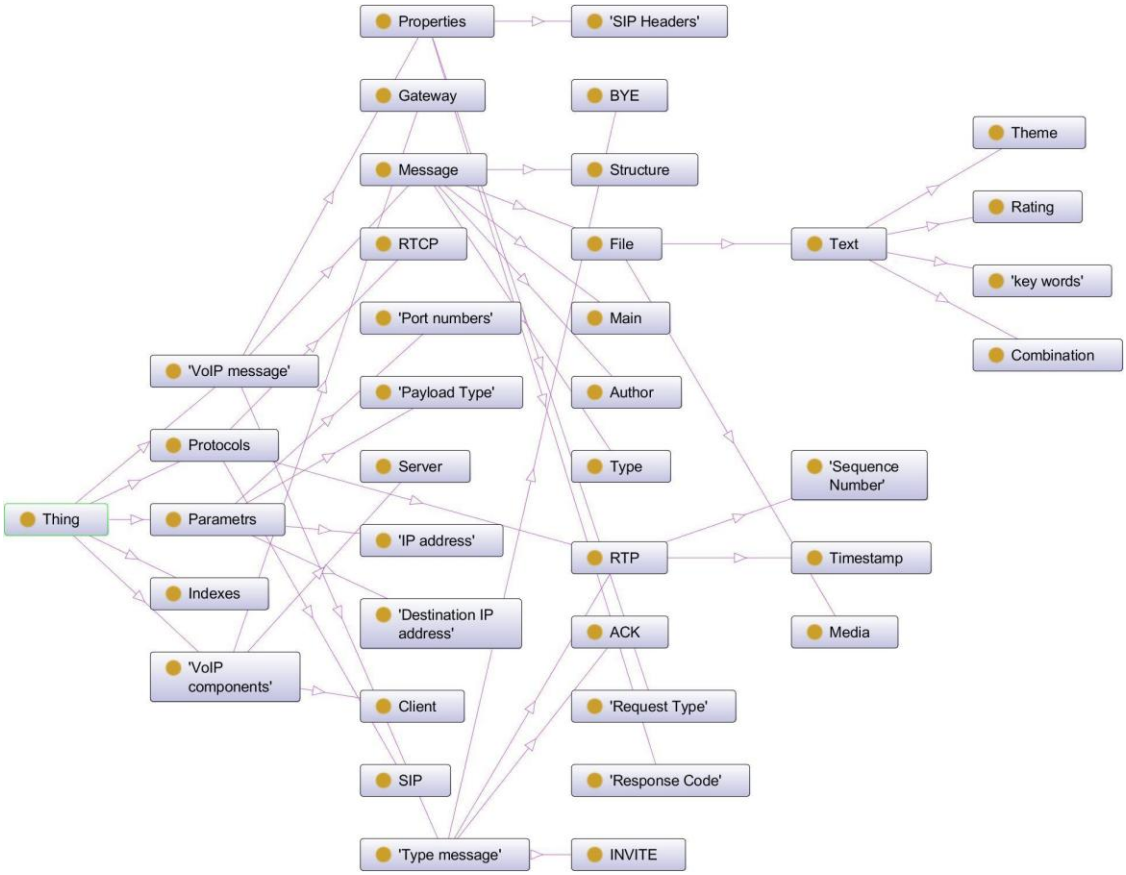

Рис. 3. Представлення формалізованої предметної області «VoIP повідомлення» у вигляді онтологічного графу

На третьому етапі формуємо опис ієрархічних зв'язків та будуємо таксономію класів, наприклад: протоколи VoIP (SIP, RTP, RTCP), компоненти VoIP (клієнт, сервер, шлюз). Будуємо відповідні відношення між класами, тобто визначаємо відношення між різними поняттями. Цей процес може включати такі зв'язки «є» (успадкування), «частина» (композиція), «має» (асоціація) тощо.

На четвертому етапі визначаємо властивості та атрибути, пов'язані з кожним поняттям. Вони представляють характеристики або ознаки, які їх описують.

На п'ятому етапі формуємо систему правил та обмеження для забезпечення узгодженості онтології. Це може включати обмеження відповідності правил домену предметної області, обмеження типів даних. Заключний етапом є перевірка онтології на реальних сценаріях VoIP, яка включає перевірку адекватності відношення до домену і відповідності до вимог практичного використання.

Метод автоматизованого наповнення онтології тематичних повідомлень в корпоративній мережі.

Онтологія описується деревоподібною системою концептів досліджуваної предметної області. Під концептами предметної області розуміються як базові поняття, так і узагальнюючі поняття, які часто є спільними у багатьох суміжних предметних областях. Онтологію для VoIP можна змоделювати деякою деревовидною структурою наступного виду

$$OC\_Voip =< IdConcept, ParentConcept, Base >, \qquad (1)$$

де $IdConcept$ - ідентифікатор концепту; $ParentConcept$ - ідентифікатор батьківського концепту; $Base$ - атрибут, який використовується для розмежування базових понять від узагальнюючих понять.

Базове поняття відрізняється від узагальнюючого тим, що узагальнюючі поняття не мають посилання на батьківський елемент, тобто $ParentConcept = 0$. Кожне поняття може описуватися деякими лінгвістичними представленнями у вигляді OC_Prs словосполучень. Онтологічні поняття описуються конкретними словоформами OC_Form, а також основами OC_Base цих понять. Словоформи можуть використовуватися для опису понять для користувачів, а основи понять – для автоматичного визначення еквівалентності представлень відповідно до вибраної мови. Атрибути таких понять групуємо в структури:

$$\begin{gathered} OC\_Base =< IdLanguage, IdBase, WordBase >, \\ OC\_Form =< IdLanguage, IdForm, IdBase, WordForm >, \\ OC_{Prs} =< IdConcept, IdLanguage, IdPhrase, \\ IdBase, IdForm, IdParentBase >, \end{gathered} \quad (2)$$

де $IdLanguage$ - ідентифікатор мови; $IdBase$ - ідентифікатор визначеної основи слова; $WordBase$ - визначена основа слова; $IdForm$ - ідентифікатор визначеної форми слова; $WordForm$ - визначена словоформа; $IdConcept$ - ідентифікатор аналізованого концепту; $IdPhrase$ - ідентифікатор аналізованої фрази; $IdParentBase$ - ідентифікатор основи поняття, що належить до батьківської категорії.

Для наповнення онтології VoIP можна використовувати:

– аналіз предметної області VoIP, включаючи протоколи, компоненти систем, типи повідомлень, характеристики зв'язку тощо;

– термінологічні словники, які використовуються у VoIP. Це включає технічні терміни, абревіатури, назви протоколів тощо;

– стандарти та специфікації VoIP, такі як SIP (Session Initiation Protocol), RTP (Real-time Transport Protocol), SDP (Session Description Protocol) тощо. Ці документи надають інформацію про протоколи, повідомлення, функції та інші аспекти VoIP;

– інформацію, що опублікована у відкритому доступі, таку як повідомлення в тематичних групах в соціальних мережах, форуми, вікі, які можуть містити корисні відомості про проблеми, вирішення, нові можливості розвитку в галузі VoIP;

– співпрацю з експертами у галузі VoIP, такими як інженери, адміністратори мереж, дослідники, які можуть допомогти у розумінні складних питань;

– програмні засоби VoIP, такі як Asterisk, FreeSWITCH, Kamailio, які містять документацію, код та інші ресурси, які можна використовувати для розуміння архітектури та принципів роботи VoIP;

– для формування бази тематичних медіа-повідомлень можна також використовувати спеціальні записи з медіа-ресурсів (Youtube, Tiktok, тощо).

Для пошуку та вибірки структурованої інформації, яка відноситься до аналізованої предметної області, формуємо множину $KeyConSet$, що містить ключові поняття, які характеризують її визначення та поведінку. Для аналізу структури VoIP повідомлень визначимо деяку допоміжну множину $AdvSet$:

$$\begin{gathered} AdvSet =< IdMessAn, IdListElem, IdElement, \\ IdBase, IdForm, IdParentBase >, \end{gathered} \quad (3)$$

де $IdMessAn$ – ідентифікатор аналізованого повідомлення; $IdListElem$ – ідентифікатор набору всіх понять; $IdElement$ – ідентифікатор окремого елемента набору.

Аналіз повторюваних понять дозволяє визначити найбільш значущі поняття і відокремити їх від несуттєвих понять. Для аналізу та контролю важливих понять використаємо деяку структуру $BaseFr$ частот основ:

$$BaseFr =< IdBase, BsFreq, IdLastMess, ConBack >, \quad (4)$$

де $IdBase$ – ідентифікатор основи, $BsFreq$ – частота появи основи в різних повідомленнях, $IdLastMess$ - ідентифікатор останнього повідомлення, де була основа поняття, $ConBack$ – відмітка фоновості поняття, що приймає $NULL$ значення за замовчуванням при деякій невизначеності.

При аналізі набору текстових повідомлень, які згруповані за відповідним критерієм, встановлюємо деякий ідентифікатор такого набору:

$$CurrMessId := \left(\pi_{IdLastMess}(BaseFr)\right) + 1. \quad (5)$$

Аналізовані елементи сформованої множини розбиваємо на набір елементарних понять із використанням роздільників, списки розбиваються на елементарні поняття $ElemCon$ із використанням роздільників.

До частин елементарного поняття необхідно застосувати процедуру $BaseConcept$ для формування їх основ. Ця процедура реалізовується шляхом відсікання закінчень від слів. При умові, що

$$BaseConcept(ElemCon_{Mess,List}) \subset BaseConcept(KeyConSet), \quad (6)$$

всі елементарні поняття з аналізованої множини з відповідними атрибутами додаємо в визначену структуру $AdvSet$. Словоформи вибираємо з елементів списку $ElemCon_{Mess,List}$ і розпізнаємо за допомогою відношення OC_Form або поповнюємо його.

Нехай $Word_k(ElemCon_{Mess,List})$ – k – те слово, яке відокремлене із елемента списку $ElemCon_{Mess,List}$. Якщо основа слова вже існує в аналізованій множині, то цей ідентифікатор $IdBaseWord$ визначаємо з відношення OC_Base:

$$IdBaseWord = \\ = \pi_{IdBase}\left(\sigma_{WordBase=BaseConcept(Word_k(ElemCon_{Mess,List}))}(OC\_Base)\right). \quad (7)$$

Якщо основа слова вже знаходиться у відношенні $BaseFr$ і номер аналізованого повідомлення не співпадає із номером зарахованого на попередньому кроці, то частотний індекс $BsFreq$ збільшуємо на 1, а номер поточного повідомлення заноситься в поле $IdLastMess$:

$$\left(Count\left(\pi_{IdBase}\left(\sigma_{IdBase=IdBaseWord\ AND\ IdLastMess<>CurrMessId}(BaseFr)\right)\right) \neq 0\right) \Rightarrow (BaseFr.BsFreq := BaseFr.BsFreq + 1) \wedge \\ \wedge (BaseFr.IdLastMess := CurrMessId).0) \quad (8)$$

Такий підхід забезпечує врахування частот використання основ у різних повідомленнях. Коли основа у відношенні $BaseFr$ не знайдена, її враховуємо з частотним індексом, який рівний 1, та додаємо у відношення, а також враховуємо номеру поточного повідомлення:

$$\left(Count\left(\pi_{IdBase}\left(\sigma_{IdBase=IdBaseWord}(BaseFr)\right)\right)=0\right)\Rightarrow\Rightarrow(BaseFr.IdBase:=IdBaseWord)\wedge(BaseFr.BsFreq:=1)\wedge\wedge(BaseFr.IdLastMess:=CurrMessId). \quad (9)$$

Якщо основа $Word_k(ElemCon_{Mess,List})$ не визначена, то вона додається у список основ, слово додається в множину словоформ, а відношення часто визначається на основі.

Для додавання поняття в онтологію експерту із знанням предметної області у IP-телефонії пропонуються тільки основи з частотою, яка більша за визначене мінімальне значення $BaseFr_0 \geq 2$, яке вибирається користувачем. Експерт із запропонованого списку основ здійснює вибір і включення в онтологію, коли дану умову задовольняє не менше $BaseFr_0$ основ

$$Count\left(\pi_{IdBase}\left(\sigma_{BaseFr>=BaseFr_0}(BaseFr)\right)\right)>BaseFr_0. \quad (10)$$

В процесі прийняття відповідного рішення, основи відображаються лише з контексту аналізованих повідомлень. Це дозволяє формувати поняття з кількох слів і не повторювати основи, які не обрані спеціалістом для включення в онтологію. Основу для контексту $ContextBase$ вибираємо виходячи із критерію максимальної частоти $MaxBsFreq$:

$$MaxBsFreq=\left(\pi_{BsFreq}(\sigma_{ConBack=Null})\right), \quad (11)$$

$$ContextBase=rand\left(\pi_{IdBase}\left(\sigma_{BsFreq=MaxBsFreq}\right)\right).$$

Формування контексту визначеної основи здійснюється із використанням двовимірного масиву $ArrayCont$, який містить словоформні ідентифікатори. Перший індекс характеризує номер фрази в контексті, а інший – ідентифікатор номера слова у фразі. $NumbPhrCont$ – показник, який характеризує кількість фраз контексту:

$$NumbPhrCont=count\left(\pi_{IdElement}\left(\sigma_{IdBase=CBase}(AdvSet)\right)\right). \quad (12)$$

Ідентифікатори усіх елементів з аналізованої структури заносимо в сформований масив $AdvPhrId$, який носить допоміжний характер:

$$AdvPhrId=\pi_{IdElement}\left(\sigma_{IdBase=CBase}(AdvSet)\right) \quad (13)$$

розмірності $NumbPhrCont$. І-та стрічка даного контекстного масиву визначається за наступним правилом:

$$AdvContext[i]=$$

$$=\left(\pi_{IdForm}\left(\sigma_{\substack{IdElement=AdvPhrN[i]\wedge\\ \wedge IdParentBase=NULL}}(AdvSet)\right)AS\ HeadLine\right)\cup$$

$$\cup\left(\pi_{IdForm}\left(\sigma_{\substack{IdElement=AdvPhrN[i]\wedge\\ IdParentBase=\\ \wedge\left(=HeadLine.IdBase\ OR\ IdParentBase=\right)\\ =HeadLine.IdBase}}(AdvSet)\right)AS\ HeadLine_1\right) \quad (14)$$

Сформований набір словоформ представляється спеціалістові в галузі VoIP та відповідної досліджуваної предметної області для формування елементів, які будуть поповнювати онтологію. Основи, які потрапляють в онтологію, отримують фоновий показник $ConBack := 2$, щоб уникнути процедур повторного аналізу. Основи, які не використовувалися жодного разу, не можуть формувати контекстні основи, і для них показник $ConBack := 1$. Вибрані елементи, які використовуються для онтологічного наповнення, можуть формувати онтологічну ієрархію, яка дозволяє формалізувати експертні знання у формі онтології.

Метод виявлення аномалій в трафіку IP-телефонії на основі групування повідомлень.

Для аналізу VoIP контенту необхідно формалізувати метод побудови базового повідомлення в IP-телефонії. Для цього необхідно спочатку здійснити перетворення VoIP медіа повідомлення в текстове представлення і таким чином сформувати базу даних вже текстових повідомлень.

Для перетворення голосового повідомлення в текстове використаємо бібліотеки SpeechRecognition для перетворення голосу в текст у мові програмування Python. Фрагмент лістингу програми такого перетворення представлено на рисунку 4.

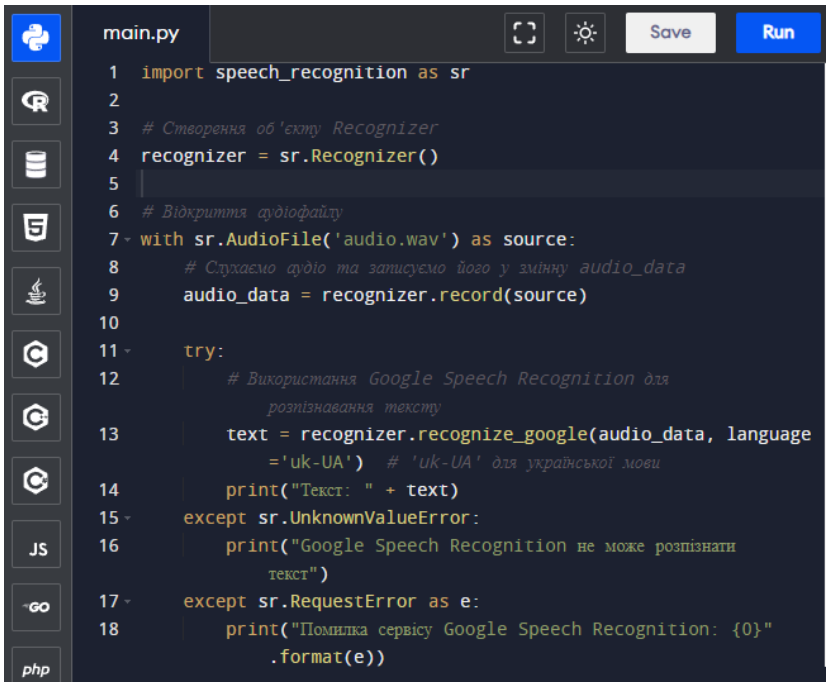

Рис. 4. Фрагмент лістингу програмного коду для перетворення VoIP медіа повідомлення в текстове представлення

Тематику таких повідомлень задаємо загальним текстовим ідентифікатором IMP предметної області та відповідним специфікатором SM. За допомогою перетворення голосового повідомлення у текстове представлення та процедури символьної конкатенації $VPS = IMP \,\&\, SM$ отримаємо множину VP текстових повідомлень, які представлені наборами відповідних компонентів конкретних понять

$$VP(VPS, P) = \{VP_i\}_{i=1}^{P}, \qquad (15)$$

де P – потужність множини VP ; VP_i – елемент множини, що визначає набір понять i - того повідомлення.

На наступному кроці для аналізу понять повідомлень використовуємо лише ті поняття, які відносяться до аналізованої предметної області, тобто належать множині VPK, а не здійснюємо повну перевірку усієї множини понять. Критерієм виконання такої умови є наявність ідентифікатора предметної області в темі повідомлення:

$$VPK = \{VP_{i*}^{*} | VP_{i*} \in VP, VP_{i*}.theme \cap \mathrm{I}MP \neq \emptyset\}_{i=1}^{P*}. \quad (16)$$

Із набору текстових тверджень необхідно вибрати інформацію, яка визначає тематику повідомлень. У першу чергу, інформація про тематику повідомлення визначається у наборі тверджень, з яких починається повідомлення.

Якщо початок повідомлення не дозволяє визначити його тематику, то інформацію про тематику вибираємо з впорядкованої множини елементів на основі частотного аналізу понять. Інформація такого типу буде впорядкованим списком $LKB(VP_i^*)$ виділених ключових понять:

$$LKB(VP_i^*) =< C_{ik}(VP_i^*) >_{k=1}^{CPi}. \quad (17)$$

Така множина може містити також і випадкову інформацію. Однак елементи множини понять, які будуть повторюватися, дають нам інформацію про структуру і тематику такого повідомлення. Тому на основі впорядкованої множини та множини ключових понять сформуємо базову BCM та узагальнену GCM множини пар поняття - частота появи поняття, які визначаємо наступним чином:

$$BCM = \{(CS_l, NCS_l) | CS_l\}, \quad (18)$$

$$GCM = \{(CS_m, NCS_m) | CS_m \in \cup_i \ LKB(VP_i^*)\}. \quad (19)$$

Для знаходження понять, які будуть релевантними до заданої предметної області, впорядкуємо елементи множини GCM у порядку спадання відповідних частот елементів, а деяку частину понять включаємо відразу в базову концептуальну множину CSC:

$$CSC = \left\{(CS_l, NCS_l) | FCL_l = \frac{NCS_l}{P^*} > FF_0\right\}, \quad (20)$$

де величина FF_0 належить інтервалу [0.200; 0.500], а конкретне значення цієї величини вибираємо із врахуванням особливостей та специфіки предметної області.

Далі аналізуємо множину понять, які мають низьку частоту, такий аналіз доцільний, якщо нам не вдалося наповнити множину понять на основі відношення. Розглянемо деяку визначену і впорядковану вибірку $SCMF$, яка також включає відповідні частоти понять у повідомленні:

$$SCMF = \{NCS_l | (CS_l, NCS_l) \in BCM\}. \quad (21)$$

Частоти з найвищими значеннями перевіряємо за критерієм 4σ на характеристику аномальності. Концепти, що відповідають критерію аномальності, включаємо в сформовану множину CSC.

При формуванні множини SK ключових концептів, що визначають тематику повідомлення, здійснюємо впорядкування цієї множини. Рангові номери присвоюємо лише ключовим концептам. Так $\mathrm{C}R\mathrm{C}_{ik}$ позначає ранг k-го концепту в і-тому повідомленні. При аналізі повідомлень необхідно також врахувати вплив тривалості голосового повідомлення (довжини перетвореного в текст повідомлення) на предметну аудиторію, оскільки є багато повідомлень, які є короткотривалими. Для того, щоб врахувати таку особливість, використаємо деяку вагову функцію $wcs(i)$.

Оскільки із спаданням тривалості голосових повідомлень їх важливість також буде плавно спадати, то чим коротше повідомлення, тим швидше відбуватиметься це спадання. Для моделювання такого процесу можна використати кубічний сплайн. Для його однозначного визначення необхідно накласти хоча б чотири умови. Зокрема, найтриваліше повідомлення буде мати важливість для цього аргументу рівну 1, якщо похідна дорівнює 0. Вводимо систему ваг $CSG(i) = \frac{csg(i)}{\sum_i \quad csg(i)}$, яка матиме нормований характер та обчислюємо усереднений ранг k-го концепту наступним співвідношенням:

$$RA_k = \sum_i \quad R_{ik} CSG(i). \tag{22}$$

Ранжування концептів здійснюємо на основі усереднення рангів. Узгодженість рангів перевіряємо із використанням коефіцієнта конкордації:

$$CSW = \frac{12 \sum_{k=1}^{K} \quad \left(\sum_{i=1}^{I} \quad R_{ik} CSG(i) - \bar{R}\right)}{I^2 (K^3 - K)}, \tag{23}$$

де $\bar{R} = \frac{1}{K} \sum_{k=1}^{K} \quad \sum_{i=1}^{I} \quad R_{ik} CSG(i)$.

Якщо виконується наступна умова $I(K-1)CSW > \chi^2_{K-1,\alpha}$, то таке ранжування є значущим.

Якщо ранжування повного списку концептів не є значущим, то відкидаємо один елемент із списку в порядку зростання відповідних ваг. Відкидання проводимо до отримання значущого ранжування, починаючи із концепту, який матиме найменшу вагу.

Запропонований метод дозволяє згрупувати повідомлення відповідно до класифікованої структури ключових понять. Таке групування доцільно використовувати для виявлення аномальних повідомлень в спеціалізованій корпоративній мережі з розгорнутою системою IP-телефонії.

У **третьому розділі** розроблено та детально описано архітектуру інтелектуалізованої системи виявлення зловмисних повідомлень IP-телефонії в умовах корпоративної мережі. Представлено узагальнену структуру інтелектуалізованої системи, яка включає низку взаємодіючих підсистем. Розроблено узагальнений алгоритм і структурно-функціональну схему інтелектуалізованої системи. Деталізація алгоритму дозволила формалізувати процеси виявлення, підвищивши їх прозорість та керованість.

Запропоновано механізм взаємозв'язку модулів для виявлення аномальних повідомлень на основі онтологічного підходу. Також запропоновано механізм автоматизованого наповнення онтології тематичних повідомлень у корпоративній мережі, що дозволяє системі постійно оновлювати свої знання про нормальні та аномальні патерни комунікації. Крім того, представлено метод виявлення аномалій у трафіку IP-телефонії на основі групування повідомлень, який дозволяє ідентифікувати відхилення у поведінці трафіку за допомогою методів кластеризації та класифікації. Деталізовано архітектуру програмного комплексу виявлення зловмисних повідомлень IP-телефонії в корпоративній мережі.

На основі проведеного вище аналізу запропоновано узагальнену структуру інтелектуалізованої системи забезпечення безпеки використання IP-телефонії в корпоративній мережі (рис. 5), серцевиною якої є п'ять базових підсистем.

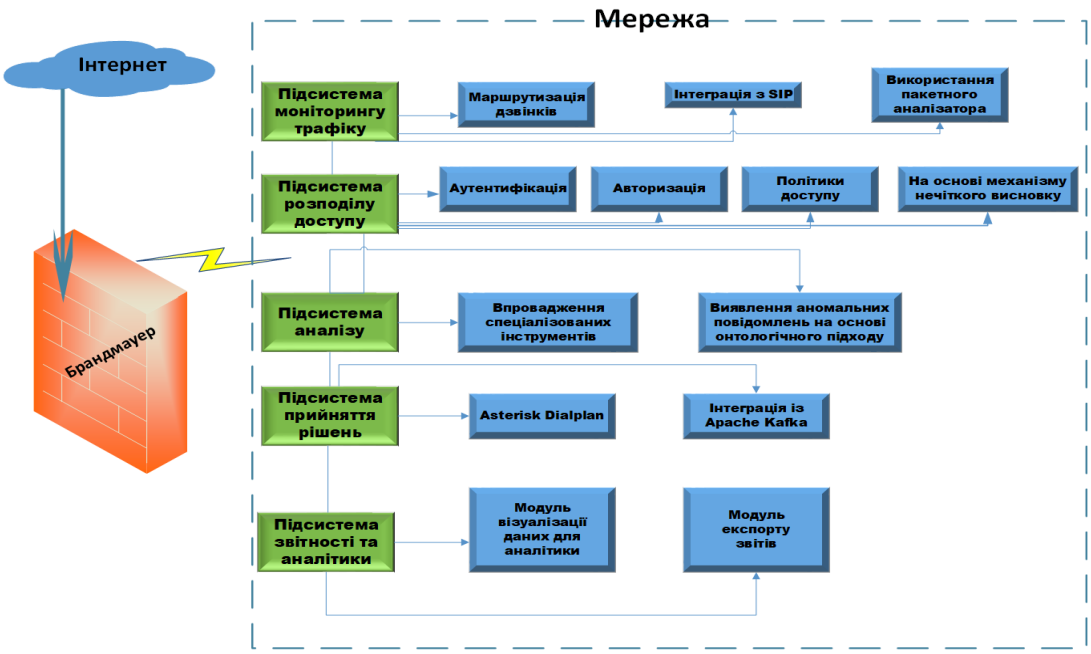

Рис. 5. Узагальнена структура інтелектуалізованої системи безпеки використання IP-телефонії в корпоративній мережі

На рисунку 6 наведена схема розподілу доступу клієнтів до мережі IP-телефонії на основі нечіткої логіки. Схема є узагальненою і може бути адаптована під конкретні вимоги та інфраструктуру.

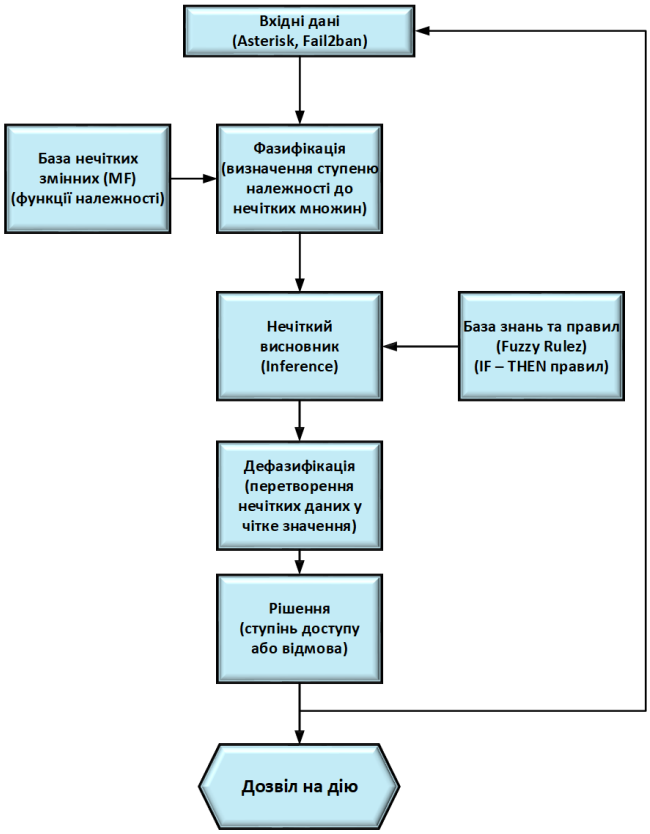

Рис. 6. Схема розподілу доступу клієнтів до мережі IP-телефонії на основі нечіткої логіки

На основі узагальненої структури інтелектуалізованої системи (див. рис. 5) та схеми розподілу доступу клієнтів (див. рис. 6) запропоновано узагальнений алгоритм функціонування інтелектуалізованої системи. Даний алгоритм описує роботу інтелектуалізованої комп'ютерної системи, яка забезпечує управління доступом та моніторингу в корпоративній мережі, також контроль доступу користувачів, моніторинг мережевого трафіку, аналіз потенційних загроз (аномалій), автоматизоване прийняття рішень щодо безпеки та формування звітності про результати.

Запропонований алгоритм є ефективним інструментом, орієнтованим на автоматизацію прийняття рішень і швидке реагування на потенційні загрози.

З врахуванням викладеного вище, розроблено структурно-функціональну схему інтелектуалізованої комп'ютерної системи виявлення зловмисних повідомлень IP-телефонії в корпоративній мережі (рис. 7.). Взаємозв'язок модулів для виявлення аномальних повідомлень на основі онтологічного підходу (рис. 8.).

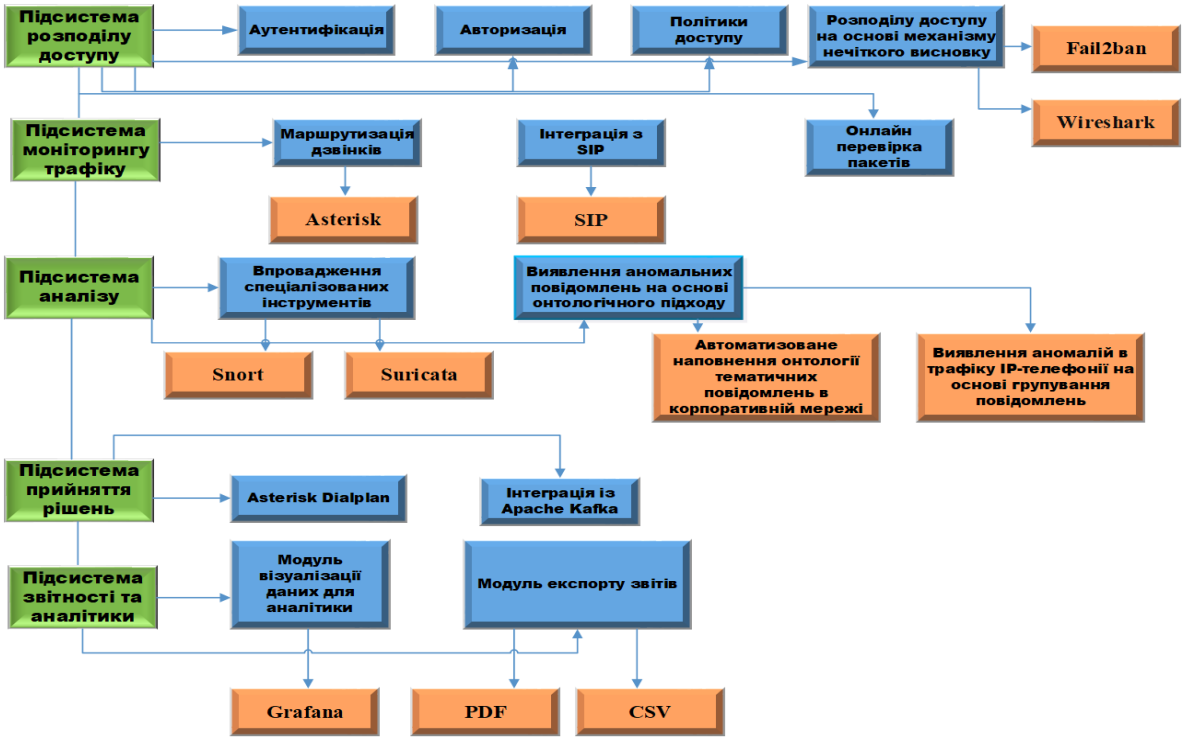

Рис. 7. Структурно-функціональна схема інтелектуалізованої системи

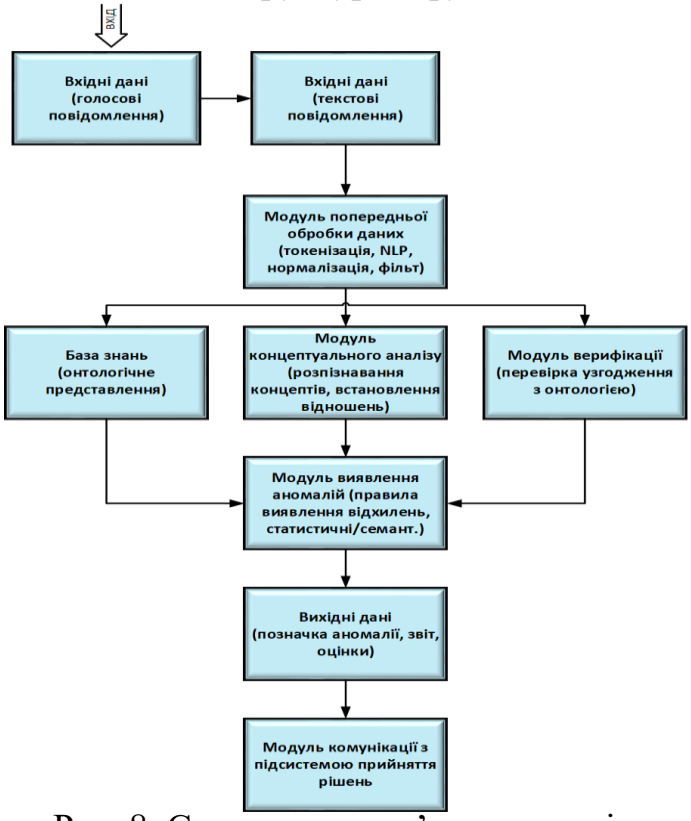

Рис. 8. Схема взаємозв'язку модулів для виявлення аномальних повідомлень на основі онтологічного підходу

Представлена схема та опис показують цілісний процес обробки та аналізу текстових повідомлень на семантичному рівні за допомогою онтологічного підходу. Він дозволяє враховувати значення, контекст та логічну узгодженість повідомлень для виявлення прихованих чи нетривіальних аномалій, які складно ідентифікувати лише статистичними чи поверхневими методами.

Архітектура програмного комплексу систем виявлення зловмисних повідомлень IP-телефонії в корпоративній мережі призначена для забезпечення безпечного функціонування мережі, захисту голосового трафіку та конфіденційності інформації, що передається через IP-мережі. Основна мета такої архітектури – мінімізувати

ризики від зовнішніх та внутрішніх загроз, забезпечуючи при цьому якісну та безперебійну роботу IP-телефонії.

На рисунку 9 представлено загальну архітектуру програмного забезпечення для захисту IP-телефонії. Необхідно відзначити, що така архітектура відображає взаємодію по рівнях.

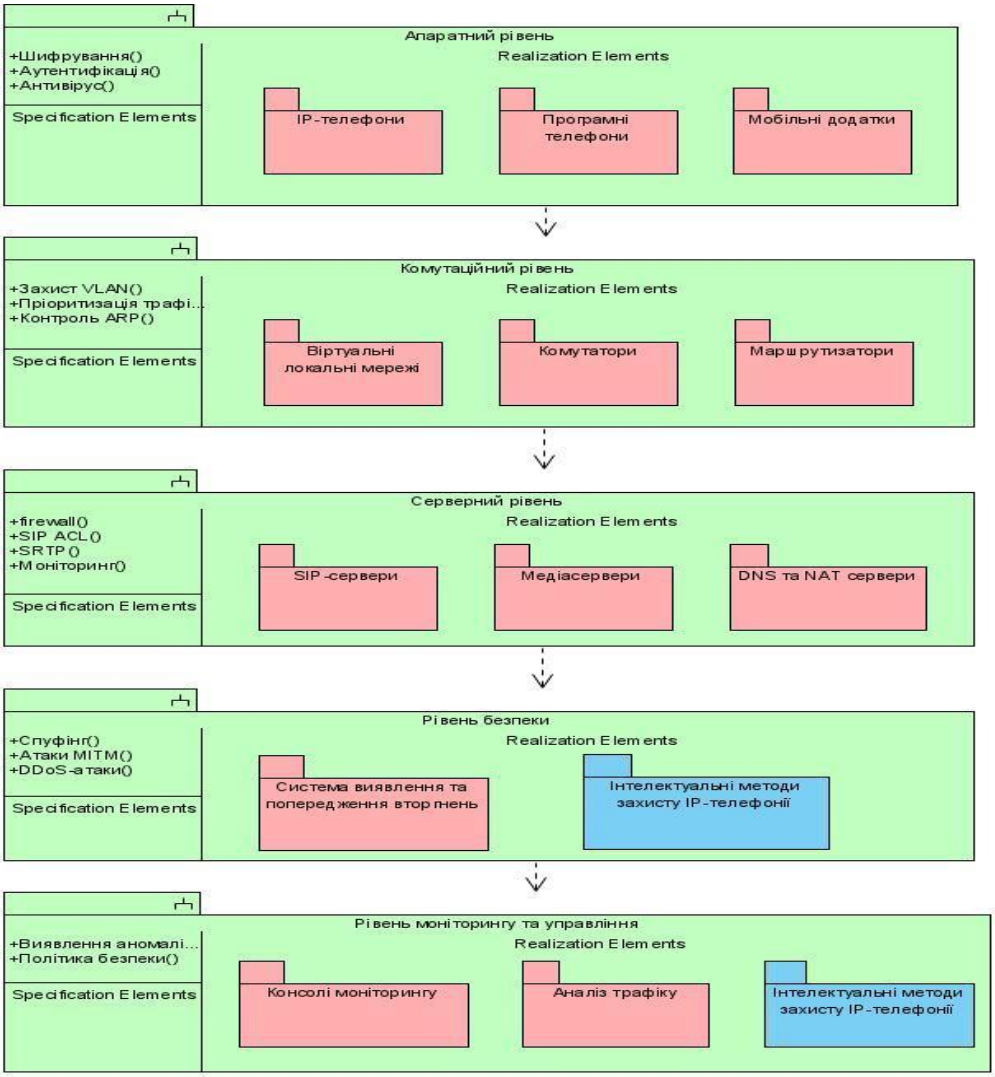

Рис. 9. Загальна архітектура програмного забезпечення для захисту IP-телефонії

Програмне забезпечення належить до класу об’єкто-орієнтованих систем, розроблене з використанням мови програмування PHP, а також СУБД MySQL. Система працює як інтегроване програмне забезпечення в середовищі Asterisk під управлінням операційною системою Linux.

Четвертий розділ присвячений дослідженню створених методів та засобів, які використовуються для підвищення ефективності захисту IP-телефонії у корпоративній мережі.

Розглянуто застосування нечіткої логіки для захисту від термінації трафіку в IP-телефонії. Основною метою є розробка нечіткої системи, яка забезпечує розподіл доступу в режимі реального часу, враховуючи різні критерії.

Моделювання нечіткої системи здійснено за допомогою Fuzzy Logic Toolbox середовища MATLAB. На вхід такої системи надходять значення рівня довіри клієнта по його IP-адресі (*IP-identification*), часу здійснення дзвінка клієнтом (*client-call-time*), часового поясу адресата дзвінка (*addr-call-time*) та якості зв'язку мережі клієнта (*connection-quality*). Загальний вигляд запропонованої нечіткої системи подано на рисунку 10.

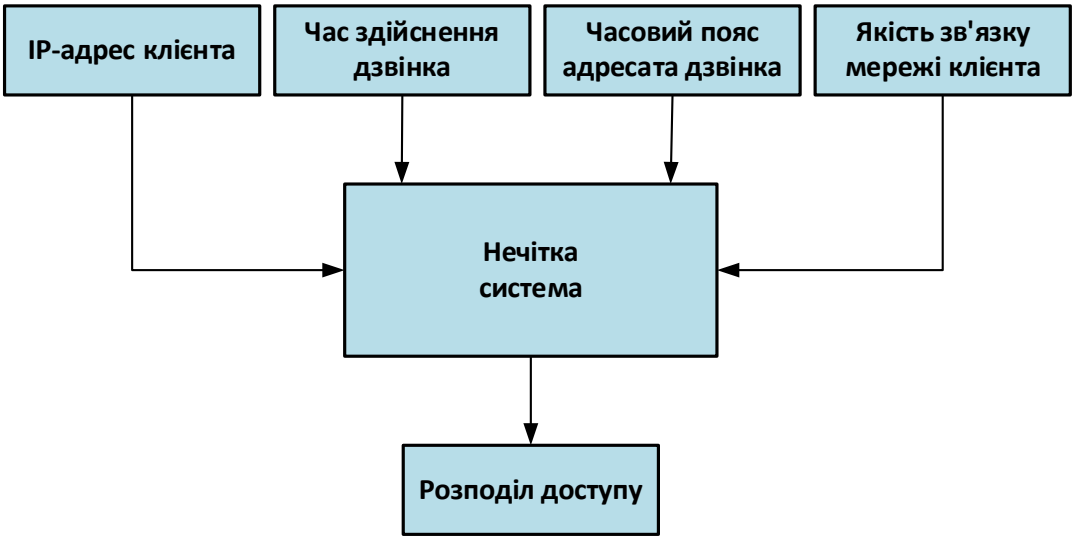

Рис. 10. Загальна схема нечіткої системи розподілу доступу в IP-телефонії

Кожна із вхідних змінних демонструє дані з певного діапазону. Над цими даними здійснюється фазифікація, тобто переведення кількісних значень у якісні.

Значення функцій належності вхідних змінних задається дзвоноподібною, а вихідної змінної – трапецевидною функцією.

Дзвоноподібна функція задається трьома числами (a, b, c), що відповідають абсцисам крайніх точок та центру кривої:

$$MF(x) = \frac{1}{1 + \left| \frac{x - c}{a} \right|^{2b}} \tag{24}$$

Функції належності вхідної змінної *IP-identification* відображаються трьома множинами: *low*, *middle*, *high*, що відображають відповідно низький, середній та високий рівень довіри клієнта IP-телефонії. Цей показник легко можна сформувати зі статистичних даних IP-адреси клієнта, його присутності в IP-телефонії, здійснених атак чи наявності збоїв при здійсненні дзвінків та правах доступу, наданих йому адміністратором мережі. Для зручності вибрано інтервал задання значень цієї змінної [0, 1].

Функції належності змінної *connection-quality*, що відображає якість зв'язку мережі, з якої здійснює запит до IP-телефонії на дзвінок клієнт, аналогічно поділені на три інтервали *low*, *middle*, *high*, що задані на відрізку [0, 1]. Низька якість зв'язку може бути у мобільних мережах через різницю у покритті та якості надання послуг

мобільними операторами. Середня якість зв'язку, як правило, можлива при застосуванні користувачем регіональної мережі Інтернет, а висока якість забезпечується використанням клієнтом локальної мережі. Для задання функцій належності змінної *client-call-time*, що показує час здійснення дзвінка клієнтом в інтервалі [0, 24], пропонується на три діапазони *morning, work time, night*.

Змінна *addr-call-time* відображає часові пояси світу, які поділені на три зони (*I-timezone*, *II-timezone*, *III-timezone*), які подані на інтервалі [-12, 12] та враховуються при побудові бази правил відносно часу здійснення дзвінка клієнта.

Трапецевидна функція задається четвіркою чисел (*a*, *b*, *c*, *d*), які позначають абсциси вершин трапеції, заданої функцією:

$$MF(x)=\begin{cases}\dfrac{b-x}{b-a}, a\le x\le b\\ 1, b\le x\le c\\ \dfrac{x-c}{d-c}, c\le x\le d\\ 0, \text{в інших випадках}\end{cases}. \qquad (25)$$

Функції належності для вихідної змінної *access* позначаються двома трапецевидними інтервалами *deny*, *allow* для точного визначення центру ваги, що позначає нечіткий висновок системи.

База знань для побудови даної нечіткої моделі складається з правил типу «якщо - то», усі вхідні змінні мають по три нечітких стани і ще один стан *none*, коли значення вхідної змінної не задане системою. Випадок, коли значення усіх вхідних змінних не задані, на практиці неможливий при функціонуванні системи IP-телефонії, тому кількість правил нечіткого висновку досліджуваної системи 4•4•4•4-1=255.

Нечіткий висновок підсистеми розподілу доступу, побудованого на основі заданих 255 правил з поточними значеннями вхідних та вихідної змінних, має вигляд, фрагмент якого представлений на рисунку 11.

На рисунку 11 зображено випадок, коли клієнт, рівень довіри якого низький, здійснює дзвінок у ранковий час в країну, яка має той самий часовий пояс, що і Україна, і якість зв'язку свідчить, що дзвінок здійснюється з локальної мережі. Висновок запропонованої нечіткої системи в даному випадку становить 0.48, що відповідає дозволу на здійснення даного дзвінка. Даний приклад підтверджує правильність роботи сформованої бази правил. Перевірка інших комбінацій значень вхідних змінних демонструє правильність роботи такої нечіткої системи. Запропонована нечітка система може реалізовуватися як програмно, так і апаратно у вигляді нечіткого контролера для успішного захисту від термінації трафіку.

Апробовано запропонований метод виявлення аномальних повідомлень у трафіку IP-телефонії був інтегрований до прототипу діючої корпоративної системи з метою оцінки його ефективності в умовах реального використання. Імплементація здійснювалася як додатковий модуль до підсистеми контролю трафіку, що дозво-

лило не лише зберегти існуючу архітектуру мережевої інфраструктури, але й забезпечити її розширення без значних витрат ресурсів або порушення стабільності основних сервісів.

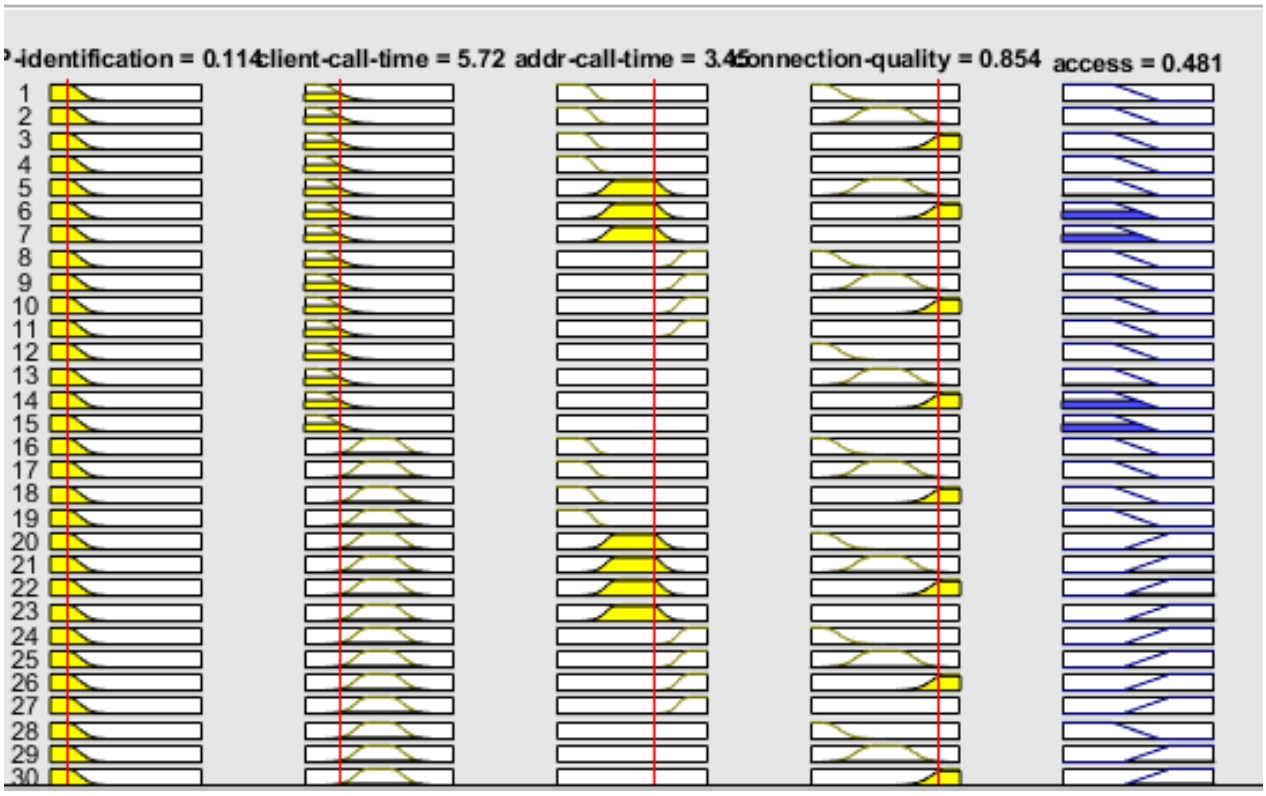

Рис. 11. Нечіткий висновок підсистеми розподілу доступу в IP-телефонії

Важливо відзначити, що завдяки модульному характеру розробленого рішення його можливо масштабувати для використання в системах різного рівня, від локальних корпоративних офісів до міжрегіональних вузлів операторів зв'язку. Таким чином, впровадження методу може здійснюватися поетапно, з мінімальним втручанням у поточні процеси функціонування інфраструктури IP-телефонії.

В процесі апробації запропонованих методів було здійснено інтеграцію даної інтелектуалізованої підсистеми в діючу корпоративну VoIP систему Західноукраїнського національного університету.

На рисунку 12 представлено сторінку, яка, окрім підсистеми аналізу ключових понять у VoIP, включає також модуль для перетворення голосових повідомлень в текстові представлення і адміністративний модуль, що використовується для прийняття рішень відносно користувачів системи та відповідне управління виявленими аномальними повідомленнями.

Дана інтеграція може також використовуватися і для виявлення голосових повідомлень, які здійснюють користувачі при дзвінках на платні сервіси, що також є дуже актуальним питанням в системі практичного використання IP-телефонії.

Загалом результати впровадження підтверджують, що запропонований метод є не лише теоретично обґрунтованим, але й практично застосовним. Його використання дозволяє значно підвищити рівень кібербезпеки в IP-телефонних системах за рахунок своєчасного виявлення аномалій, зменшення кількості хибнопозитивних спрацьовувань та зниження витрат часу на ручний аналіз мережевого трафіку.

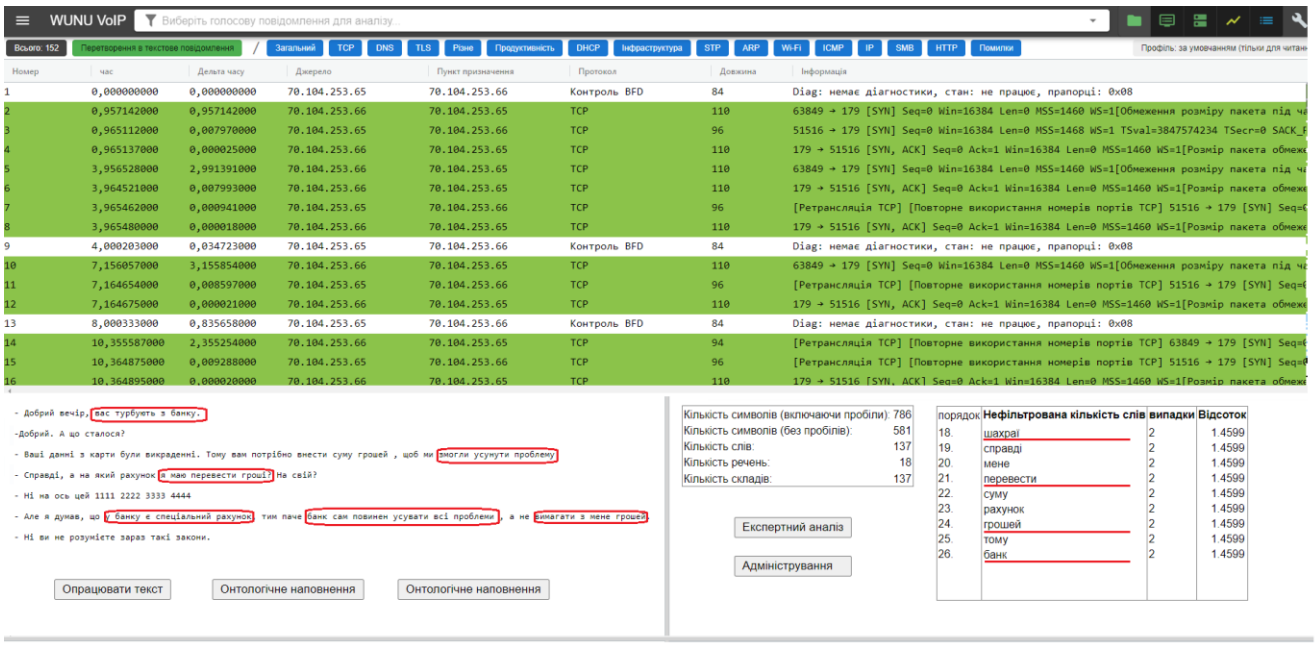

Рис. 12. Інтерфейс інтеграції підсистеми аналізу VoIP повідомлень в систему управління корпоративною IP-телефонією

З метою оцінювання ефективності розробленого методу виявлення аномальних повідомлень у трафіку IP-телефонії було проведено серію експериментів, що базуються на аналізі VoIP-повідомлень у режимі реального часу. Запропонований підхід передбачає групування тематично схожих SIP-повідомлень з використанням структурованої онтології домену IP-телефонії, яка автоматично оновлюється в процесі аналізу трафіку.

Для оцінки продуктивності системи було змодельовано ситуацію, в якій порівнюється швидкість обробки повідомлень оператором вручну та системою автоматичного виявлення.

При обробці 50 повідомлень система витратила 1185 секунд, у той час як оператор – 6720 секунд. Аналогічна тенденція спостерігалася при обробці 100 повідомлень: автоматизована система завершила процес за 1730 секунд, тоді як вручну оператору для цього було потрібно 14356 секунд. Отже, із зростанням кількості повідомлень різниця в часі між ручною та автоматизованою обробкою суттєво зростає, що свідчить про хорошу масштабованість запропонованого рішення, рисунок 13.

Проведені експериментальні дослідження підтвердили ефективність методу виявлення аномальних VoIP-повідомлень, заснованого на групуванні та онтологічному аналізі трафіку IP-телефонії. Результати показали суттєве зниження часу обробки повідомлень у порівнянні з ручною роботою оператора, при цьому ефективність системи зростає зі збільшенням обсягів даних.

Ключовим чинником, що забезпечує приріст продуктивності, є механізм автоматичного поповнення онтології, який дає змогу враховувати нові шаблони та семантичні характеристики аномальних повідомлень у корпоративній мережі. Таким чином, розроблена система демонструє високий потенціал до впровадження в інфраструктуру IP-телефонії для забезпечення оперативного виявлення аномалій, підвищення надійності комунікацій та зниження навантаження на персонал.

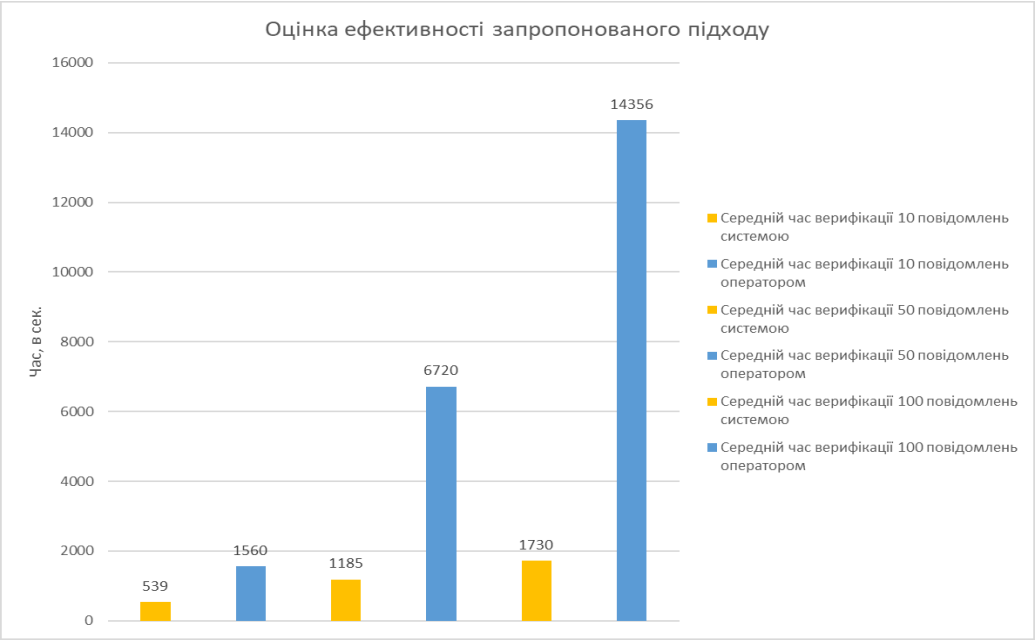

Рис. 13. Результат: скорочення часу опрацювання повідомлень в середньому від 2.89 до 8.3 рази

На основі аналізу узагальненої структури інтелектуалізованої комп'ютерної системи виявлення зловмисних повідомлень IP-телефонії в корпоративній мережі ICSDMM-VoIP (Intelligent Computer System for Detection of Malicious Messages), виконано порівняльну оцінку основних показників, які впливають на ефективність контролю нетипової поведінки користувачів IP-телефонії (таблиця 1).

Таблиця 1

Порівняння розробленої системи контролю нетипової поведінки користувачів IP-телефонії з відомими рішеннями

Техніко-економічні характеристики	Системи				
	Wireshark	Suricata	Snort	Asterisk	ICSDMM - VoIP
Аналітика та звітність	+/–	–	+/-	+/-	+
Багатопотокова обробка даних	+/–	+	–	+	+
Адаптивна інтеграція	+/–	–	+	+	+
Масштабованість	–	–	–	+	+
Інтелектуальний аналіз даних	–	–	–	–	+
Вартість	Freeware	Freeware	Shareware	Freeware	Freeware

Результати впровадження підтверджують, що використання розробленої інтелектуалізованої системи виявлення зловмисних повідомлень дозволяє суттєво підвищити рівень кібербезпеки в IP-телефонії за рахунок своєчасного виявлення аномалій, зменшення кількості хибних спрацьовувань та зниження витрат часу на рутинний аналіз мережевого трафіку.

ВИСНОВКИ

У дисертаційній роботі вирішено актуальну науково-прикладну задачу створення інтелектуалізованої комп'ютерної системи своєчасного виявлення зловмисних VoIP-повідомлень в корпоративній мережі з метою підвищення ефективності захисту IP-телефонії, що базується на онтологічній моделі її функціонування та методах інтелектуального аналізу даних.

Основні наукові та практичні результати дослідження дозволяють сформулювати наступні висновки:

1. Проведено аналіз сучасних методів та засобів захисту IP-телефонії у корпоративних мережах. Проведено систематизацію загроз, притаманних IP-телефонії, зокрема атак типу spoofing, sniffing, DoS/DDoS, маніпулювання сигналізаційним трафіком (SIP), термінацію VoIP-трафіку. Встановлено, що традиційні засоби безпеки, такі як брандмауери та VPN, є недостатньо ефективними проти складних, багаторівневих атак у VoIP-середовищі. Доведено доцільність застосування методів інтелектуального аналізу даних для підвищення ефективності захисту VoIP-інфраструктури.

2. Запропоновано онтологію опису VoIP-повідомлень у межах функціонування системи IP-телефонії. В межах онтологічного моделювання здійснено формалізацію ключових понять предметної області, які репрезентовано у вигляді окремих концептів. Кожному концепту присвоєно чітке визначення, атрибутивний склад та типові екземпляри, що забезпечує однозначність інтерпретації інформації як людиною, так і програмними агентами. Розроблена онтологія є гнучкою основою для подальшого розширення та адаптації під конкретні прикладні задачі, зокрема для виявлення аномалій у трафіку, забезпечення семантичного аналізу повідомлень, генерації сигнатур для систем захисту від VoIP-атак, а також інтеграції з інструментами інтелектуального аналізу даних у сфері мережевої безпеки. Її використання дозволяє підвищити рівень інтероперабельності, автоматизації та формалізації у процесах контролю, аудиту та реагування на події в IP-телефонії.

3. Запропоновано метод автоматизованого наповнення онтології тематичних повідомлень у корпоративній системі IP-телефонії, що є суттєвим внеском у підвищення ефективності побудови онтологічних баз знань для аналізу комунікаційного трафіку. Основою методу є формалізоване представлення VoIP-повідомлень у вигляді деревовидних структур, які моделюють логічну та ієрархічну організацію повідомлень із розподілом на рівні заголовків, полів, параметрів і вкладених елементів. Такий підхід дозволяє здійснювати семантичний аналіз та структуризацію повідомлень із врахуванням протоколів SIP, SDP, RTP та інших, що використовуються в IP-телефонії. Додатково, операції аналізу та взаємодії між структурними компонентами повідомлень описано за допомогою алгебри кортежів, яка забезпечує формальну основу для їх порівняння, об’єднання, фільтрації та трансформації у відповідні концепти онтології. Це дозволяє автоматизувати процес ідентифікації нових сутностей і зв’язків, які мають бути внесені до онтологічної моделі, що в свою чергу дає можливість зменшити час наповнення онтологічної бази знань.

4. Розроблено метод виявлення аномалій у трафіку IP-телефонії, заснований на групуванні VoIP-повідомлень з використанням онтологічного підходу та контекстно-частотного аналізу, що дозволяє значно підвищити точність і швидкість

виявлення нестандартної або шкідливої активності у корпоративних мережах зв’язку. Метод базується на інтеграції знання-орієнтованої моделі, у якій повідомлення описані за допомогою спеціалізованої онтології, що відображає типи повідомлень, їхні структури, допустимі послідовності взаємодії та семантичні зв’язки між учасниками сесії зв’язку. Така онтологічна база дозволяє порівнювати вхідні повідомлення з відомими шаблонами «нормальної» поведінки. Розроблений метод забезпечив підвищення ефективності виявлення аномальних повідомлень, у середньому 5, 6 разів, порівняно з традиційними підходами. Це досягнуто за рахунок глибшого урахування контексту комунікаційних подій, семантики повідомлень та гнучкого онтологічного представлення знань про мережеві взаємодії.

5. Розроблено підсистему для захисту від термінації трафіку в IP-телефонії на основі використання нечіткої логіки. Підсистема здатна гнучко адаптуватися до змінних умов і приймати рішення на основі нечітких даних. Це дозволяє враховувати численні критерії, такі як IP-адреса клієнта, час здійснення дзвінка, часовий пояс і якість зв’язку, що сприяє підвищенню рівня безпеки та якості зв'язку.

6. Розроблено архітектуру, структурно-функціональні схеми та узагальнений алгоритм функціонування інтелектуалізованої системи виявлення аномалій у трафіку IP-телефонії, яка об’єднує окремі підсистеми в єдиний цілісний інформаційно-аналітичний контур. Система охоплює всі етапи обробки VoIP-трафіку — від первинного збору даних до прийняття рішень щодо наявності аномальних або потенційно загрозливих повідомлень. Завдяки такій архітектурі забезпечено модульність, гнучкість і масштабованість розробленого рішення, а також можливість інтеграції з зовнішніми системами інформаційної безпеки, SIEM-платформами та системами моніторингу.

7. Проведено експериментальні дослідження запропонованих методів і засобів для виявлення аномалій у трафіку IP-телефонії, що дозволило кількісно оцінити їх ефективність та практичну придатність. З метою перевірки працездатності розроблених методів імплементовано програмну підсистему виявлення аномальних VoIP-повідомлень, яка інтегрована в університетську корпоративну систему IP-телефонії. Програмна підсистема реалізує компоненти онтологічного аналізу повідомлень, модуль контекстно-частотної обробки, інтерфейси інтеграції з потоковими джерелами SIP/SDP/RTP-трафіку, а також механізми прийняття рішень щодо ідентифікації аномальних подій. Це дозволило здійснювати обробку повідомлень у режимі, наближеному до реального часу, що критично важливо для виявлення атак типу DoS, SIP flooding, spoofing, несанкціонованого доступу та інших загроз.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. A. Melnyk, R. Shevchuk, I. Romanets, I. Yakymenko, S. Voznyak, and V. Luchyk, "A Method of Detecting Anomalies in IP Phone Traffic based on Ontology of Voip Messages," *2024 14th International Conference on Advanced Computer Information Technologies (ACIT)*, Ceske Budejovice, Czech Republic, 2024, pp. 485-489, DOI: 10.1109/ACIT62333.2024.10712505.
2. Balyk, A., Karpinski, M., Naglik, A., Shangytbayeva, G., & Romanets, I. (2017). Using Graphic Network Simulator 3 For DDoS Attacks Simulation. International Journal of Computing, 16(4), 219-225 DOI: 10.47839/ijc.16.4.910
3. Todd J.B. Blayone, Olena Mykhailenko, Svetlana Usca, Ihor Romanets, Exploring technology attitudes and personal–cultural orientations as student readiness factors for digitalised work. *Higher Education, Skills and Work-based Learning* 17 June 2021; 11 (3): 649–671. DOI:10.1108/HESWBL-03-2020-0041
4. Романець І. Захист від термінації трафіку в IP-мережі на основі нечіткої логіки: Науковий журнал "Вимірювальна та обчислювальна техніка в технологічних процесах"(1), Хмельницький національний університет, Хмельницький, 2024 с. 186-193, DOI: 10.31891/2219-9365-2024-77-23.
5. І. Романець, «Онтологічний підхід в системі забезпечення безпеки використання IP-телефонії», *Опт-ел. інф-енерг. техн.*, вип. 47, вип. 1, с. 240–252, Чер. 2024. DOI: 10.31649/1681-7893-2024-47-1-240-252.
6. І. Романець, "Архітектура інтелектуалізованої системи забезпечення безпеки використання IP-телефонії в корпоративній мережі", *Вісник Хмельницького національного університету. Серія: Технічні науки*, vol. 337, no. 3(2), pp. 425–436, May 2024, DOI: 10.31891/2307-5732-2024-337-3-65.
7. Markowsky, G., Sachenko, A., Voznyak, S., Spilchuk, V., Romanyak, R., Turchenko, V. and Romanets, I. 2014. The Ternopil Educational Communication Center: A Nato Project To Integrate Regional Information Technology Resources. *International Journal of Computing*. 7, 1 (Aug. 2014), 185-190. DOI:10.47839/ijc.7.1.503.
8. I. Romanets, A. Sachenko and L. Dubchak, "Method of Protection Against Traffic Termination in VoIP," 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Iasi, Romania, 2018, pp. 1-5, DOI: 10.1109/ECAI.2018.8678992.
9. Blayone, T., O. Mikhailenko, S. Ušča, A. Abuže, Ihor Romanets, and Mykhailo Oleksiiv. "Exploring the dispositional readiness of Latvian and Ukrainian university learners for digitalised work toward Industry 4.0." *Preprint* (2019).
10. M. Dyvak, A. Pukas, A. Melnyk, I. Voytyuk, S. Valchyshyn, and I. Romanets, "Software Architecture for Modeling the Interval Static and Dynamic Objects," 2021 11th International Conference on Advanced Computer Information Technologies (ACIT), Deggendorf, Germany, 2021, pp. 572-575, DOI: 10.1109/ ACIT52158.2021.9548577.
11. M. Dyvak, A. Melnyk, L. Dostalek, V. Ostroverkhov, L. Honchar, and I. Romanets, "Repository of Interval Models of Dynamics of Concentrations of Harmful Emissions of Motor Transport," 2022 12th International Conference on Advanced Computer Information Technologies (ACIT), Ruzomberok, Slovakia, 2022, pp. 89-94, DOI: 10.1109/ACIT54803.2022.9913123.

12. O. Kochan, O. Osolinskyi, A. Sachenko, V. Kochan, and I. Romanets, "Simulator of Microcontroller's Power Consumption," 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Dortmund, Germany, 2023, pp. 787-792, DOI: 10.1109/IDAACS58523.2023.10348884.
13. V. Tymchyshyn, B. Tymchyshyn, A. Melnyk, V. Manzhula, V. Faifura, and I. Romanets, "The System Architecture of the Software for Modeling Harmful Emissions in Soil," 2023 13th International Conference on Advanced Computer Information Technologies (ACIT), Wrocław, Poland, 2023, pp. 58-62, DOI: 10.1109/ACIT58437.2023.10275416.
14. A. Sartiukova, O. Markiv, V. Vysotska, I. Shakleina, N. Sokulska, and I. Romanets, "Remote Voice Control of Computer Based on Convolutional Neural Network," 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Dortmund, Germany, 2023, pp. 1058–1064, DOI: 10.1109/IDAACS58523. 2023.10348808.

АНОТАЦІЯ

Романець І.Є. Інтелектуалізована комп'ютерна система виявлення зловмисних повідомлень IP-телефонії в корпоративній мережі – Рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.05 – Комп’ютерні системи та компоненти». – Західноукраїнський національний університет, Тернопіль, 2025.

Дисертаційна робота присвячена актуальній науково-технічній задачі підвищення рівня захищеності корпоративних мереж шляхом розробки та дослідження інтелектуалізованої комп'ютерної системи, призначеної для виявлення зловмисних повідомлень в IP-телефонії.

В роботі проаналізовано сучасні загрози та вразливості систем IP-телефонії в корпоративному середовищі, такі як спам через інтернет-телефонію, шахрайство з оплатою дзвінків та атаки на відмову в обслуговуванні (DoS). Виявлено недостатню ефективність існуючих методів та засобів захисту перед новими та адаптивними атаками, що зумовлює необхідність застосування інтелектуальних підходів.

Запропоновано архітектуру інтелектуалізованої комп'ютерної системи, яка використовує методи машинного навчання для аналізу трафіку (SIP-повідомлень) та метаданих медіа-потоків (RTP/RTCP) IP-телефонії. Розроблено моделі та алгоритми для виділення інформативних ознак зловмисних повідомлень, класифікації трафіку та виявлення аномалій в реальному часі. Особливу увагу приділено здатності системи до самонавчання та адаптації до нових, раніше невідомих типів загроз.

Досліджено ефективність різних алгоритмів машинного навчання для вирішення задачі класифікації та виявлення аномалій в IP-телефонії. Проведено експериментальну перевірку розробленої системи на реальних та змодельованих наборах даних, що підтвердило її здатність виявляти зловмисні повідомлення з високою точністю та низьким рівнем помилкових спрацювань.

Наукова новизна роботи полягає у розробці комплексного підходу до виявлення зловмисних повідомлень в IP-телефонії на основі інтелектуального аналізу багатовимірних даних, що враховує специфіку корпоративних мереж, а також у

вдосконаленні методів адаптивного виявлення загроз на основі онтологічного підходу та нечіткої логіки.

Практичне значення отриманих результатів полягає у розробленому програмному прототипі системи, рекомендаціях щодо її впровадження та налаштування в корпоративних мережах для підвищення рівня безпеки IP-телефонії, запобігання фінансовим втратам та захисту конфіденційної інформації.

Ключові слова: IP-телефонія, VoIP, зловмисні повідомлення, вішинг, корпоративна мережа, інтелектуалізована система, машинне навчання, виявлення аномалій, кібербезпека, захист інформації.

ANNOTATION

Romanets I.Y. Intelligent computer system for detecting malicious IP-telephony messages in a corporate network – Manuscript.

Thesis for a Ph.D. degree by specialty 05.13.05 – Computer Systems and Components – West Ukrainian National University, Ternopil, 2025.

The dissertation is devoted to the pressing scientific and technical challenge of enhancing the security level of corporate networks by developing and researching an intellectualized computer system designed to detect malicious messages in IP telephony.

The work analyzes modern threats and vulnerabilities in IP telephony systems within corporate environments, such as spam over internet telephony (SPIT), call payment fraud, and denial-of-service (DoS) attacks. The insufficient effectiveness of existing protection methods and tools against new and adaptive attacks has been identified, necessitating the application of intelligent approaches.

An architecture for an intellectualized computer system is proposed, utilizing machine learning methods to analyze traffic (e.g., SIP messages) and metadata of media streams (RTP/RTCP) in VoIP. Models and algorithms have been developed to extract informative features of malicious messages, classify traffic, and detect anomalies in real time. Special attention is given to the system's ability to self-learn and adapt to new, previously unknown types of threats.

The effectiveness of various machine learning algorithms for solving the tasks of classification and anomaly detection in VoIP has been investigated. Experimental testing of the developed system on real and simulated datasets confirmed its ability to detect malicious messages with high accuracy and a low false-positive rate.

The scientific novelty of the work lies in the development of a comprehensive approach to detecting malicious messages in VoIP based on intelligent analysis of multidimensional data, considering the specifics of corporate networks, as well as in the improvement of adaptive threat detection methods using ontological approaches and fuzzy logic.

The practical significance of the obtained results is reflected in the developed software prototype of the system, recommendations for its implementation, and configuration in corporate networks to enhance IP-telephony security, prevent financial losses, and protect confidential information.

Keywords: IP-telephony, VoIP, malicious messages, vishing, corporate network, intellectualized system, machine learning, anomaly detection, cybersecurity, information protection.

Підписано до друку 24.07.2025 р.
Формат 60х90/16. Гарнітура Times.
Папір офсетний. Друк на дуплікаторі.
Умов. друк. арк. 0,9. Обл.-вид. арк. 1,0.
Зам. № А001-25. Тираж 100 прим.

Видавець та виготовлювач
Західноукраїнський національний університет
вул. Львівська, 11, м. Тернопіль 46009

Свідоцтво про внесення суб'єкта видавничої справи
до Державного реєстру видавців ДК № 3467 від 23.04.2009 р.

Видавничо-поліграфічний центр «Університетська думка»
вул. Сергія Короля, 2, м. Тернопіль, 46009
тел. (0352) 51-75-72
E-mail: edition@wunu.edu.ua

Дисертацією є рукопис.

Робота виконана у Західноукраїнському національному університеті Міністерства освіти і науки України.

Науковий керівник: доктор технічних наук, професор
САЧЕНКО Анатолій Олексійович,
Західноукраїнський національний університет,
(м. Тернопіль),
професор кафедри інформаційно-обчислювальних систем і управління.

Офіційні опоненти: доктор технічних наук, професор,
ОПАНАСЕНКО Володимир Миколайович
Інститут кібернетики ім. В.М. Глушкова НАН України,
(м. Київ),
Завідувач відділу мікропроцесорної техніки №205

кандидат технічних наук, доцент,
КЛЬОЦ Юрій Павлович
Хмельницький національний університет (м. Хмельницький),
завідувач кафедри кібербезпеки

Захист відбудеться **28 серпня 2025 р. о $14^{\underline{00}}$** годині на засіданні спеціалізованої вченої ради Д 58.082.02 у Західноукраїнському національному університеті за адресою: 46009, м. Тернопіль, вул. Львівська, 11А (корпус 11, зал засідань вченої ради).

З дисертацією можна ознайомитись у бібліотеці Західноукраїнського національного університету за адресою: 46009, м. Тернопіль, вул. Сергія. Короля, 4.

Вчений секретар
спеціалізованої вченої ради Д 58.082.02
кандидат технічних наук, доцент Леся ДУБЧАК

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Сучасні корпоративні мережі все частіше використовують IP-телефонію (VoIP) як ефективний засіб комунікації через її гнучкість, масштабованість та економічні переваги. Однак зростання популярності VoIP супроводжується збільшенням кількості кіберзагроз, зокрема зловмисних повідомлень (спам, фішинг, атаки на відмову в обслуговуванні тощо), що становлять серйозну небезпеку для конфіденційності, цілісності та доступності комунікаційних сервісів.

IP-телефонія забезпечує гнучкість, масштабованість та економічність корпоративного зв'язку, дозволяє інтегрувати голосові сервіси з іншими інформаційними системами, включаючи CRM, ERP та інструменти управління знаннями.

Однак, разом із перевагами, IP-телефонія стає привабливою ціллю для зловмисників, що обумовлює виникнення нових загроз, зокрема – розсилання зловмисних SIP-повідомлень, фішингу, спуфінгу, DoS-атак, перехоплення конфіденційної інформації.

Значний внесок у сферу дослідження безпеки IP-телефонії (VoIP) та виявлення зловмисних повідомлень внесли зарубіжні та вітчизняні вчені, список яких наведено у дисертаційні роботі.

У контексті сучасних викликів, зумовлених гібридними загрозами та цифровими трансформаціями, питання підвищення ефективності виявлення зловмисних повідомлень в IP-телефонії набуває особливої актуальності. Це особливо важливо для корпоративного середовища, де забезпечення інформаційної безпеки прямо впливає на стійкість бізнесу, збереження критичних даних та довіру клієнтів.

На сьогодні існуючі методи, технічні засоби та програмні рішення не дозволяють своєчасно виявити та локалізувати зловмисну активність у VoIP-мережах. Більшість з них не враховують динаміку змін у поведінкових шаблонах трафіку, мають низьку адаптивність до нових типів загроз і не здатні до самонавчання. Це вимагає впровадження нових підходів, що базуються на використанні технологій штучного інтелекту, машинного навчання, експертних систем та аналізу аномалій.

У зв'язку з цим актуальним є науково обґрунтоване розв'язання задачі підвищення ефективності захисту IP-телефонії із врахуванням своєчасності виявлення зловмисних VoIP-повідомлень шляхом створення інтелектуалізованої комп'ютерної системи, що відповідає сучасним тенденціям у сфері кібербезпеки та цифрової трансформації.

Розв'язання цієї задачі сприятиме підвищенню рівня захищеності корпоративних інформаційних систем, забезпеченню безперервності бізнес-процесів, зниженню ризиків фінансових та репутаційних втрат.

Зв'язок роботи з науковими програмами, планами, темами. Науково-дослідна робота за темою дисертації проводилася у відповідності з координаційним планом науково-дослідних робіт і науково-технічних програм Міністерства освіти і науки України, виконанні фундаментальної держбюджетної роботи «Математичне та комп'ютерне моделювання об'єктів з розподіленими параметрами на основі поєднання онтологічного та інтервального аналізу» (держреєстраційний номер 0122U001497).

Мета і задачі дослідження. Метою роботи є підвищення ефективності захисту IP-телефонії у корпоративній мережі шляхом створення інтелектуалізованої комп'ютерної системи виявлення зловмисних повідомлень, що базується на онтологічній моделі її функціонування та методах інтелектуального аналізу даних.

Для досягнення цієї мети у дисертаційній роботі поставлено наступні завдання:

1) аналіз сучасних методів та засобів захисту IP-телефонії у корпоративній мережі, які включають компоненти інтелектуального аналізу даних;
2) розробити онтологічну модель функціонування IP-телефонії, яка включатиме формалізацію основних понять та зв’язків між ними;
3) розробити метод автоматизованого наповнення онтології тематичних повідомлень в корпоративній мережі;
4) розробити метод виявлення аномалій в трафіку IP-телефонії на основі групування повідомлень;
5) розробити засіб захисту від термінації трафіку в IP-телефонії на основі використані нечіткої логіки, що сприятиме підвищенню рівня безпеки та якості зв'язку.
6) розробити архітектуру та структурно-функціональні схеми базових підсистем інтелектуалізованої системи;
7) здійснити експериментальні дослідження розроблених методів та засоби виявлення аномальних повідомлень у IP-телефонії корпоративної мережі.

Об’єкт досліджень: процеси захисту IP-телефонії з використанням інтелектуальних методів аналізу даних.

Предмет досліджень: методи та засоби захисту IP-телефонії в корпоративній мережі.

Методи дослідження. Для вирішення поставлених задач у дисертаційній роботі використовувалися такі методи: методи системного аналізу та огляду літератури – методи штучного інтелекту, зокрема онтологічного інжинірингу та представлення знань; методи машинного навчання, зокрема кластерного аналізу (групування) та виявлення аномалій; методи об'єктно-орієнтованого проектування та програмування; методи експериментальних досліджень та статистичної обробки даних.

Наукова новизна отриманих результатів.

Основні наукові результати, висунуті на захист.

Вперше розроблено:

– онтологічну модель опису VoIP повідомлень в системі IP-телефонії, яка включає формалізацію основних понять у формі окремих концептів та опис зв’язків між цими поняттями, що у сукупності створило можливість розробки нових, більш ефективних методів виявлення аномалій в трафіку IP-телефонії;
– метод виявлення аномалій в трафіку IP-телефонії, який, ґрунтується на поєднані методів групування VoIP-повідомлень та контекстно-частотного аналізу, що у сукупності дозволило підвищити ефективність виявлення аномальних повідомлень.

Удосконалено:

- підсистему захисту від термінації трафіку в IP-телефонії, яка відрізняється від відомих тим, що здатна гнучко адаптуватися до змінних умов і приймати рішення на основі нечітких даних, що дозволяє враховувати численні критерії, такі як IP-адреса клієнта, час здійснення дзвінка, часовий пояс і якість зв'язку, а також сприяє підвищенню рівня безпеки та якості зв'язку.
- архітектуру системи виявлення зловмисних повідомлень IP-телефонії в корпоративній мережі, яка відрізняється від відомих наявністю блоків і процедур виявлення аномальних повідомлень в трафіку IP-телефонії на основі онтологічного підходу та розподілу доступу в режимі реального часу з використанням нечіткої логіки, що дозволило спростити функції адміністратора системи та зменшити час оброблення повідомлень.

Практичне значення отриманих результатів полягає в тому, що за результатами проведених автором теоретичних досліджень:

– розроблено систему, яка демонструє підвищення рівня виявлення зловмисних повідомлень при зниженні кількості хибних спрацьовувань в корпоративній мережі, що забезпечує значне зменшення фінансових втрат від VoIP-шахрайства та підвищення надійності зв'язку;

– створено прототип програмного рішення для Asterisk із відкритим інтерфейсом налаштування правил реагування, що спрощує його впровадження та адаптацію під потреби підприємств різного масштабу;

– запропоновані методи нечіткого виведення та онтологічного аналізу можуть бути використані як окремі модулі в існуючих системах безпеки мережі, що підвищує масштабованість та гнучкість рішення;

– результати дослідження можуть лягти в основу технічних рекомендацій із захисту IP-телефонії для IT-відділів корпоративних замовників, а також можуть бути враховані при формуванні політик інформаційної безпеки в галузевих стандартах.

Теоретичні та практичні результати роботи використані у ПП «Колумбус», компанії «Infineon Technologies AG» та у Тернопільській обласній військовій адміністрації в управління цифрової трансформації для захисту від несанкціонованого доступу до інформації, а також у навчальному процесі при викладанні дисциплін "Комп'ютерні системи", «Комп'ютерні мережі», «Програмно-апаратні засоби захисту інформації».

Особистий внесок здобувача. Усі наукові результати, викладені у дисертаційній роботі, отримані автором особисто і повністю розкрито у публікаціях. У друкованих працях, опублікованих у співавторстві, автору належить: метод виявлення аномалій в трафіку IP-телефонії за допомогою групування VoIP повідомлень на основі контекстно-частотного аналізу; метод автоматизованого наповнення онтології тематичних повідомлень у корпоративній системі IP-телефонії, який ґрунтується на формалізованому представленні повідомлень за допомогою деревовидних структур та на описі операцій взаємодії засобами алгебри кортежів; запропоновано онтологію для опису повідомлень у системах IP-телефонії, яка формалізує основні концепції як окремі сутності та деталізує зв'язки між ними.

Апробація результатів дисертації. Основні положення та результати дисертаційної роботи доповідались та обговорювались на 7 міжнародних науково-практичних конференціях, зокрема:

1. IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications – IDAACS (м. Бухарест, Румунія, 2017 р.);
2. 10th International Conference on Electronics, Computers and Artificial Intelligence – ECAI (м. Яси, Румунія, 2018 р.);
3. 11th International Conference on Advanced Computer Information Technologies – ACIT (м. Деггендорф, Німеччина, 2021 р.);
4. 12th International Conference on Advanced Computer Information Technologies – ACIT (м. Руженберок, Словаччина, 2022 р.);
5. IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications – IDAACS (м. Дортмунд, Німеччина, 2023 р.);
6. 13th International Conference on Advanced Computer Information Technologies – ACIT (м. Вроцлав, Польща, 2023 р.);
7. 14th International Conference on Advanced Computer Information Technologies – ACIT (м. Чеське Будейовіце, Чехія, 2024 р.).

Публікації. За темою дисертації опубліковано 14 наукових праць, з них: 3 одноосібних статті у наукових фахових виданнях України, рекомендованих МОН України, 4 – у виданнях, що індексуються в міжнародних наукометричних базах Scopus/Web of Science, 7 тез доповідей на міжнародних конференціях, що індексуються в міжнародних наукометричних базах Scopus/Web of Science.

Структура та обсяг роботи. Дисертація складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Робота містить 138 сторінок основного тексту. Загальний об’єм дисертації – 170 сторінок, 1 таблиця, 68 рисунків, 4 додатки, 112 найменування використаних джерел.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** наведено загальну характеристику роботи, обґрунтовано актуальність теми, визначено об’єкт та предмет дослідження, сформульовані мета і задачі розкрито наукову та практичну цінність отриманих результатів, а також подані відомості про апробацію дисертаційної роботи.

У **першому розділі** досліджено структуру IP-телефонії, а також її особливості, проведено аналіз існуючих методів і засобів виявлення зловмисних повідомлень IP-телефонії, досліджено методи штучного інтелекту для виявлення зловмисних повідомлень, методи виявлення нетипової поведінки при передачі голосу. Проведено всебічний аналіз сучасного стану та перспектив розвитку систем виявлення зловмисних повідомлень в IP-телефонії. Розглянуто ключові аспекти, що стосуються структури IP-телефонії та її особливостей, які створюють як переваги, так і нові виклики у сфері кібербезпеки. Визначено, що попри значні економічні та функціональні вигоди, притаманні IP-телефонії, її інтеграція з мережею Інтернет робить її вразливою до широкого спектру кібератак, включаючи DoS/DDoS, спуфінг, фішинг

та інші зловмисні дії, спрямовані на порушення конфіденційності, цілісності та доступності голосових комунікацій.

Проведено аналіз існуючих методів і засобів виявлення зловмисних повідомлень IP-телефонії. Детально розглянуто застосування методів штучного інтелекту, таких як машинне навчання (зокрема, методи класифікації та кластеризації) та глибоке навчання, які демонструють значний потенціал у розпізнаванні складних патернів аномальної поведінки та зловмисних повідомлень. Окрему увагу приділено методам виявлення нетипової поведінки при передачі голосу, що дозволяє ідентифікувати загрози, які не завжди підпадають під відомі сигнатури атак, базуючись на відхиленнях від нормального трафіку та поведінкових шаблонів користувачів. Виявлено, що, попри значні досягнення, існуючі підходи часто стикаються з проблемами масштабованості, високого рівня хибно позитивних спрацьовувань та складності адаптації до нових типів атак.

Визначено перспективні шляхи розвитку предметної області та здійснено постановку задачі дослідження. Встановлено, що подальші дослідження повинні бути зосереджені на розробці та вдосконаленні гібридних методів, які комбінують сигнатурний аналіз з методами штучного інтелекту та аналізом поведінки для підвищення ефективності виявлення зловмисних повідомлень. Особлива увага має бути приділена розробці адаптивних систем, здатних до самонавчання та оперативного реагування на динамічно змінювані загрози. Це дозволить підвищити рівень безпеки та стійкості систем IP-телефонії до сучасних кіберзагроз.

У **другому розділі** обґрунтовано доцільність використання методів аналізу мовлення та відповідної обробки природної мови, які дозволяють створювати більш точні та ефективні системи виявлення аномального трафіку та потенційно небезпечних комунікацій в системі VoIP, виділено напрямок використання інтелектуальних засобів для аналізу контенту в системі VoIP. Запропоновано онтологію опису VoIP повідомлень в системі IP-телефонії, здійснено формалізацію основних понять у формі окремих концептів та описано зв’язки між цими поняттями, запропоновано метод виявлення аномалій в трафіку IP-телефонії на основі групування VoIP повідомлень, який, на відміну від існуючих, ґрунтується на поєднанні знання-орієнтованого підходу з використанням онтології та контекстно-частотного аналізу, запропоновано метод автоматизованого наповнення онтології тематичних повідомлень в корпоративній системі IP-телефонії, який ґрунтується на формалізованому представленні повідомлень за допомогою деревовидних структур та на описі операцій взаємодії засобами алгебри кортежів.

Завдяки постійному розвитку мережевих технологій та протоколів передачі даних, які використовуються в VoIP, якість голосу та зв'язку постійно покращується, що безпосередньо впливає на якість обслуговування.

Сучасні VoIP-системи надають широкий спектр додаткових функцій, таких як конференц-зв'язок, зміна голосу, відеодзвінки, інтеграція з CRM-системами.

Зростання мобільності відбувається завдяки використанню мобільних додатків та можливості робити виклики через Інтернет з будь-якого місця. Це дозволяє мобільним користувачам мати доступ до VoIP-зв'язку незалежно від їхнього місця перебування.

Інтеграція з іншими технологіями – VoIP стає складовою частиною інтегрованих комунікаційних рішень, які включають в себе електронну пошту, чат, відеоконференції та інші засоби спілкування.

Особливо необхідно відзначити такий напрям, як забезпечення безпеки. Розвиток технологій шифрування, аутентифікації та контролю доступу допомагає забезпечити безпеку VoIP-зв'язку та захистити дані користувачів від несанкціонованого доступу та перехоплення. Підсумовуючи, необхідно відзначити, що розвиток VoIP відбувається швидкими темпами, відкриваючи нові можливості для спілкування та покращення бізнес-комунікацій, а напрямок забезпечення безпеки VoIP є особливо актуальним в сучасних умовах.

Оскільки забезпечення безпеки функціонування VoIP є одним із ключових аспектів, які безпосередньо впливають на якість IP-телефонії та розширення сфер її використання, то проаналізуємо більш детально відомі методи та засоби, які використовуються для захисту інформації у VoIP (рисунок 1).

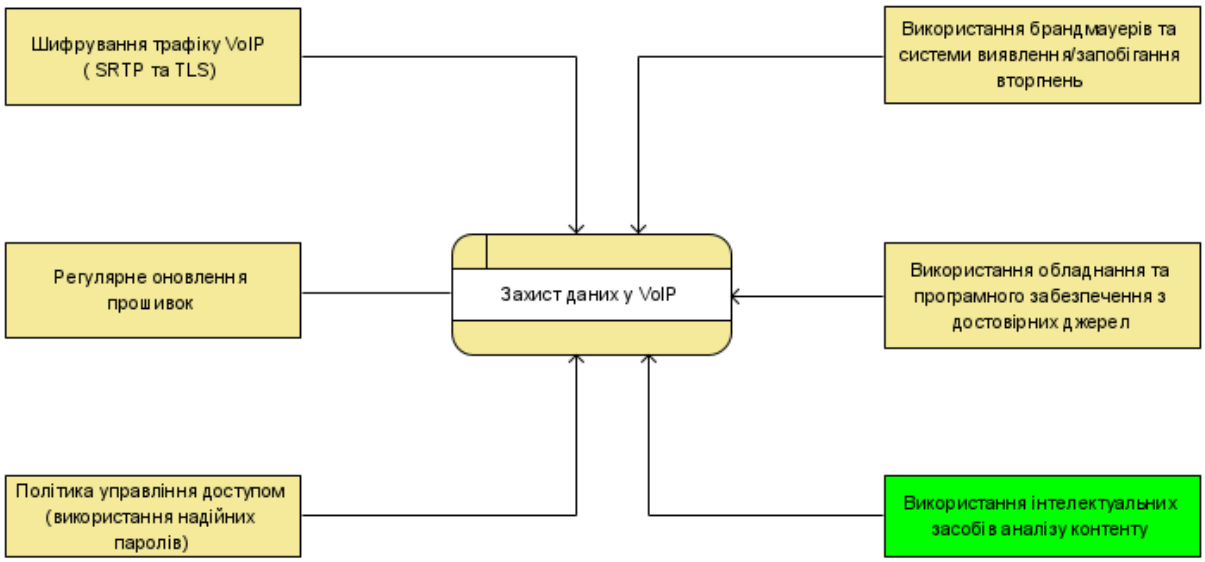

Рис. 1. Підходи до захисту інформації в системі VoIP

Системи VoIP вразливі до різних загроз безпеки, включаючи підслуховування, фішинг і атаки на відмову в обслуговуванні (DoS). Відомими підходами до зміцнення безпеки в IP-телефонії є наступні:

- шифрування трафіку VoIP за допомогою різних протоколів, таких як безпечний транспортний протокол реального часу (SRTP) і безпечний протокол транспортного рівня (TLS);
- ефективна політика управління доступом, що включає використання надійних паролей та контроль доступу, щоб запобігти несанкціонованому проникненню;
- регулярне оновлення прошивок та програмного забезпечення для виправлення вразливостей безпеки;
- використання брандмауерів та систем виявлення/запобігання вторгненням для захисту від кіберзагроз;
- використання обладнання та програмного забезпечення VoIP від перевірених постачальників, які дотримуються відповідних стандартів та протоколів;
- використання інтелектуальних засобів аналізу контенту.

З постійним розвитком технологій штучного інтелекту, особливо цікавим є напрямок використання інтелектуальних засобів для аналізу контенту в системі VoIP. Розглянемо основні сучасні підходи, які можна використовувати для аналізу VoIP контенту (рисунок 2).

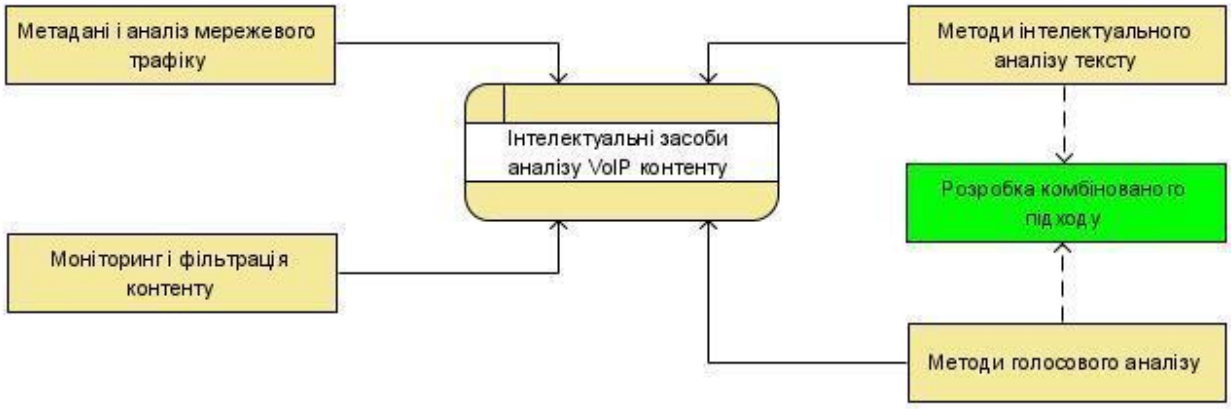

Рис. 2. Підходи до інтелектуального аналізу контенту в системі VoIP

Використання алгоритмів машинного навчання та обробки природної мови (NLP) для сканування тексту VoIP-повідомлень на ключові слова, фрази або семантичні зв'язки дозволяють знаходити інформацію, яка носить зловмисний характер. Використання методів та спеціалізованого програмного забезпечення для голосового аналізу допомагає визначати специфічні звуки, мовні ознаки або відтінки, що характеризують пропаганду чи загрози. Дослідження метаданих VoIP-повідомлень, таких як IP-адреси, час і тривалість дзвінка, здійснюється для виявлення аномалій або звичайних патернів, що можуть вказувати на можливу злочинну діяльність.

Важливо відзначити, що ефективність аналізу VoIP-повідомлень (як голосових, так і текстових) на вміст злочинної інформації може залежати від комбінації цих методів та технологій, а також від постійного вдосконалення і адаптації до нових загроз.

Створення онтології функціонування VoIP вимагає співпраці між експертами в предметній області, інженерами та дослідниками, які знайомі з технологіями, стандартами та найкращими підходами до організації VoIP. Така онтологія сприятиме полегшенню взаємодії, семантичній інтеграції та допомагатиме в системному проектуванні, управлінні, виправленні несправностей, а також сприятиме покращенню системи захисту інформації в VoIP.

Онтологія даної предметної області формується на основі структурованого представлення концептів, сутностей, зв’язків у сфері технології VoIP. Вона дозволить отримати формалізовану та семантично насичену структуру для розуміння та представлення знань про протоколи, компоненти та їх взаємодію.

Онтологія VoIP повинна включати:

- поняття та сутності: стек протоколів VoIP, представлення багаторівневої архітектури протоколів VoIP, включаючи такі протоколи, як SIP (протокол ініціації сеансу), RTP (транспортний протокол реального часу), SDP (протокол опису сеансу) та інші; компоненти VoIP: об’єкти, залучені до зв’язку VoIP, такі як користувацькі додатки (мобільні телефони з додатками, IP-телефони), шлюзи VoIP, проксі-

сервери, реєстратори та медіа-сервери; опис життєвого циклу VoIP виклику: представлення різних етапів VoIP виклику, включаючи встановлення виклику, узгодження, розрив виклику та систему управління викликом;

- відношення: комунікаційні зв'язки між об'єктами VoIP, такі як зв'язок користувач-користувач, зв'язок користувач-сервер, зв'язок сервер-сервер і зв'язок медіа-потоку; відношення залежностей між компонентами та протоколами VoIP, такі як залежність SIP від базових транспортних протоколів, таких як UDP або TCP, і залежність RTP від RTCP для зворотного зв'язку та управління;
- властивості та атрибути: властивості протоколу VoIP, такі як формати повідомлень, заголовки, методи, параметри та коди стану, визначені SIP, RTP та іншими відповідними протоколами; атрибути компонентів VoIP, наприклад IP-адреси, порти, кодеки, підтримувані функції, можливості та конфігурації; події, пов'язані зі зв'язком VoIP, такі як ініціювання виклику, завершення виклику, утримання/відновлення виклику, переадресація виклику, переадресація виклику та події медіапотоку; дії, які виконують об'єкти VoIP у відповідь на події, такі як надсилання SIP-повідомлень, узгодження параметрів медіа, встановлення та розрив медіа-сеансів і обробка сигналів управління викликами;
- ієрархічна таксономія: класифікація об'єктів і протоколів VoIP в ієрархічні таксономії на основі їхніх ролей, функцій і зв'язків; категоризація технологій VoIP на основі моделей розгортання (локальні, хмарні), мережевих архітектур (однорангові, клієнт-сервер) і сценаріїв використання (корпоративні VoIP, мобільний VoIP);
- семантичні обмеження: обмеження на зв'язки та взаємодію між об'єктами VoIP, що забезпечують узгодженість в онтології; правила, що регулюють поведінку та використання протоколів VoIP, включаючи правила обробки повідомлень, встановлення сеансу та узгодження медіа для окремих протоколів.

При побудові онтології опису VoIP повідомлень в системі IP-телефонії необхідно здійснити формалізацію основних понять у формі окремих концептів та описати зв'язки між цими поняттями. Розробка онтології для повідомлень VoIP передбачає створення структурованого представлення концептів, сутностей та зв'язків у межах відповідного домену. На рисунку 3 представлено онтологічний граф для формалізованого опису VoIP повідомлень.

На першому етапі здійснюємо ідентифікацію домену, яка включає визначення області та меж описуваної онтології. У нашому випадку це буде зв'язок за протоколом передачі голосових повідомлень через Інтернет , що охоплює концепції, пов'язані з передачею голосових даних через IP-мережі.

На другому етапі визначаємо основні поняття та сутності, які пов'язані з VoIP повідомленнями:

- протоколи VoIP (SIP, RTP, RTCP);
- компоненти VoIP (клієнт VoIP, сервер VoIP, шлюз VoIP);
- типи повідомлень VoIP (пакети INVITE, ACK, BYE, RTP);
- параметри VoIP (IP-адреса клієнта, IP-адреса призначення, номери портів, тип корисного навантаження);
- повідомлення (голосові та текстові), текстові повідомлення характеризуються основної частиною, автором, переліком ключових слів, оцінкою та комбінаційним представленням.

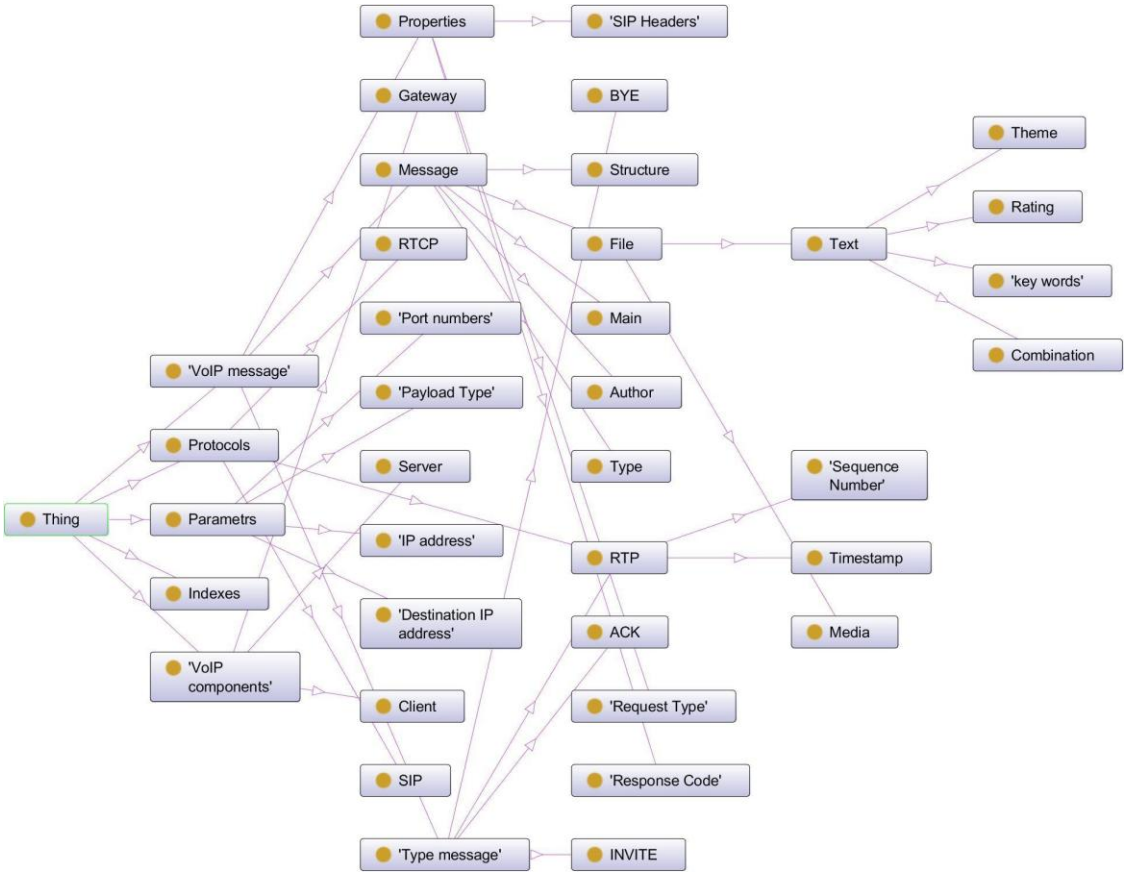

Рис. 3. Представлення формалізованої предметної області «VoIP повідомлення» у вигляді онтологічного графу

На третьому етапі формуємо опис ієрархічних зв'язків та будуємо таксономію класів, наприклад: протоколи VoIP (SIP, RTP, RTCP), компоненти VoIP (клієнт, сервер, шлюз). Будуємо відповідні відношення між класами, тобто визначаємо відношення між різними поняттями. Цей процес може включати такі зв'язки «є» (успадкування), «частина» (композиція), «має» (асоціація) тощо.

На четвертому етапі визначаємо властивості та атрибути, пов'язані з кожним поняттям. Вони представляють характеристики або ознаки, які їх описують.

На п'ятому етапі формуємо систему правил та обмеження для забезпечення узгодженості онтології. Це може включати обмеження відповідності правил домену предметної області, обмеження типів даних. Заключний етапом є перевірка онтології на реальних сценаріях VoIP, яка включає перевірку адекватності відношення до домену і відповідності до вимог практичного використання.

Метод автоматизованого наповнення онтології тематичних повідомлень в корпоративній мережі.

Онтологія описується деревоподібною системою концептів досліджуваної предметної області. Під концептами предметної області розуміються як базові поняття, так і узагальнюючі поняття, які часто є спільними у багатьох суміжних предметних областях. Онтологію для VoIP можна змоделювати деякою деревовидною структурою наступного виду

$$OC\_Voip =< IdConcept, ParentConcept, Base >, \qquad (1)$$

де $IdConcept$ - ідентифікатор концепту; $ParentConcept$ - ідентифікатор батьківського концепту; $Base$ - атрибут, який використовується для розмежування базових понять від узагальнюючих понять.

Базове поняття відрізняється від узагальнюючого тим, що узагальнюючі поняття не мають посилання на батьківський елемент, тобто $ParentConcept = 0$. Кожне поняття може описуватися деякими лінгвістичними представленнями у вигляді OC_Prs словосполучень. Онтологічні поняття описуються конкретними словоформами OC_Form, а також основами OC_Base цих понять. Словоформи можуть використовуватися для опису понять для користувачів, а основи понять – для автоматичного визначення еквівалентності представлень відповідно до вибраної мови. Атрибути таких понять групуємо в структури:

$$\begin{gathered} OC\_Base =< IdLanguage, IdBase, WordBase >, \\ OC\_Form =< IdLanguage, IdForm, IdBase, WordForm >, \\ OC_{Prs} =< IdConcept, IdLanguage, IdPhrase, \\ IdBase, IdForm, IdParentBase >, \end{gathered} \quad (2)$$

де $IdLanguage$ - ідентифікатор мови; $IdBase$ - ідентифікатор визначеної основи слова; $WordBase$ - визначена основа слова; $IdForm$ - ідентифікатор визначеної форми слова; $WordForm$ - визначена словоформа; $IdConcept$ - ідентифікатор аналізованого концепту; $IdPhrase$ - ідентифікатор аналізованої фрази; $IdParentBase$ - ідентифікатор основи поняття, що належить до батьківської категорії.

Для наповнення онтології VoIP можна використовувати:

– аналіз предметної області VoIP, включаючи протоколи, компоненти систем, типи повідомлень, характеристики зв'язку тощо;

– термінологічні словники, які використовуються у VoIP. Це включає технічні терміни, абревіатури, назви протоколів тощо;

– стандарти та специфікації VoIP, такі як SIP (Session Initiation Protocol), RTP (Real-time Transport Protocol), SDP (Session Description Protocol) тощо. Ці документи надають інформацію про протоколи, повідомлення, функції та інші аспекти VoIP;

– інформацію, що опублікована у відкритому доступі, таку як повідомлення в тематичних групах в соціальних мережах, форуми, вікі, які можуть містити корисні відомості про проблеми, вирішення, нові можливості розвитку в галузі VoIP;

– співпрацю з експертами у галузі VoIP, такими як інженери, адміністратори мереж, дослідники, які можуть допомогти у розумінні складних питань;

– програмні засоби VoIP, такі як Asterisk, FreeSWITCH, Kamailio, які містять документацію, код та інші ресурси, які можна використовувати для розуміння архітектури та принципів роботи VoIP;

– для формування бази тематичних медіа-повідомлень можна також використовувати спеціальні записи з медіа-ресурсів (Youtube, Tiktok, тощо).

Для пошуку та вибірки структурованої інформації, яка відноситься до аналізованої предметної області, формуємо множину $KeyConSet$, що містить ключові поняття, які характеризують її визначення та поведінку. Для аналізу структури VoIP повідомлень визначимо деяку допоміжну множину $AdvSet$:

$$\begin{gathered} AdvSet =< IdMessAn, IdListElem, IdElement, \\ IdBase, IdForm, IdParentBase >, \end{gathered} \quad (3)$$

де $IdMessAn$ – ідентифікатор аналізованого повідомлення; $IdListElem$ – ідентифікатор набору всіх понять; $IdElement$ – ідентифікатор окремого елемента набору.

Аналіз повторюваних понять дозволяє визначити найбільш значущі поняття і відокремити їх від несуттєвих понять. Для аналізу та контролю важливих понять використаємо деяку структуру $BaseFr$ частот основ:

$$BaseFr =< IdBase, BsFreq, IdLastMess, ConBack >, \quad (4)$$

де $IdBase$ – ідентифікатор основи, $BsFreq$ – частота появи основи в різних повідомленнях, $IdLastMess$ - ідентифікатор останнього повідомлення, де була основа поняття, $ConBack$ – відмітка фоновості поняття, що приймає $NULL$ значення за замовчуванням при деякій невизначеності.

При аналізі набору текстових повідомлень, які згруповані за відповідним критерієм, встановлюємо деякий ідентифікатор такого набору:

$$CurrMessId := \left(\pi_{IdLastMess}(BaseFr)\right) + 1. \quad (5)$$

Аналізовані елементи сформованої множини розбиваємо на набір елементарних понять із використанням роздільників, списки розбиваються на елементарні поняття $ElemCon$ із використанням роздільників.

До частин елементарного поняття необхідно застосувати процедуру $BaseConcept$ для формування їх основ. Ця процедура реалізовується шляхом відсікання закінчень від слів. При умові, що

$$BaseConcept(ElemCon_{Mess,List}) \subset BaseConcept(KeyConSet), \quad (6)$$

всі елементарні поняття з аналізованої множини з відповідними атрибутами додаємо в визначену структуру $AdvSet$. Словоформи вибираємо з елементів списку $ElemCon_{Mess,List}$ і розпізнаємо за допомогою відношення OC_Form або поповнюємо його.

Нехай $Word_k(ElemCon_{Mess,List})$ – k – те слово, яке відокремлене із елемента списку $ElemCon_{Mess,List}$. Якщо основа слова вже існує в аналізованій множині, то цей ідентифікатор $IdBaseWord$ визначаємо з відношення OC_Base:

$$IdBaseWord = \\ = \pi_{IdBase}\left(\sigma_{WordBase=BaseConcept(Word_k(ElemCon_{Mess,List}))}(OC\_Base)\right). \quad (7)$$

Якщо основа слова вже знаходиться у відношенні $BaseFr$ і номер аналізованого повідомлення не співпадає із номером зарахованого на попередньому кроці, то частотний індекс $BsFreq$ збільшуємо на 1, а номер поточного повідомлення заноситься в поле $IdLastMess$:

$$\left(Count\left(\pi_{IdBase}\left(\sigma_{IdBase=IdBaseWord\ AND\ IdLastMess<>CurrMessId}(BaseFr)\right)\right) \neq 0\right) \Rightarrow (BaseFr.BsFreq := BaseFr.BsFreq + 1) \wedge \\ \wedge (BaseFr.IdLastMess := CurrMessId).0) \quad (8)$$

Такий підхід забезпечує врахування частот використання основ у різних повідомленнях. Коли основа у відношенні $BaseFr$ не знайдена, її враховуємо з частотним індексом, який рівний 1, та додаємо у відношення, а також враховуємо номеру поточного повідомлення:

$$\left(Count\left(\pi_{IdBase}\left(\sigma_{IdBase=IdBaseWord}(BaseFr)\right)\right)=0\right)\Rightarrow\Rightarrow(BaseFr.IdBase:=IdBaseWord)\wedge(BaseFr.BsFreq:=1)\wedge\wedge(BaseFr.IdLastMess:=CurrMessId). \quad (9)$$

Якщо основа $Word_k(ElemCon_{Mess,List})$ не визначена, то вона додається у список основ, слово додається в множину словоформ, а відношення часто визначається на основі.

Для додавання поняття в онтологію експерту із знанням предметної області у IP-телефонії пропонуються тільки основи з частотою, яка більша за визначене мінімальне значення $BaseFr_0 \geq 2$, яке вибирається користувачем. Експерт із запропонованого списку основ здійснює вибір і включення в онтологію, коли дану умову задовольняє не менше $BaseFr_0$ основ

$$Count\left(\pi_{IdBase}\left(\sigma_{BaseFr>=BaseFr_0}(BaseFr)\right)\right)>BaseFr_0. \quad (10)$$

В процесі прийняття відповідного рішення, основи відображаються лише з контексту аналізованих повідомлень. Це дозволяє формувати поняття з кількох слів і не повторювати основи, які не обрані спеціалістом для включення в онтологію. Основу для контексту $ContextBase$ вибираємо виходячи із критерію максимальної частоти $MaxBsFreq$:

$$MaxBsFreq=\left(\pi_{BsFreq}(\sigma_{ConBack=Null})\right), \quad (11)$$

$$ContextBase=rand\left(\pi_{IdBase}\left(\sigma_{BsFreq=MaxBsFreq}\right)\right).$$

Формування контексту визначеної основи здійснюється із використанням двовимірного масиву $ArrayCont$, який містить словоформні ідентифікатори. Перший індекс характеризує номер фрази в контексті, а інший – ідентифікатор номера слова у фразі. $NumbPhrCont$ – показник, який характеризує кількість фраз контексту:

$$NumbPhrCont=count\left(\pi_{IdElement}\left(\sigma_{IdBase=CBase}(AdvSet)\right)\right). \quad (12)$$

Ідентифікатори усіх елементів з аналізованої структури заносимо в сформований масив $AdvPhrId$, який носить допоміжний характер:

$$AdvPhrId=\pi_{IdElement}\left(\sigma_{IdBase=CBase}(AdvSet)\right) \quad (13)$$

розмірності $NumbPhrCont$. І-та стрічка даного контекстного масиву визначається за наступним правилом:

$$AdvContext[i]=$$

$$=\left(\pi_{IdForm}\left(\sigma_{\substack{IdElement=AdvPhrN[i]\wedge\\ \wedge IdParentBase=NULL}}(AdvSet)\right)AS\ HeadLine\right)\cup$$

$$\cup\left(\pi_{IdForm}\left(\sigma_{\substack{IdElement=AdvPhrN[i]\wedge\\ IdParentBase=\\ \wedge\left(=HeadLine.IdBase\ OR\ IdParentBase=\right)\\ =HeadLine.IdBase}}(AdvSet)\right)AS\ HeadLine_1\right) \quad (14)$$

Сформований набір словоформ представляється спеціалістові в галузі VoIP та відповідної досліджуваної предметної області для формування елементів, які будуть поповнювати онтологію. Основи, які потрапляють в онтологію, отримують фоновий показник $ConBack := 2$, щоб уникнути процедур повторного аналізу. Основи, які не використовувалися жодного разу, не можуть формувати контекстні основи, і для них показник $ConBack := 1$. Вибрані елементи, які використовуються для онтологічного наповнення, можуть формувати онтологічну ієрархію, яка дозволяє формалізувати експертні знання у формі онтології.

Метод виявлення аномалій в трафіку IP-телефонії на основі групування повідомлень.

Для аналізу VoIP контенту необхідно формалізувати метод побудови базового повідомлення в IP-телефонії. Для цього необхідно спочатку здійснити перетворення VoIP медіа повідомлення в текстове представлення і таким чином сформувати базу даних вже текстових повідомлень.

Для перетворення голосового повідомлення в текстове використаємо бібліотеки SpeechRecognition для перетворення голосу в текст у мові програмування Python. Фрагмент лістингу програми такого перетворення представлено на рисунку 4.

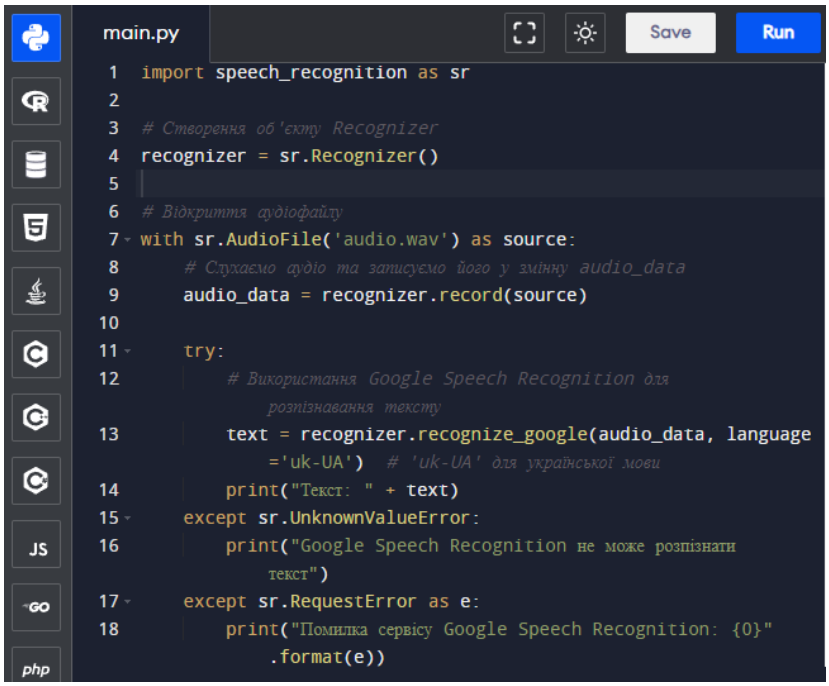

Рис. 4. Фрагмент лістингу програмного коду для перетворення VoIP медіа повідомлення в текстове представлення

Тематику таких повідомлень задаємо загальним текстовим ідентифікатором IMP предметної області та відповідним специфікатором SM. За допомогою перетворення голосового повідомлення у текстове представлення та процедури символьної конкатенації $VPS = IMP \,\&\, SM$ отримаємо множину VP текстових повідомлень, які представлені наборами відповідних компонентів конкретних понять

$$VP(VPS, P) = \{VP_i\}_{i=1}^{P}, \qquad (15)$$

де P – потужність множини VP ; VP_i – елемент множини, що визначає набір понять i - того повідомлення.

На наступному кроці для аналізу понять повідомлень використовуємо лише ті поняття, які відносяться до аналізованої предметної області, тобто належать множині VPK, а не здійснюємо повну перевірку усієї множини понять. Критерієм виконання такої умови є наявність ідентифікатора предметної області в темі повідомлення:

$$VPK = \{VP_{i*}^{*} | VP_{i*} \in VP, VP_{i*}.theme \cap \mathrm{I}MP \neq \emptyset\}_{i=1}^{P*}. \quad (16)$$

Із набору текстових тверджень необхідно вибрати інформацію, яка визначає тематику повідомлень. У першу чергу, інформація про тематику повідомлення визначається у наборі тверджень, з яких починається повідомлення.

Якщо початок повідомлення не дозволяє визначити його тематику, то інформацію про тематику вибираємо з впорядкованої множини елементів на основі частотного аналізу понять. Інформація такого типу буде впорядкованим списком $LKB(VP_i^*)$ виділених ключових понять:

$$LKB(VP_i^*) =< C_{ik}(VP_i^*) >_{k=1}^{CPi}. \quad (17)$$

Така множина може містити також і випадкову інформацію. Однак елементи множини понять, які будуть повторюватися, дають нам інформацію про структуру і тематику такого повідомлення. Тому на основі впорядкованої множини та множини ключових понять сформуємо базову BCM та узагальнену GCM множини пар поняття - частота появи поняття, які визначаємо наступним чином:

$$BCM = \{(CS_l, NCS_l) | CS_l\}, \quad (18)$$

$$GCM = \{(CS_m, NCS_m) | CS_m \in \cup_i \ LKB(VP_i^*)\}. \quad (19)$$

Для знаходження понять, які будуть релевантними до заданої предметної області, впорядкуємо елементи множини GCM у порядку спадання відповідних частот елементів, а деяку частину понять включаємо відразу в базову концептуальну множину CSC:

$$CSC = \left\{(CS_l, NCS_l) | FCL_l = \frac{NCS_l}{P^*} > FF_0\right\}, \quad (20)$$

де величина FF_0 належить інтервалу [0.200; 0.500], а конкретне значення цієї величини вибираємо із врахуванням особливостей та специфіки предметної області.

Далі аналізуємо множину понять, які мають низьку частоту, такий аналіз доцільний, якщо нам не вдалося наповнити множину понять на основі відношення. Розглянемо деяку визначену і впорядковану вибірку $SCMF$, яка також включає відповідні частоти понять у повідомленні:

$$SCMF = \{NCS_l | (CS_l, NCS_l) \in BCM\}. \quad (21)$$

Частоти з найвищими значеннями перевіряємо за критерієм 4σ на характеристику аномальності. Концепти, що відповідають критерію аномальності, включаємо в сформовану множину CSC.

При формуванні множини SK ключових концептів, що визначають тематику повідомлення, здійснюємо впорядкування цієї множини. Рангові номери присвоюємо лише ключовим концептам. Так $\mathrm{C}R\mathrm{C}_{ik}$ позначає ранг k-го концепту в і-тому повідомленні. При аналізі повідомлень необхідно також врахувати вплив тривалості голосового повідомлення (довжини перетвореного в текст повідомлення) на предметну аудиторію, оскільки є багато повідомлень, які є короткотривалими. Для того, щоб врахувати таку особливість, використаємо деяку вагову функцію $wcs(i)$.

Оскільки із спаданням тривалості голосових повідомлень їх важливість також буде плавно спадати, то чим коротше повідомлення, тим швидше відбуватиметься це спадання. Для моделювання такого процесу можна використати кубічний сплайн. Для його однозначного визначення необхідно накласти хоча б чотири умови. Зокрема, найтриваліше повідомлення буде мати важливість для цього аргументу рівну 1, якщо похідна дорівнює 0. Вводимо систему ваг $CSG(i) = \frac{csg(i)}{\sum_i \quad csg(i)}$, яка матиме нормований характер та обчислюємо усереднений ранг k-го концепту наступним співвідношенням:

$$RA_k = \sum_i \quad R_{ik} CSG(i). \tag{22}$$

Ранжування концептів здійснюємо на основі усереднення рангів. Узгодженість рангів перевіряємо із використанням коефіцієнта конкордації:

$$CSW = \frac{12 \sum_{k=1}^{K} \quad \left(\sum_{i=1}^{I} \quad R_{ik} CSG(i) - \bar{R}\right)}{I^2 (K^3 - K)}, \tag{23}$$

де $\bar{R} = \frac{1}{K} \sum_{k=1}^{K} \quad \sum_{i=1}^{I} \quad R_{ik} CSG(i)$.

Якщо виконується наступна умова $I(K-1)CSW > \chi^2_{K-1,\alpha}$, то таке ранжування є значущим.

Якщо ранжування повного списку концептів не є значущим, то відкидаємо один елемент із списку в порядку зростання відповідних ваг. Відкидання проводимо до отримання значущого ранжування, починаючи із концепту, який матиме найменшу вагу.

Запропонований метод дозволяє згрупувати повідомлення відповідно до класифікованої структури ключових понять. Таке групування доцільно використовувати для виявлення аномальних повідомлень в спеціалізованій корпоративній мережі з розгорнутою системою IP-телефонії.

У **третьому розділі** розроблено та детально описано архітектуру інтелектуалізованої системи виявлення зловмисних повідомлень IP-телефонії в умовах корпоративної мережі. Представлено узагальнену структуру інтелектуалізованої системи, яка включає низку взаємодіючих підсистем. Розроблено узагальнений алгоритм і структурно-функціональну схему інтелектуалізованої системи. Деталізація алгоритму дозволила формалізувати процеси виявлення, підвищивши їх прозорість та керованість.

Запропоновано механізм взаємозв'язку модулів для виявлення аномальних повідомлень на основі онтологічного підходу. Також запропоновано механізм автоматизованого наповнення онтології тематичних повідомлень у корпоративній мережі, що дозволяє системі постійно оновлювати свої знання про нормальні та аномальні патерни комунікації. Крім того, представлено метод виявлення аномалій у трафіку IP-телефонії на основі групування повідомлень, який дозволяє ідентифікувати відхилення у поведінці трафіку за допомогою методів кластеризації та класифікації. Деталізовано архітектуру програмного комплексу виявлення зловмисних повідомлень IP-телефонії в корпоративній мережі.

На основі проведеного вище аналізу запропоновано узагальнену структуру інтелектуалізованої системи забезпечення безпеки використання IP-телефонії в корпоративній мережі (рис. 5), серцевиною якої є п'ять базових підсистем.

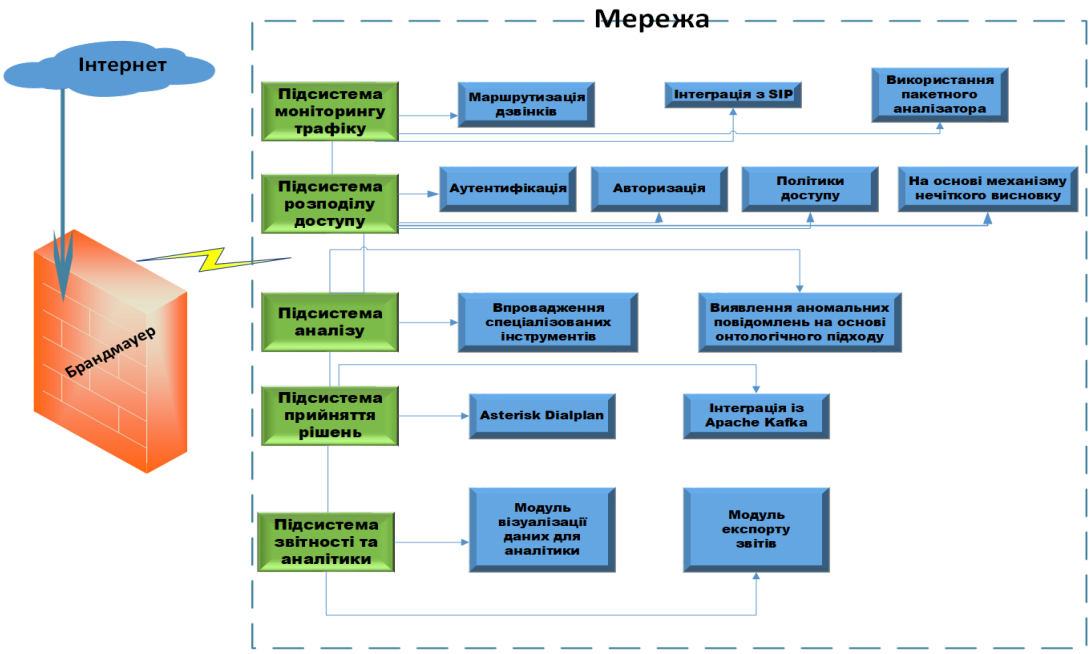

Рис. 5. Узагальнена структура інтелектуалізованої системи безпеки використання IP-телефонії в корпоративній мережі

На рисунку 6 наведена схема розподілу доступу клієнтів до мережі IP-телефонії на основі нечіткої логіки. Схема є узагальненою і може бути адаптована під конкретні вимоги та інфраструктуру.

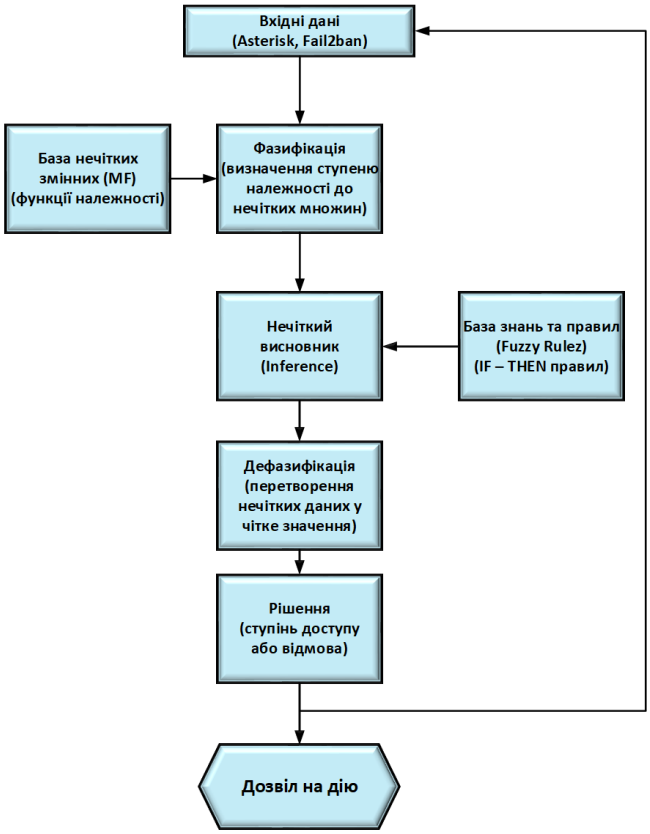

Рис. 6. Схема розподілу доступу клієнтів до мережі IP-телефонії на основі нечіткої логіки

На основі узагальненої структури інтелектуалізованої системи (див. рис. 5) та схеми розподілу доступу клієнтів (див. рис. 6) запропоновано узагальнений алгоритм функціонування інтелектуалізованої системи. Даний алгоритм описує роботу інтелектуалізованої комп'ютерної системи, яка забезпечує управління доступом та моніторингу в корпоративній мережі, також контроль доступу користувачів, моніторинг мережевого трафіку, аналіз потенційних загроз (аномалій), автоматизоване прийняття рішень щодо безпеки та формування звітності про результати.

Запропонований алгоритм є ефективним інструментом, орієнтованим на автоматизацію прийняття рішень і швидке реагування на потенційні загрози.

З врахуванням викладеного вище, розроблено структурно-функціональну схему інтелектуалізованої комп'ютерної системи виявлення зловмисних повідомлень IP-телефонії в корпоративній мережі (рис. 7.). Взаємозв'язок модулів для виявлення аномальних повідомлень на основі онтологічного підходу (рис. 8.).

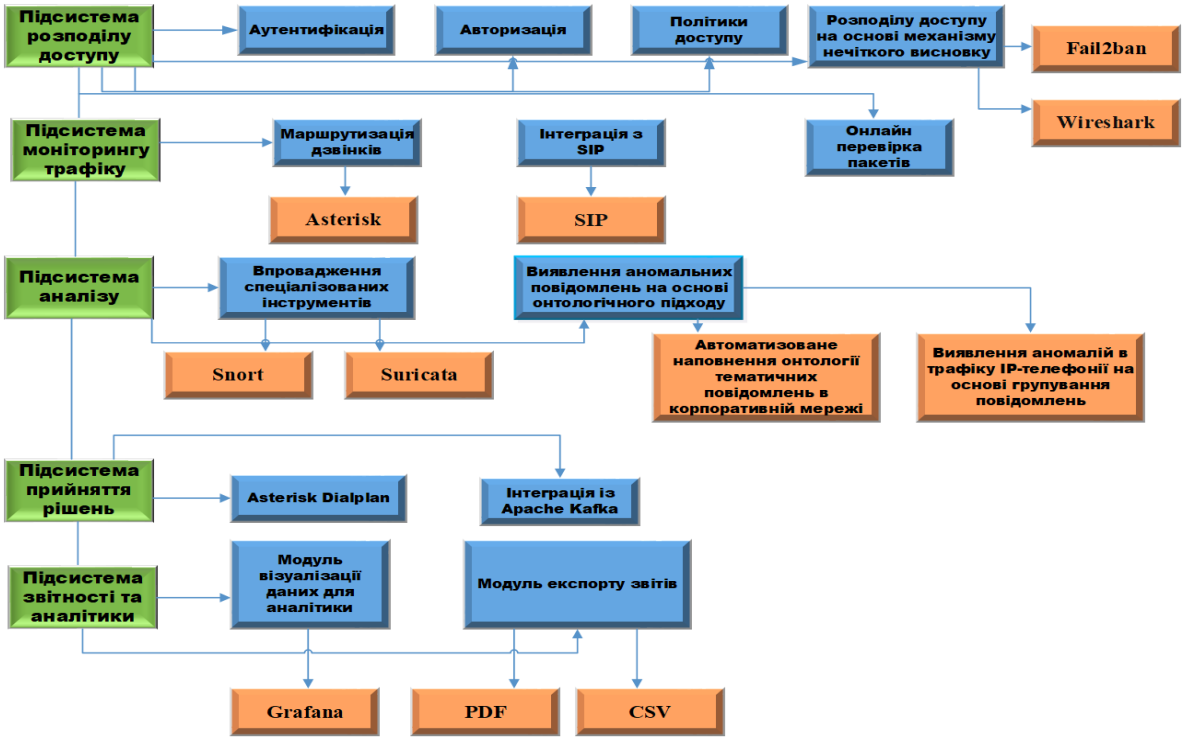

Рис. 7. Структурно-функціональна схема інтелектуалізованої системи

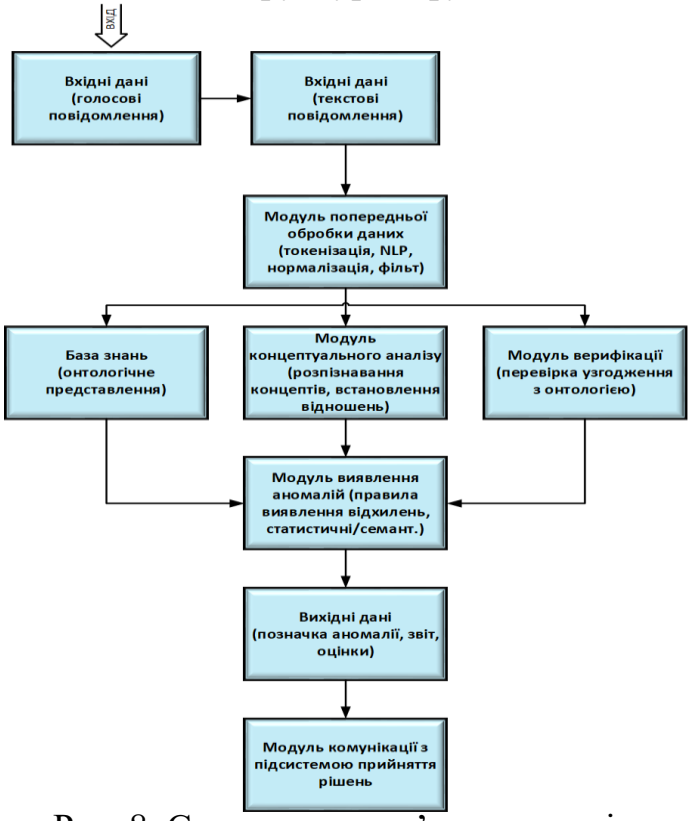

Рис. 8. Схема взаємозв'язку модулів для виявлення аномальних повідомлень на основі онтологічного підходу

Представлена схема та опис показують цілісний процес обробки та аналізу текстових повідомлень на семантичному рівні за допомогою онтологічного підходу. Він дозволяє враховувати значення, контекст та логічну узгодженість повідомлень для виявлення прихованих чи нетривіальних аномалій, які складно ідентифікувати лише статистичними чи поверхневими методами.

Архітектура програмного комплексу систем виявлення зловмисних повідомлень IP-телефонії в корпоративній мережі призначена для забезпечення безпечного функціонування мережі, захисту голосового трафіку та конфіденційності інформації, що передається через IP-мережі. Основна мета такої архітектури – мінімізувати

ризики від зовнішніх та внутрішніх загроз, забезпечуючи при цьому якісну та безперебійну роботу IP-телефонії.

На рисунку 9 представлено загальну архітектуру програмного забезпечення для захисту IP-телефонії. Необхідно відзначити, що така архітектура відображає взаємодію по рівнях.

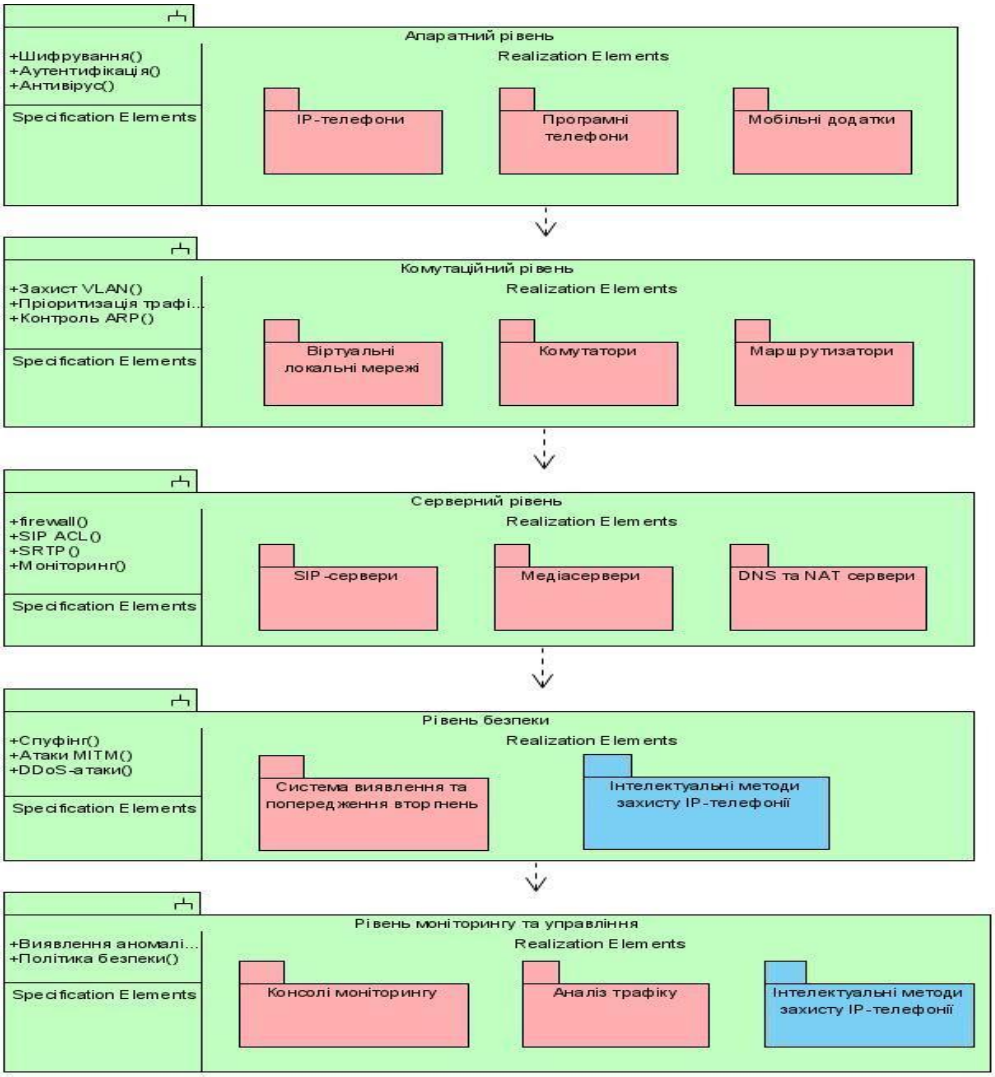

Рис. 9. Загальна архітектура програмного забезпечення для захисту IP-телефонії

Програмне забезпечення належить до класу об’єкто-орієнтованих систем, розроблене з використанням мови програмування PHP, а також СУБД MySQL. Система працює як інтегроване програмне забезпечення в середовищі Asterisk під управлінням операційною системою Linux.

Четвертий розділ присвячений дослідженню створених методів та засобів, які використовуються для підвищення ефективності захисту IP-телефонії у корпоративній мережі.

Розглянуто застосування нечіткої логіки для захисту від термінації трафіку в IP-телефонії. Основною метою є розробка нечіткої системи, яка забезпечує розподіл доступу в режимі реального часу, враховуючи різні критерії.

Моделювання нечіткої системи здійснено за допомогою Fuzzy Logic Toolbox середовища MATLAB. На вхід такої системи надходять значення рівня довіри клієнта по його IP-адресі (*IP-identification*), часу здійснення дзвінка клієнтом (*client-call-time*), часового поясу адресата дзвінка (*addr-call-time*) та якості зв'язку мережі клієнта (*connection-quality*). Загальний вигляд запропонованої нечіткої системи подано на рисунку 10.

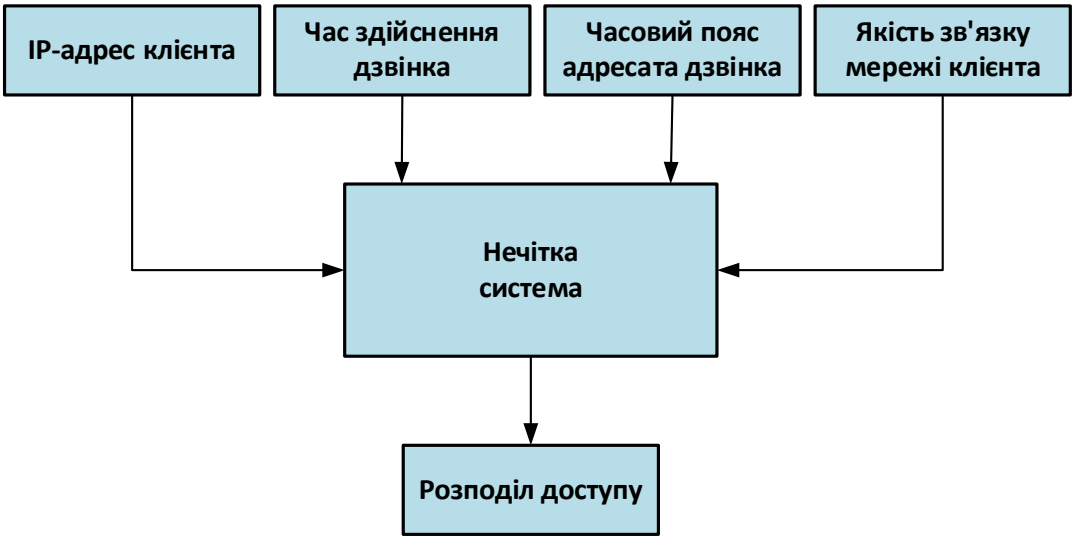

Рис. 10. Загальна схема нечіткої системи розподілу доступу в IP-телефонії

Кожна із вхідних змінних демонструє дані з певного діапазону. Над цими даними здійснюється фазифікація, тобто переведення кількісних значень у якісні.

Значення функцій належності вхідних змінних задається дзвоноподібною, а вихідної змінної – трапецевидною функцією.

Дзвоноподібна функція задається трьома числами (a, b, c), що відповідають абсцисам крайніх точок та центру кривої:

$$MF(x) = \frac{1}{1 + \left| \frac{x - c}{a} \right|^{2b}} \tag{24}$$

Функції належності вхідної змінної *IP-identification* відображаються трьома множинами: *low*, *middle*, *high*, що відображають відповідно низький, середній та високий рівень довіри клієнта IP-телефонії. Цей показник легко можна сформувати зі статистичних даних IP-адреси клієнта, його присутності в IP-телефонії, здійснених атак чи наявності збоїв при здійсненні дзвінків та правах доступу, наданих йому адміністратором мережі. Для зручності вибрано інтервал задання значень цієї змінної [0, 1].

Функції належності змінної *connection-quality*, що відображає якість зв'язку мережі, з якої здійснює запит до IP-телефонії на дзвінок клієнт, аналогічно поділені на три інтервали *low*, *middle*, *high*, що задані на відрізку [0, 1]. Низька якість зв'язку може бути у мобільних мережах через різницю у покритті та якості надання послуг

мобільними операторами. Середня якість зв'язку, як правило, можлива при застосуванні користувачем регіональної мережі Інтернет, а висока якість забезпечується використанням клієнтом локальної мережі. Для задання функцій належності змінної *client-call-time*, що показує час здійснення дзвінка клієнтом в інтервалі [0, 24], пропонується на три діапазони *morning, work time, night*.

Змінна *addr-call-time* відображає часові пояси світу, які поділені на три зони (*I-timezone*, *II-timezone*, *III-timezone*), які подані на інтервалі [-12, 12] та враховуються при побудові бази правил відносно часу здійснення дзвінка клієнта.

Трапецевидна функція задається четвіркою чисел (*a*, *b*, *c*, *d*), які позначають абсциси вершин трапеції, заданої функцією:

$$MF(x)=\begin{cases}\dfrac{b-x}{b-a}, a\le x\le b\\ 1, b\le x\le c\\ \dfrac{x-c}{d-c}, c\le x\le d\\ 0, \text{в інших випадках}\end{cases}. \qquad (25)$$

Функції належності для вихідної змінної *access* позначаються двома трапецевидними інтервалами *deny*, *allow* для точного визначення центру ваги, що позначає нечіткий висновок системи.

База знань для побудови даної нечіткої моделі складається з правил типу «якщо - то», усі вхідні змінні мають по три нечітких стани і ще один стан *none*, коли значення вхідної змінної не задане системою. Випадок, коли значення усіх вхідних змінних не задані, на практиці неможливий при функціонуванні системи IP-телефонії, тому кількість правил нечіткого висновку досліджуваної системи 4•4•4•4-1=255.

Нечіткий висновок підсистеми розподілу доступу, побудованого на основі заданих 255 правил з поточними значеннями вхідних та вихідної змінних, має вигляд, фрагмент якого представлений на рисунку 11.

На рисунку 11 зображено випадок, коли клієнт, рівень довіри якого низький, здійснює дзвінок у ранковий час в країну, яка має той самий часовий пояс, що і Україна, і якість зв'язку свідчить, що дзвінок здійснюється з локальної мережі. Висновок запропонованої нечіткої системи в даному випадку становить 0.48, що відповідає дозволу на здійснення даного дзвінка. Даний приклад підтверджує правильність роботи сформованої бази правил. Перевірка інших комбінацій значень вхідних змінних демонструє правильність роботи такої нечіткої системи. Запропонована нечітка система може реалізовуватися як програмно, так і апаратно у вигляді нечіткого контролера для успішного захисту від термінації трафіку.

Апробовано запропонований метод виявлення аномальних повідомлень у трафіку IP-телефонії був інтегрований до прототипу діючої корпоративної системи з метою оцінки його ефективності в умовах реального використання. Імплементація здійснювалася як додатковий модуль до підсистеми контролю трафіку, що дозво-

лило не лише зберегти існуючу архітектуру мережевої інфраструктури, але й забезпечити її розширення без значних витрат ресурсів або порушення стабільності основних сервісів.

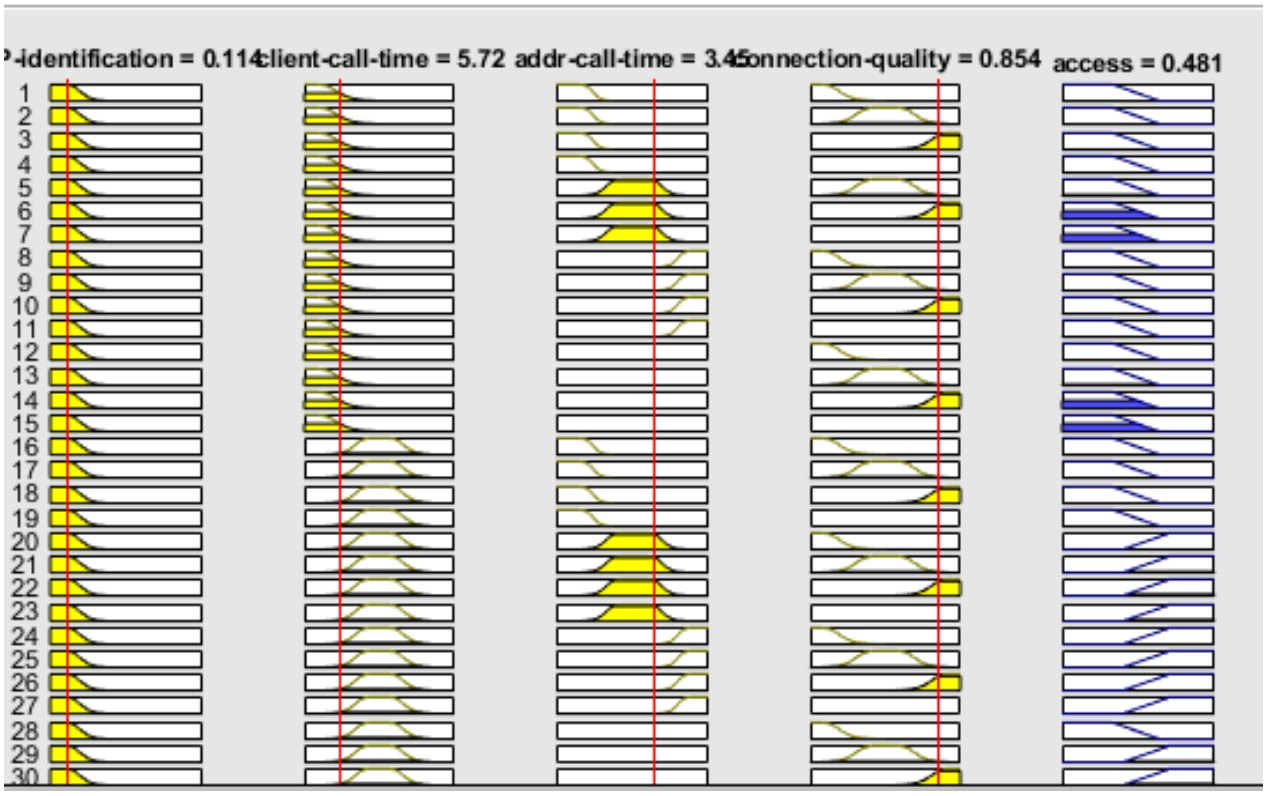

Рис. 11. Нечіткий висновок підсистеми розподілу доступу в IP-телефонії

Важливо відзначити, що завдяки модульному характеру розробленого рішення його можливо масштабувати для використання в системах різного рівня, від локальних корпоративних офісів до міжрегіональних вузлів операторів зв'язку. Таким чином, впровадження методу може здійснюватися поетапно, з мінімальним втручанням у поточні процеси функціонування інфраструктури IP-телефонії.

В процесі апробації запропонованих методів було здійснено інтеграцію даної інтелектуалізованої підсистеми в діючу корпоративну VoIP систему Західноукраїнського національного університету.

На рисунку 12 представлено сторінку, яка, окрім підсистеми аналізу ключових понять у VoIP, включає також модуль для перетворення голосових повідомлень в текстові представлення і адміністративний модуль, що використовується для прийняття рішень відносно користувачів системи та відповідне управління виявленими аномальними повідомленнями.

Дана інтеграція може також використовуватися і для виявлення голосових повідомлень, які здійснюють користувачі при дзвінках на платні сервіси, що також є дуже актуальним питанням в системі практичного використання IP-телефонії.

Загалом результати впровадження підтверджують, що запропонований метод є не лише теоретично обґрунтованим, але й практично застосовним. Його використання дозволяє значно підвищити рівень кібербезпеки в IP-телефонних системах за рахунок своєчасного виявлення аномалій, зменшення кількості хибнопозитивних спрацьовувань та зниження витрат часу на ручний аналіз мережевого трафіку.

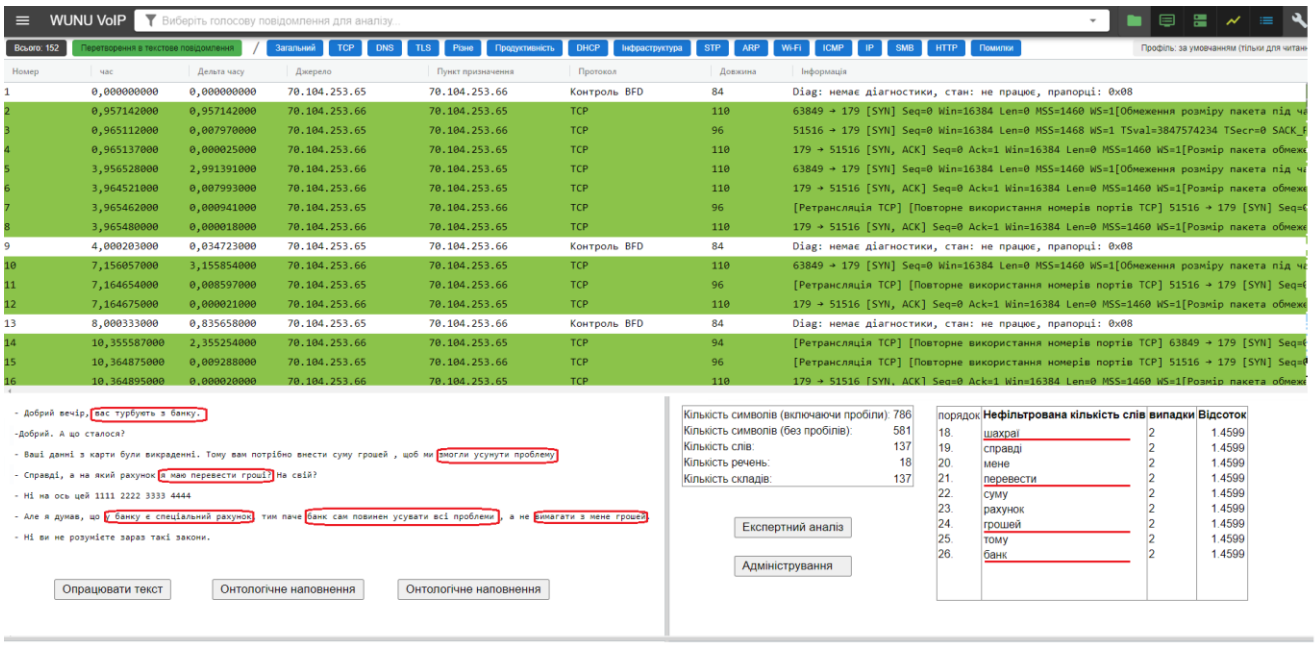

Рис. 12. Інтерфейс інтеграції підсистеми аналізу VoIP повідомлень в систему управління корпоративною IP-телефонією

З метою оцінювання ефективності розробленого методу виявлення аномальних повідомлень у трафіку IP-телефонії було проведено серію експериментів, що базуються на аналізі VoIP-повідомлень у режимі реального часу. Запропонований підхід передбачає групування тематично схожих SIP-повідомлень з використанням структурованої онтології домену IP-телефонії, яка автоматично оновлюється в процесі аналізу трафіку.

Для оцінки продуктивності системи було змодельовано ситуацію, в якій порівнюється швидкість обробки повідомлень оператором вручну та системою автоматичного виявлення.

При обробці 50 повідомлень система витратила 1185 секунд, у той час як оператор – 6720 секунд. Аналогічна тенденція спостерігалася при обробці 100 повідомлень: автоматизована система завершила процес за 1730 секунд, тоді як вручну оператору для цього було потрібно 14356 секунд. Отже, із зростанням кількості повідомлень різниця в часі між ручною та автоматизованою обробкою суттєво зростає, що свідчить про хорошу масштабованість запропонованого рішення, рисунок 13.

Проведені експериментальні дослідження підтвердили ефективність методу виявлення аномальних VoIP-повідомлень, заснованого на групуванні та онтологічному аналізі трафіку IP-телефонії. Результати показали суттєве зниження часу обробки повідомлень у порівнянні з ручною роботою оператора, при цьому ефективність системи зростає зі збільшенням обсягів даних.

Ключовим чинником, що забезпечує приріст продуктивності, є механізм автоматичного поповнення онтології, який дає змогу враховувати нові шаблони та семантичні характеристики аномальних повідомлень у корпоративній мережі. Таким чином, розроблена система демонструє високий потенціал до впровадження в інфраструктуру IP-телефонії для забезпечення оперативного виявлення аномалій, підвищення надійності комунікацій та зниження навантаження на персонал.

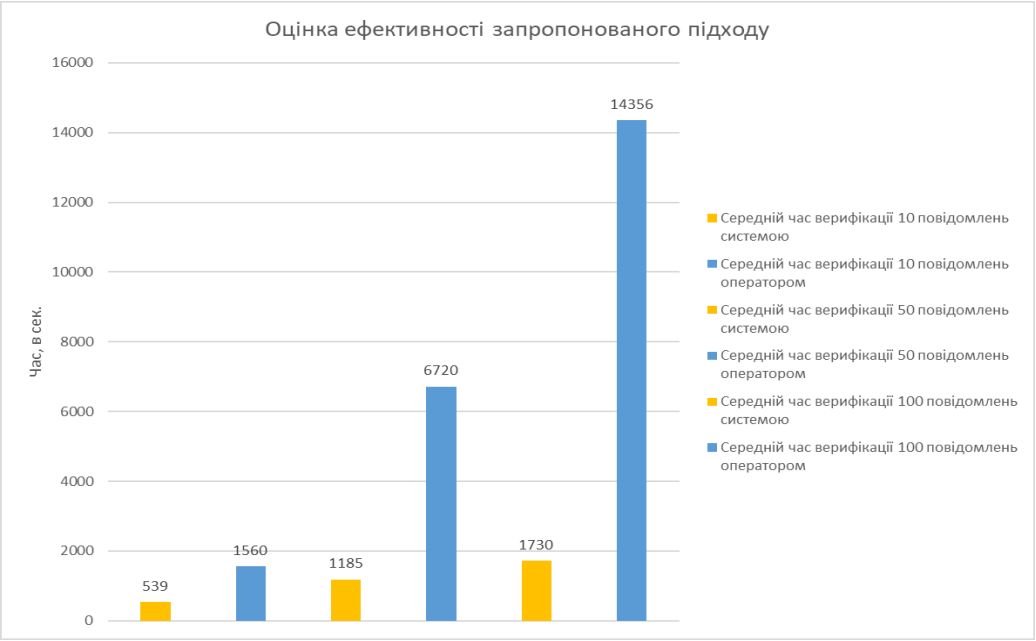

Рис. 13. Результат: скорочення часу опрацювання повідомлень в середньому від 2.89 до 8.3 рази

На основі аналізу узагальненої структури інтелектуалізованої комп'ютерної системи виявлення зловмисних повідомлень IP-телефонії в корпоративній мережі ICSDMM-VoIP (Intelligent Computer System for Detection of Malicious Messages), виконано порівняльну оцінку основних показників, які впливають на ефективність контролю нетипової поведінки користувачів IP-телефонії (таблиця 1).

Таблиця 1

Порівняння розробленої системи контролю нетипової поведінки користувачів IP-телефонії з відомими рішеннями

Техніко-економічні характеристики	Системи				
	Wireshark	Suricata	Snort	Asterisk	ICSDMM - VoIP
Аналітика та звітність	+/–	–	+/-	+/-	+
Багатопотокова обробка даних	+/–	+	–	+	+
Адаптивна інтеграція	+/–	–	+	+	+
Масштабованість	–	–	–	+	+
Інтелектуальний аналіз даних	–	–	–	–	+
Вартість	Freeware	Freeware	Shareware	Freeware	Freeware

Результати впровадження підтверджують, що використання розробленої інтелектуалізованої системи виявлення зловмисних повідомлень дозволяє суттєво підвищити рівень кібербезпеки в IP-телефонії за рахунок своєчасного виявлення аномалій, зменшення кількості хибних спрацьовувань та зниження витрат часу на рутинний аналіз мережевого трафіку.

ВИСНОВКИ

У дисертаційній роботі вирішено актуальну науково-прикладну задачу створення інтелектуалізованої комп'ютерної системи своєчасного виявлення зловмисних VoIP-повідомлень в корпоративній мережі з метою підвищення ефективності захисту IP-телефонії, що базується на онтологічній моделі її функціонування та методах інтелектуального аналізу даних.

Основні наукові та практичні результати дослідження дозволяють сформулювати наступні висновки:

1. Проведено аналіз сучасних методів та засобів захисту IP-телефонії у корпоративних мережах. Проведено систематизацію загроз, притаманних IP-телефонії, зокрема атак типу spoofing, sniffing, DoS/DDoS, маніпулювання сигналізаційним трафіком (SIP), термінацію VoIP-трафіку. Встановлено, що традиційні засоби безпеки, такі як брандмауери та VPN, є недостатньо ефективними проти складних, багаторівневих атак у VoIP-середовищі. Доведено доцільність застосування методів інтелектуального аналізу даних для підвищення ефективності захисту VoIP-інфраструктури.

2. Запропоновано онтологію опису VoIP-повідомлень у межах функціонування системи IP-телефонії. В межах онтологічного моделювання здійснено формалізацію ключових понять предметної області, які репрезентовано у вигляді окремих концептів. Кожному концепту присвоєно чітке визначення, атрибутивний склад та типові екземпляри, що забезпечує однозначність інтерпретації інформації як людиною, так і програмними агентами. Розроблена онтологія є гнучкою основою для подальшого розширення та адаптації під конкретні прикладні задачі, зокрема для виявлення аномалій у трафіку, забезпечення семантичного аналізу повідомлень, генерації сигнатур для систем захисту від VoIP-атак, а також інтеграції з інструментами інтелектуального аналізу даних у сфері мережевої безпеки. Її використання дозволяє підвищити рівень інтероперабельності, автоматизації та формалізації у процесах контролю, аудиту та реагування на події в IP-телефонії.

3. Запропоновано метод автоматизованого наповнення онтології тематичних повідомлень у корпоративній системі IP-телефонії, що є суттєвим внеском у підвищення ефективності побудови онтологічних баз знань для аналізу комунікаційного трафіку. Основою методу є формалізоване представлення VoIP-повідомлень у вигляді деревовидних структур, які моделюють логічну та ієрархічну організацію повідомлень із розподілом на рівні заголовків, полів, параметрів і вкладених елементів. Такий підхід дозволяє здійснювати семантичний аналіз та структуризацію повідомлень із врахуванням протоколів SIP, SDP, RTP та інших, що використовуються в IP-телефонії. Додатково, операції аналізу та взаємодії між структурними компонентами повідомлень описано за допомогою алгебри кортежів, яка забезпечує формальну основу для їх порівняння, об’єднання, фільтрації та трансформації у відповідні концепти онтології. Це дозволяє автоматизувати процес ідентифікації нових сутностей і зв’язків, які мають бути внесені до онтологічної моделі, що в свою чергу дає можливість зменшити час наповнення онтологічної бази знань.

4. Розроблено метод виявлення аномалій у трафіку IP-телефонії, заснований на групуванні VoIP-повідомлень з використанням онтологічного підходу та контекстно-частотного аналізу, що дозволяє значно підвищити точність і швидкість

виявлення нестандартної або шкідливої активності у корпоративних мережах зв’язку. Метод базується на інтеграції знання-орієнтованої моделі, у якій повідомлення описані за допомогою спеціалізованої онтології, що відображає типи повідомлень, їхні структури, допустимі послідовності взаємодії та семантичні зв’язки між учасниками сесії зв’язку. Така онтологічна база дозволяє порівнювати вхідні повідомлення з відомими шаблонами «нормальної» поведінки. Розроблений метод забезпечив підвищення ефективності виявлення аномальних повідомлень, у середньому 5, 6 разів, порівняно з традиційними підходами. Це досягнуто за рахунок глибшого урахування контексту комунікаційних подій, семантики повідомлень та гнучкого онтологічного представлення знань про мережеві взаємодії.

5. Розроблено підсистему для захисту від термінації трафіку в IP-телефонії на основі використання нечіткої логіки. Підсистема здатна гнучко адаптуватися до змінних умов і приймати рішення на основі нечітких даних. Це дозволяє враховувати численні критерії, такі як IP-адреса клієнта, час здійснення дзвінка, часовий пояс і якість зв’язку, що сприяє підвищенню рівня безпеки та якості зв'язку.

6. Розроблено архітектуру, структурно-функціональні схеми та узагальнений алгоритм функціонування інтелектуалізованої системи виявлення аномалій у трафіку IP-телефонії, яка об’єднує окремі підсистеми в єдиний цілісний інформаційно-аналітичний контур. Система охоплює всі етапи обробки VoIP-трафіку — від первинного збору даних до прийняття рішень щодо наявності аномальних або потенційно загрозливих повідомлень. Завдяки такій архітектурі забезпечено модульність, гнучкість і масштабованість розробленого рішення, а також можливість інтеграції з зовнішніми системами інформаційної безпеки, SIEM-платформами та системами моніторингу.

7. Проведено експериментальні дослідження запропонованих методів і засобів для виявлення аномалій у трафіку IP-телефонії, що дозволило кількісно оцінити їх ефективність та практичну придатність. З метою перевірки працездатності розроблених методів імплементовано програмну підсистему виявлення аномальних VoIP-повідомлень, яка інтегрована в університетську корпоративну систему IP-телефонії. Програмна підсистема реалізує компоненти онтологічного аналізу повідомлень, модуль контекстно-частотної обробки, інтерфейси інтеграції з потоковими джерелами SIP/SDP/RTP-трафіку, а також механізми прийняття рішень щодо ідентифікації аномальних подій. Це дозволило здійснювати обробку повідомлень у режимі, наближеному до реального часу, що критично важливо для виявлення атак типу DoS, SIP flooding, spoofing, несанкціонованого доступу та інших загроз.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. A. Melnyk, R. Shevchuk, I. Romanets, I. Yakymenko, S. Voznyak, and V. Luchyk, "A Method of Detecting Anomalies in IP Phone Traffic based on Ontology of Voip Messages," *2024 14th International Conference on Advanced Computer Information Technologies (ACIT)*, Ceske Budejovice, Czech Republic, 2024, pp. 485-489, DOI: 10.1109/ACIT62333.2024.10712505.
2. Balyk, A., Karpinski, M., Naglik, A., Shangytbayeva, G., & Romanets, I. (2017). Using Graphic Network Simulator 3 For DDoS Attacks Simulation. International Journal of Computing, 16(4), 219-225 DOI: 10.47839/ijc.16.4.910
3. Todd J.B. Blayone, Olena Mykhailenko, Svetlana Usca, Ihor Romanets, Exploring technology attitudes and personal–cultural orientations as student readiness factors for digitalised work. *Higher Education, Skills and Work-based Learning* 17 June 2021; 11 (3): 649–671. DOI:10.1108/HESWBL-03-2020-0041
4. Романець І. Захист від термінації трафіку в IP-мережі на основі нечіткої логіки: Науковий журнал "Вимірювальна та обчислювальна техніка в технологічних процесах"(1), Хмельницький національний університет, Хмельницький, 2024 с. 186-193, DOI: 10.31891/2219-9365-2024-77-23.
5. І. Романець, «Онтологічний підхід в системі забезпечення безпеки використання IP-телефонії», *Опт-ел. інф-енерг. техн.*, вип. 47, вип. 1, с. 240–252, Чер. 2024. DOI: 10.31649/1681-7893-2024-47-1-240-252.
6. І. Романець, "Архітектура інтелектуалізованої системи забезпечення безпеки використання IP-телефонії в корпоративній мережі", *Вісник Хмельницького національного університету. Серія: Технічні науки*, vol. 337, no. 3(2), pp. 425–436, May 2024, DOI: 10.31891/2307-5732-2024-337-3-65.
7. Markowsky, G., Sachenko, A., Voznyak, S., Spilchuk, V., Romanyak, R., Turchenko, V. and Romanets, I. 2014. The Ternopil Educational Communication Center: A Nato Project To Integrate Regional Information Technology Resources. *International Journal of Computing*. 7, 1 (Aug. 2014), 185-190. DOI:10.47839/ijc.7.1.503.
8. I. Romanets, A. Sachenko and L. Dubchak, "Method of Protection Against Traffic Termination in VoIP," 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Iasi, Romania, 2018, pp. 1-5, DOI: 10.1109/ECAI.2018.8678992.
9. Komar, M., Kochan, V., Dubchak, L., Sachenko, A., Golovko, V., Bezobrazov, S., Romanets, I. "High performance adaptive system for cyber attacks detection", 2017 IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2017, Bucharest, Romania, 2, art. no. 8095208, pp. 853-858. DOI: 10.1109/IDAACS.2017.8095208.
10. M. Dyvak, A. Pukas, A. Melnyk, I. Voytyuk, S. Valchyshyn, and I. Romanets, "Software Architecture for Modeling the Interval Static and Dynamic Objects," 2021 11th International Conference on Advanced Computer Information Technologies (ACIT), Deggendorf, Germany, 2021, pp. 572-575, DOI: 10.1109/ ACIT52158.2021.9548577.
11. M. Dyvak, A. Melnyk, L. Dostalek, V. Ostroverkhov, L. Honchar, and I. Romanets, "Repository of Interval Models of Dynamics of Concentrations of Harmful Emissions of Motor Transport," 2022 12th International Conference on Advanced Computer Information

Technologies (ACIT), Ruzomberok, Slovakia, 2022, pp. 89-94, DOI: 10.1109/ACIT54803.2022.9913123.

12. O. Kochan, O. Osolinskyi, A. Sachenko, V. Kochan, and I. Romanets, "Simulator of Microcontroller's Power Consumption," 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Dortmund, Germany, 2023, pp. 787-792, DOI: 10.1109/IDAACS58523.2023.10348884.
13. V. Tymchyshyn, B. Tymchyshyn, A. Melnyk, V. Manzhula, V. Faifura, and I. Romanets, "The System Architecture of the Software for Modeling Harmful Emissions in Soil," 2023 13th International Conference on Advanced Computer Information Technologies (ACIT), Wrocław, Poland, 2023, pp. 58-62, DOI: 10.1109/ACIT58437.2023.10275416.
14. A. Sartiukova, O. Markiv, V. Vysotska, I. Shakleina, N. Sokulska, and I. Romanets, "Remote Voice Control of Computer Based on Convolutional Neural Network," 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Dortmund, Germany, 2023, pp. 1058–1064, DOI: 10.1109/IDAACS58523. 2023.10348808.

АНОТАЦІЯ

Романець І.Є. Інтелектуалізована комп'ютерна система виявлення зловмисних повідомлень IP-телефонії в корпоративній мережі – Рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.05 – Комп’ютерні системи та компоненти». – Західноукраїнський національний університет, Тернопіль, 2025.

Дисертаційна робота присвячена актуальній науково-технічній задачі підвищення рівня захищеності корпоративних мереж шляхом розробки та дослідження інтелектуалізованої комп'ютерної системи, призначеної для виявлення зловмисних повідомлень в IP-телефонії.

В роботі проаналізовано сучасні загрози та вразливості систем IP-телефонії в корпоративному середовищі, такі як спам через інтернет-телефонію, шахрайство з оплатою дзвінків та атаки на відмову в обслуговуванні (DoS). Виявлено недостатню ефективність існуючих методів та засобів захисту перед новими та адаптивними атаками, що зумовлює необхідність застосування інтелектуальних підходів.

Запропоновано архітектуру інтелектуалізованої комп'ютерної системи, яка використовує методи машинного навчання для аналізу трафіку (SIP-повідомлень) та метаданих медіа-потоків (RTP/RTCP) IP-телефонії. Розроблено моделі та алгоритми для виділення інформативних ознак зловмисних повідомлень, класифікації трафіку та виявлення аномалій в реальному часі. Особливу увагу приділено здатності системи до самонавчання та адаптації до нових, раніше невідомих типів загроз.

Досліджено ефективність різних алгоритмів машинного навчання для вирішення задачі класифікації та виявлення аномалій в IP-телефонії. Проведено експериментальну перевірку розробленої системи на реальних та змодельованих наборах даних, що підтвердило її здатність виявляти зловмисні повідомлення з високою точністю та низьким рівнем помилкових спрацювань.

Наукова новизна роботи полягає у розробці комплексного підходу до виявлення зловмисних повідомлень в IP-телефонії на основі інтелектуального аналізу багатовимірних даних, що враховує специфіку корпоративних мереж, а також у вдосконаленні методів адаптивного виявлення загроз на основі онтологічного підходу та нечіткої логіки.

Практичне значення отриманих результатів полягає у розробленому програмному прототипі системи, рекомендаціях щодо її впровадження та налаштування в корпоративних мережах для підвищення рівня безпеки IP-телефонії, запобігання фінансовим втратам та захисту конфіденційної інформації.

Ключові слова: IP-телефонія, VoIP, зловмисні повідомлення, вішинг, інтелектуалізована система, машинне навчання, виявлення аномалій, кібербезпека, захист інформації.

ANNOTATION

Romanets I.Y. Intelligent computer system for detecting malicious IP-telephony messages in a corporate network – Manuscript.

Thesis for a Ph.D. degree by specialty 05.13.05 – Computer Systems and Components – West Ukrainian National University, Ternopil, 2025.

The dissertation is devoted to the pressing scientific and technical challenge of enhancing the security level of corporate networks by developing and researching an intellectualized computer system designed to detect malicious messages in IP telephony.

The work analyzes modern threats and vulnerabilities in IP telephony systems within corporate environments, such as spam over internet telephony (SPIT), call payment fraud, and denial-of-service (DoS) attacks. The insufficient effectiveness of existing protection methods and tools against new and adaptive attacks has been identified, necessitating the application of intelligent approaches.

An architecture for an intellectualized computer system is proposed, utilizing machine learning methods to analyze traffic (e.g., SIP messages) and metadata of media streams (RTP/RTCP) in VoIP. Models and algorithms have been developed to extract informative features of malicious messages, classify traffic, and detect anomalies in real time. Special attention is given to the system's ability to self-learn and adapt to new, previously unknown types of threats.

The effectiveness of various machine learning algorithms for solving the tasks of classification and anomaly detection in VoIP has been investigated. Experimental testing of the developed system on real and simulated datasets confirmed its ability to detect malicious messages with high accuracy and a low false-positive rate.

The scientific novelty of the work lies in the development of a comprehensive approach to detecting malicious messages in VoIP based on intelligent analysis of multidimensional data, considering the specifics of corporate networks, as well as in the improvement of adaptive threat detection methods using ontological approaches and fuzzy logic.

The practical significance of the obtained results is reflected in the developed software prototype of the system, recommendations for its implementation, and configuration in corporate networks to enhance IP-telephony security, prevent financial losses, and protect confidential information.

Keywords: IP-telephony, VoIP, malicious messages, vishing, intellectualized system, machine learning, anomaly detection, cybersecurity, information protection.