УДК 004.891

д.т.н., проф. **Лєнков С.В.** (ВІКНУ)
к.т.н., доц., **Джулій В.М.** (ХмНУ)
к.т.н., с.н.с. **Мірошніченко О.В.** (ВІКНУ)
к.т.н., доц. **Браун В.О.** (ВІКНУ)
Прохорський С.І. (ВІТІ)

DOI: https://doi.org/10.17721/2519-481X/2023/79-11

ІНФОРМАЦІЙНО-АНАЛІТИЧНА СИСТЕМИ ПРОГНОЗУВАННЯ ВРАЗЛИВОСТЕЙ ТА ЗАГРОЗ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

В роботі запропоновано структурну схему інформаційно-аналітична системи прогнозування вразливостей та загроз інформаційної безпеки. Аналіз проведених досліджень дозволяє зробити висновок, що для вирішення задачі дослідження та розробки інформаційно-аналітичної нечіткої системи для логічного нечіткого виводу про появу вразливостей та загроз інформаційної безпеки, автоматизації проведення аналізу потоку повідомлень тематичних інтернет-ресурсів, про доцільність використання експертних систем прогнозування. Для вирішення задач прогнозування вразливостей та загроз інформаційної безпеки конфіденційних даних на основі потоку тематичних повідомлень інтернет-ресурсів, з використанням запропонованих алгоритмів та методу, можуть використовуватися експертні системи прогнозування гібридного типу, призначені для використання на інформаційно - обчислювальній техніці загального призначення.

Реалізований в інформаційно-аналітичній системі алгоритм прогнозування вразливостей та загроз безпеки інформації на основі аналізу потоку даних тематичних інтернет-ресурсів дозволяє автоматизувати інформаційний процес виявлення нових вразливостей, загроз, надає фахівцям інформаційної безпеки можливість оцінити своєчасно ступінь захищеності ресурсів та при необхідності вжити відповідних заходів щодо нейтралізації можливих загроз та вразливостей, тим самим підвищити інформаційну безпеку обчислювальних комп'ютерних систем від реалізації нових мережевих комп'ютерних атак. Проведено аналіз систем нечіткого логічного виводу, сучасних засобів обробки великих об'ємів даних, засобів морфологічного аналізу тексту, редакторів онтологій. Для проведення логічного моделювання інформаційної системи прогнозування вразливостей та загроз інформаційної безпеки побудовані UML–діаграми діяльності, послідовності дій, класів. Для фізичного моделювання системи розроблено UML-діаграми розгортання та компонентів. Обґрунтовано можливість реалізації інформаційно-аналітичної системи прогнозування вразливостей та загроз безпеки інформації на основі аналізу текстових повідомлень тематичних інтернет-ресурсів з використанням наступних програмних продуктів: СКБД MySQL, редактор онтології – Protégé, системи нечіткого логічного виводу – Fuzzy Logic Designer, засобів морфологічного аналізу даних – Mystem. Для проведення оцінки отриманих результатів обчисленні показники MAPE, MAE, RMSE для значень прогнозування виникнення вразливостей та загроз інформаційної безпеки, а також розраховані на їх основі згладжені часові рядки з періодом три та п'ять діб.

Ключові слова: інформаційна безпека, тематичні інтернет-ресурси соціальні мережі, джерела повідомлень, вразливості, атаки, інформаційна система.

Вступ. На сучасному етапі, проблеми інформаційної безпеки розвитку суспільства у більшості сфер їх діяльності виходять на передній план. Це пов'язано зі значним зростанням кількості реалізованих проектів інформатизації. Більшість реалізованих проектів інформатизації спрямовані на побудову єдиного телекомунікаційного та інформаційного простору з метою оптимізації процесів обробки різноманітної інформації великих об'ємів, наприклад забезпечення оперативного доступу до інформації, надійного зберігання даних для користувачів інформаційного обміну [1-17].

Проблемою інформаційної безпеки суспільства є шкідлива інформація, злочинні та терористичні угрупування беруть на озброєння засоби інформаційного впливу, розробляють та пишуть стратегії, спрямовані на залучення нових адептів та розширення сфери впливу через

соціальні мережі. Однією зі складових надійного забезпечення інформаційної безпеки держави є проведення аналізу, виявлення, моніторинг та активна протидія розповсюдженню шкідливої інформації в соціальних мережах [1,2, 5 -7, 9,14,16].

Важливість проблеми пов'язана з наступними факторами: зростанням різноманітності та кількості засобів комп'ютерної техніки та сфер людської діяльності їх застосування; високим рівнем довіри до інформаційно-пошукових систем обробки та управління даними; зростанням числа користувачів інформаційного простору взаємодії; накопиченням великих об'ємів різнотипної інформації, інтенсивним обміном потоком даних в мережі між користувачами, з використанням широкого спектра механізмів доступу до конфіденційних ресурсів, інформаційних процесів; промисловим шпигунством та конкурентною боротьбою у сфері інформаційних послуг суспільства; недостатньою кількістю, на сучасному етапі, фахівців високої кваліфікації в області інформаційної безпеки, ринковими відношеннями в області розробки програмного забезпечення, обслуговування, розповсюдження, виробництва обчислювальної комп'ютерної техніки для реалізації інформаційної безпеки; різноманіттям атак, загроз і різнотипних каналів отримання несанкціонованого доступу до конфіденційних ресурсів та диференціацією негативних наслідків [3, 10-12,16,17].

Виникає потреба у проведенні захисту комп'ютерних систем та інформаційних ресурсів від блокування, несанкціонованого доступу до конфіденційних даних, знищення та інших злочинних, небажаних загроз, різноманіття та кількість яких постійно зростає. За оцінками, проведеними експертними організаціями, збитки в інформаційній сфері від злочинів в мережі Інтернет щорічно оцінюються в мільярди доларів [6,8,9,14,15].

Аналіз останніх досліджень та постановка задачі. З метою автоматизації широкого спектру вирішення задач застосовуються експертні системи та інформаційні технології. Аналіз проведених досліджень дозволяє зробити висновок, що для вирішення задачі дослідження та розробки інформаційно-аналітичної нечіткої системи для логічного нечіткого виводу про появу вразливостей та загроз інформаційної безпеки, автоматизації проведення аналізу потоку повідомлень тематичних інтернет-ресурсів, про доцільність використання експертних систем прогнозування. Для вирішення задач прогнозування вразливостей та загроз інформаційної безпеки даних на основі потоку тематичних повідомлень інтернет-ресурсів, з використанням запропонованих алгоритмів та методу, можуть використовуватися експертні системи прогнозування гібридного типу, призначені для використання на інформаційно - обчислювальній техніці загального призначення [1-4,7,9,14,16].

Прогнозування вразливостей та загроз інформаційної безпеки конфіденційним даним на основі отриманих повідомлень тематичних інтернет-форумів використання онтології предметної області відіграє ключову позицію. Успіх проведення аналізу повідомлень форуму залежить від способу побудови онтології предметної області. Робота з онтологіями передбачає використання методологій та методів їх побудови, із застосуванням прикладних інструментів та спеціалізованих мов програмування, також вирішення задач, пов'язаних з забезпеченням життєвого циклу та їхньою розробкою [1-4,8,9].

В якості основні програмних інструментів для розробки онтологій предметної області виступають редактори онтологій. Основна їхня функція - забезпечення можливостей формалізації знань про предметну область, для якої проводиться аналіз у заданому форматі онтологічної структури. До основних функцій сучасних редакторів онтології відносяться: імпорт онтології із зовнішніх форматів та експорт у потрібний формат; інтерактивна розробка онтології; редагування метаданих онтології (версії формалізації, загального опису, простору імен); робота з елементами онтології: видалення, редагування, створення відношень онтології, аксіом, об'єктів, класів. Функціональні можливості редакторів онтології можуть бути розширені, підключенням додаткових модулів та плагінів для візуалізації онтологій, несуперечності, перевірки логічної цілісності. Сучасні редактори онтологій розрізняються реалізованих у них наборами функцій, форматами представлення та зберігання даних, можливостями проведення модифікації вхідного коду [3,4,11,17].

Результати порівняння характеристик та параметрів сучасних та популярних редакторів онтологій наведені в табл. 1.

Таблиця 1

Таблиця характеристик редакторів онтології

№ п/п	Редактор онтології	Модель	Мова ПЗ	Мова представлення	Зберігання онтологій	Розширення
1	Protégé	Local	Java	OKBC	Файли, СКБД	Плагіни
2	OntoEdit	Local	Java	OXML	Файли	Плагіни
3	Oiled	Local	Java	DAWL+OIL	Файли	-

Найбільш популярним редактором онтологій - редактор Protégé. Архітектура редактора Protégé легко розширюється, вільно поширюється, підтримка модулів розширення, має відкритий вихідний код. Редактор Protégé використовується для розробки бази знань (онтології) вразливостей та загроз інформаційної безпеки конфіденційних даних, застосовується для проведення обчислення результатів експериментів, які є основою для проведення оцінки ефективності запропонованих алгоритмів та моделей, в основу яких покладено тезаурус інформаційної безпеки, класифікацію вразливостей і загроз [6-8,11,13].

Для реалізації семантичної фільтрації текстових повідомлень тематичних форумів інтернет-ресурсів проведено аналіз сучасних програмних інструментів, які забезпечують функціями семантичного та морфологічного аналізу текстових повідомлень, порівняльна характеристика яких наведена в табл. 2. Морфологічні процесори виконують відповідні функції лематизації словоформ.

Таблиця 2

Порівняльна характеристика програмних засобів морфологічного аналізу текстстових повідомлень

№ п/п	Система	Відкриті вихідні коди	Швидкість слів в сек.	Підключення словників	Обєм словника тис. слів
1	AOT	Так	60-90 тис.	Ні	160
2	MyStem	Ні	100-120 тис.	Так	>250
3	Pymorphy2	Так	80-100 тис.	Ні	250

Словник процесора TreeTagger доступний у вигляді бінарного файлу, закритим є словник системи MyStem. Швидкість обробки слів у наведених програмних платформ є достатньо високою. Для роботи з обмеженими предметними областями, особливо важливою задачею є можливість підключення словника даних. Дана функція реалізована у процесорі MyStem. Кожна морфологічна програмна система використовує власну систему морфологічних тегів, таким чином, у зв'язку з цим порівняти результати роботи процесорів на однакових потоках текстів достатньо складно. На рис. 1 наведено етапи нечіткого виводу роботи інформаційно-аналітичної системи прогнозування вразливостей та загроз інформаційної безпеки.

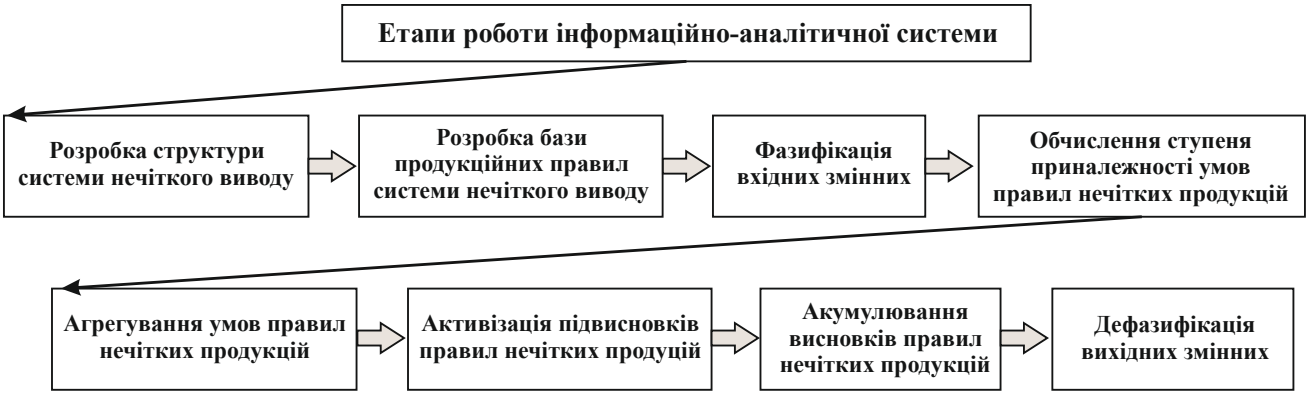

Рисунок 1 - Етапи нечіткого виводу роботи інформаційно-аналітичної системи

Інформаційно-аналітична система прогнозування вразливостей та загроз інформаційної безпеки. На теперішній час існують програмні інструменти, які надають функції збору текстових повідомлень, що розміщуються на тематичних інтернет-сервісах. Їх застосування в практичному використанню дозволяє реалізувати функції нечіткої інформаційно-аналітичної системи формування інтернет - потоку повідомлень різних дискусійних тематичних інтернет-ресурсів. Так як, форуми тематичних інтернет-ресурсів представляють сховища неформалізованих даних, щодо інформаційної безпеки та технологій, містять нечіткі поняття та знання (відсутні будь-які формати викладу текстових повідомлень, учасниками застосовується специфічний сленг, який є у користувачів інтернет-дискусій), доцільно, в даній ситуації використовувати для роботи з даними форумами механізми нечіткої логіки. Обґрунтованість застосування нечітких логічних моделей пов'язана зі значною часткою невизначеності потоку повідомлень, обумовленої складністю предметної області та неповнотою інформації. Як інструмент для досягнення поставленої задачі пропонується використовувати нечітку інформаційно-аналітичну систему прогнозування вразливостей та загроз інформаційної безпеки, що надає функціональні можливості, які представленні на рис. 2.

Ефективність роботи інформаційно-аналітичної системи прогнозування вразливостей та загроз інформаційної безпеки в значній мірі залежить від якості використовуваної у ній бази продкуційних правил (бази знань). База знань є сполучною ланкою між ключовими модулями системи та сховищем даних. До бази продукційних правил включено список тематичних форумів та онтологію предметної області [5,6,8,11,16,17].

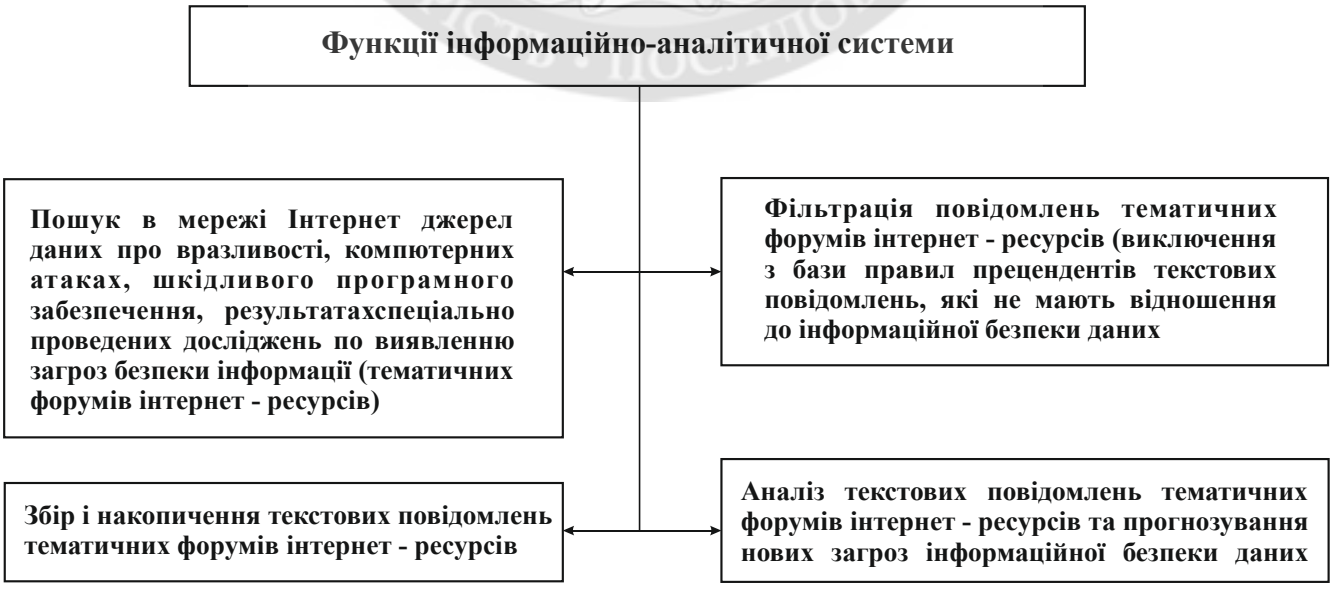

Рисунок 2 – Функції інформаційно-аналітичної системи аналізу потоку повідомлень тематичних інтернет-форумів

Організацію процесу аналізу потоку текстових повідомлень тематичних форумів інтернет-ресурсів та прогнозування вразливостей та загроз інформаційної безпеки конфіденційних даних відображає структура нечіткої інформаційно-аналітичної системи, яка представлена на рис. 3. Стрілками темного кольору позначені потоки текстових повідомлень інтернет-ресурсів в процесі пошуку джерел даних тематичних форумів, предметної області що представляє інтерес. Світлими стрілками (рис. 3) позначений потік повідомлень тематичних форумів інтернет-ресурсів у процесі прогнозування вразливостей та загроз інформаційної безпеки конфіденційним даним

Список тематичних джерел форумів містить адреси інтернет-ресурсів, на яких розміщуються текстові публікації про шкідливе програмне забезпечення, вразливості та комп'ютерні атаки [7,8,11,15,17].

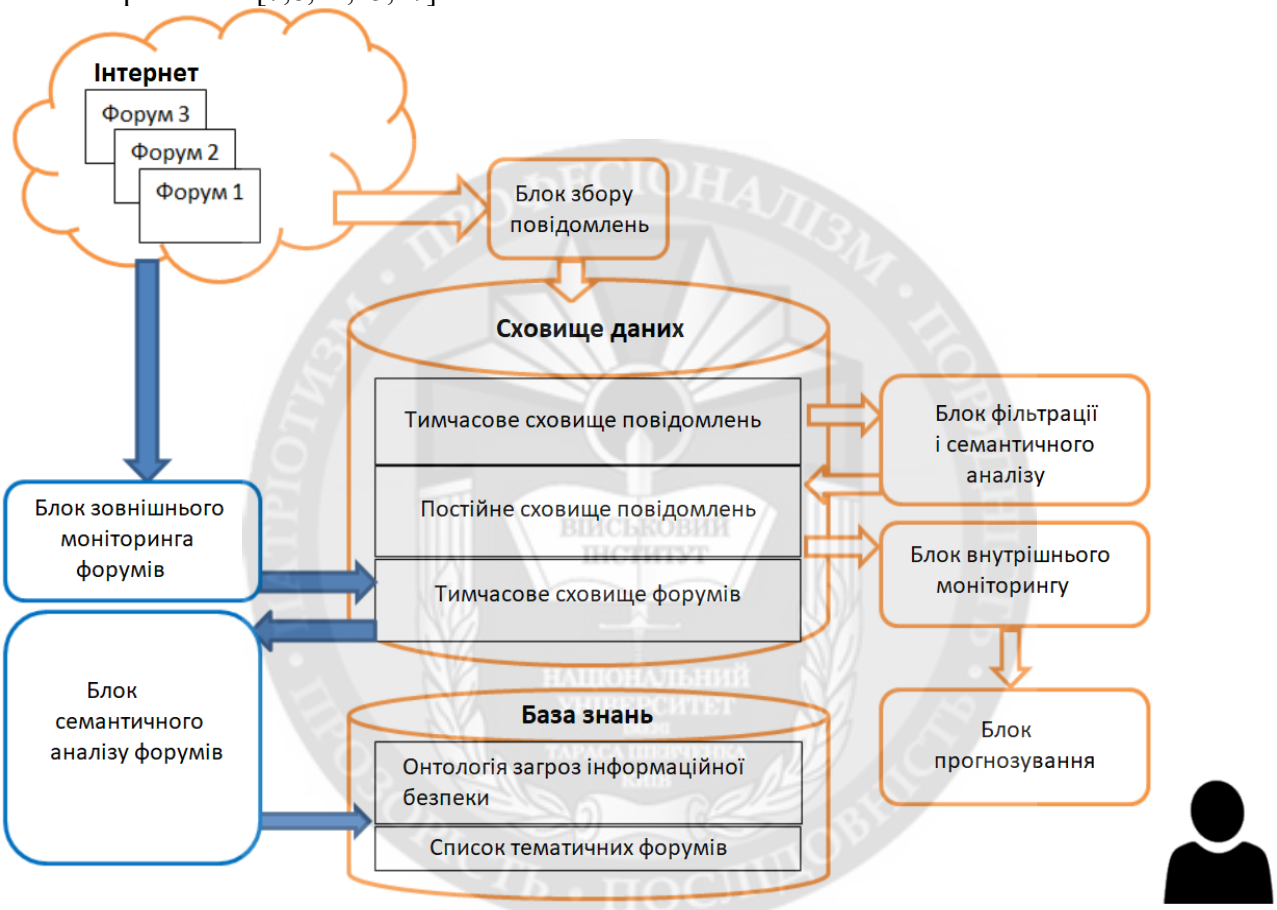

Рисунок 3 - Структура нечіткої інформаційно-аналітичної системи

На початковому етапі роботи інформаційно-аналітичної системи список формується експертним шляхом, із загальної кількості форумів тематичних інтернет–ресурсів виділяються ті, тематика яких дозволяє інтернет-інформацію віднести до хакерських (інформація містить результати спеціалізованих досліджень з виявлення вразливостей та загроз інформаційної безпеки конфіденційних даних, повідомлення про комп'ютерні атаки, вразливості, шкідливе програмне забезпечення). Автоматизоване виявлення нових форумів тематичних інтернет-ресурсів, в даній ситуації, можливо, шляхом проведення аналізу різноманітних форумів інтернет-ресурсів, з використанням запропонованих критеріїв відбору текстових повідомлень, що належать до заданої предметної області, для якої проводиться аналіз з використання онтології.

На теперішній час для вирішення задачі проектування та розробки інформаційно-обчислювальних систем використовується універсальна мова моделювання UML - дозволяє реалізувати об'єктно-орієнтований підхід до проектування систем, будувати моделі систем із зазначеннями їх основних якостей.

Для розробки концептуальної моделі інформаційно-обчислювальної системи застосовуються моделі бізнес-об'єктів: діаграми діяльності, варіантів використання, послідовностей дій. Під час проектування логічної моделі інформаційно-обчислювальної системи вимоги до системи формуються на основі застосування діаграм варіантів використання. На етапі попереднього проектування інформаційної системи використовуються діаграми послідовностей, станів, класів. На етапі проектування фізичної моделі інформаційної системи детальне проектування виконується із використанням діаграм класів, компонентів та розгортання. Для опису функціонального призначення інформаційної системи застосовуються діаграми UML варіанти використання. Діаграми є концептуальним представленням інформаційної системи (вхідними моделями) у процесі розробки та проектування. В залежності від розв'язуваних задач, роботу з інформаційною системою, можуть здійснювати користувачі двох типів:

1. Експерт, для роботи доступні чотири модулі інформаційної системи: редактор онтології (Protégé) - використовується для формалізації накопичуваних експертних знань в форматі онтології, що відносяться до заданої предметної області; редактор функцій приналежності вхідних та вихідних параметрів, правил логічних нечітких продукцій (Fuzzy Logic Designer) - використовується для здійснення логічного нечіткого виводу про вразливості та загрози інформаційної безпеки конфіденційних даних; підсистема розширення ядра онтології - для вилучення термінології з текстів предметної області, експерту надається можливість оцінювати вилучення на термінологічність і вносити їх до онтології предметної області, розширюючи базу знань правил продукцій; редактор списку тематичних форумів інтернет-ресурсів - використовується для формування потоку текстових повідомлень, що аналізуються.

2. Спеціаліст з інформаційної безпеки - необхідно прийняти рішення про достатність заходів, для здійснення захисту конфіденційних даних. Користувачу надається доступ до підсистеми логічного нечіткого виводу про виникнення вразливостей та загроз безпеки інформації, в основі лежить проведення аналізу текстових повідомлень тематичних форумів інтернет-ресурсів. Під час отримання виводу про виникнення вразливості чи загрози безпеки інформації, користувачу надана можливість оцінити актуальність безпеки для інформаційно-обчислювальної системи, та прийняти відповідні запобіжні заходи усунення негативних чинників. Нечітка інформаційно-аналітична система розробляється для обробки потоку інтернет повідомлень тематичних форумів інтернет-ресурсів, фільтрації текстових повідомлень на основі онтології заданої предметної області, проведення статистичного аналізу, сематичного аналізу потоку даних, логічного нечіткого виводу, базується на результатах потоку текстових повідомлень.

При використанні об'єктно-орієнтованого підходу розробки системи, центральне місце займає розробка моделей, представлених у вигляді діаграми класів. Діаграми класів розглядаються на початкових етапах розробки та моделювання інформаційної системи. Для представлення інформаційно-обчислювальної аналітичної системи задіяно сім класів: Автори повідомлень, Форуми, Онтологія, Теми, Розділи, Інформаційні ресурси, Повідомлення. На рис. 4. наведено діаграму класів інформаційно-аналітичної системи.

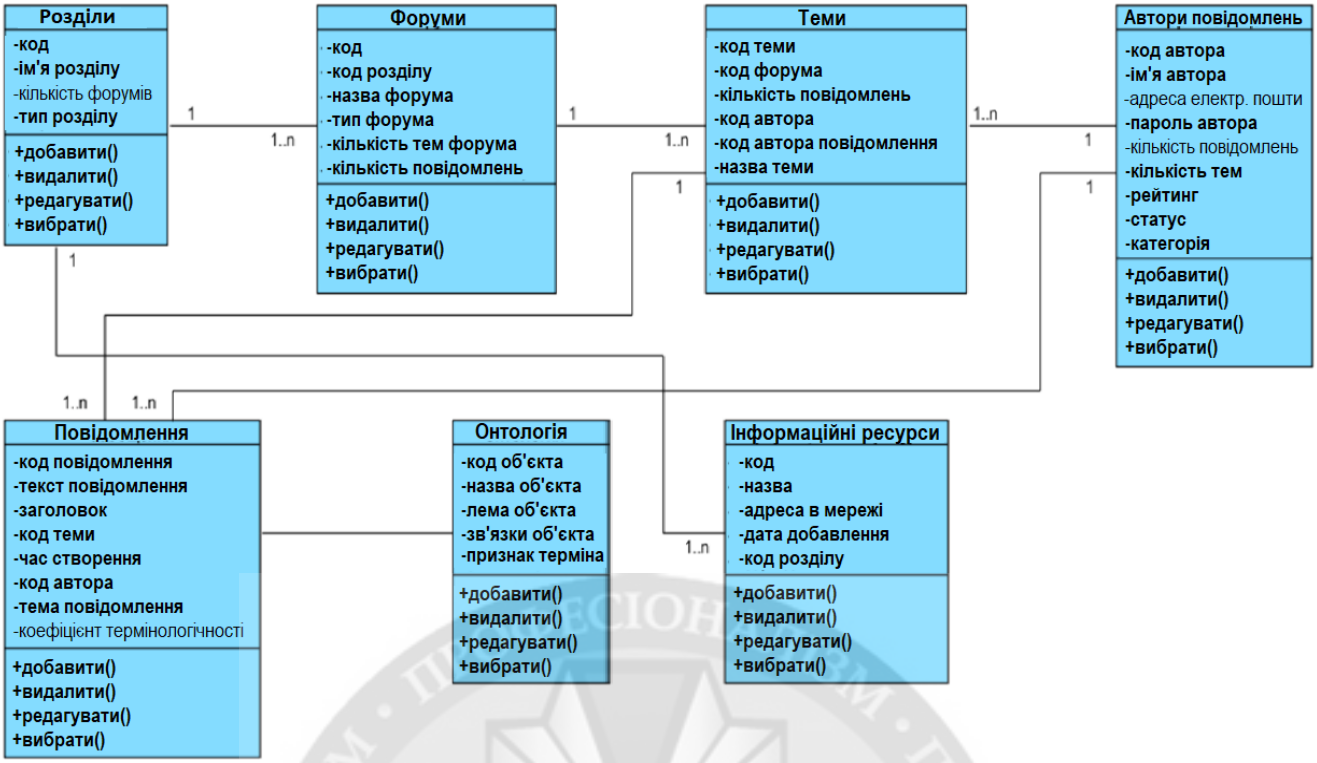

Рисунок 4 – Діаграма класів нечіткої інформаційно-аналітичної системи

Для опису взаємодії інформаційної системи об'єктів використанні діаграми послідовності дій. Діаграми описують послідовності, у яких об'єкти діаграми отримують та надсилають повідомлення на протязі часу. Діаграма діяльності інформаційно-обчислювальної аналітичної системи, наведена на рис. 5.

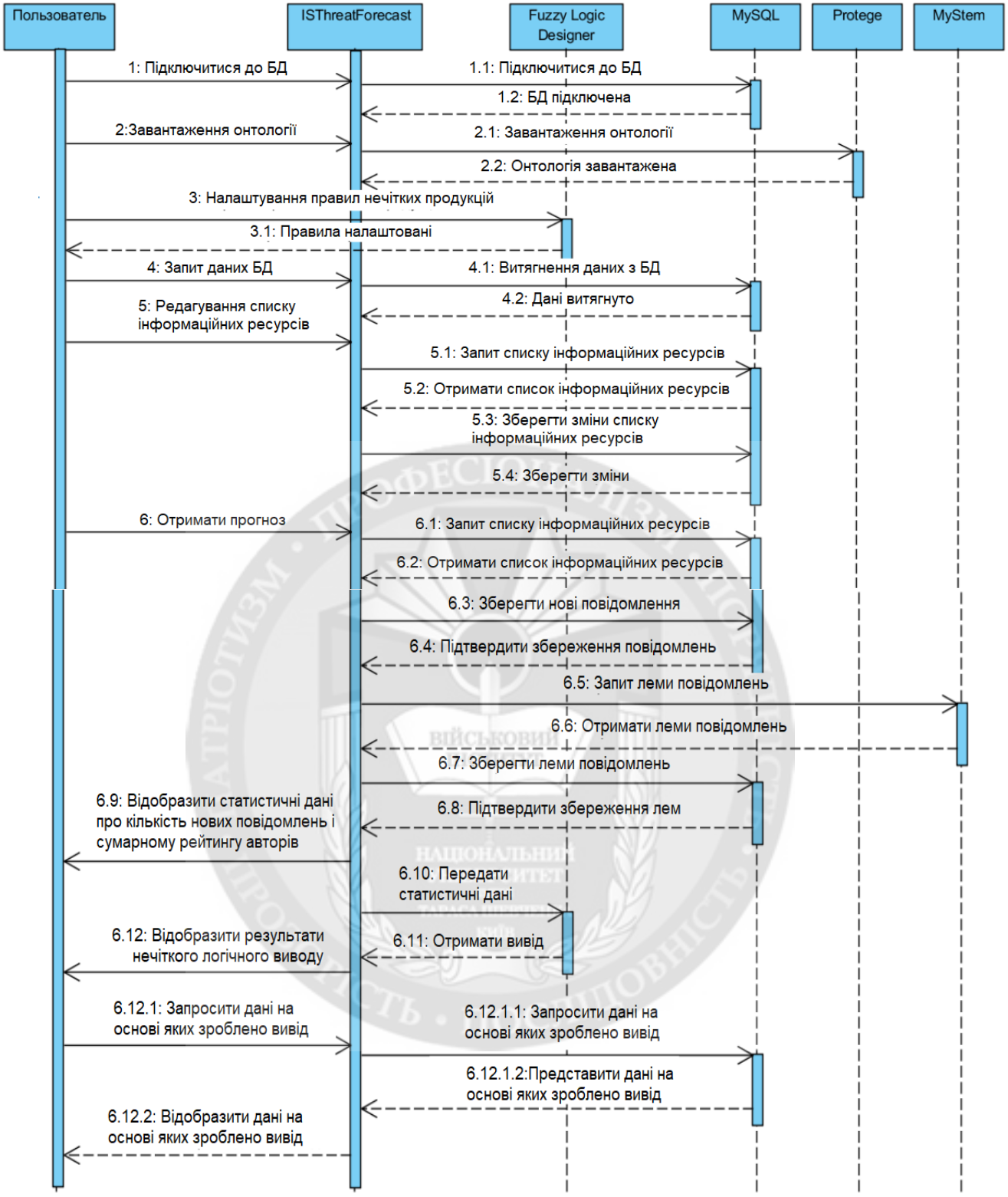

Рисунок 5 – Діаграма діяльності інформаційно-аналітичної системи

Діаграми послідовності визначають основні повідомлення, на які реагують об'єкти, компоненти, відображають динамічну складову інформаційно системи.

Запропонована діаграма описує логічну модель, в якій як засіб для реалізації нечіткого логічного виводу використовується Fuzzy Logic Designer, MySQL, редактор онтології – Protégé, засоби семантичної обробки MyStem.

Розглянуті діаграми реалізують концептуальну сторону побудови моделі інформаційно-аналітичної системи, подання системи здійснюється на логічному рівні. Для реалізації

фізичної системи, потрібно реалізувати в матеріальні сутності всі елементи логічного представлення.

Для фізичного представлення моделі використовується діаграма UML розгортання, відображається загальна топологія та конфігурація інформаційно-обчислювальної системи, а також розподіл за окремими вузлами компонентів. Вузлами діаграми інформаційно-обчислювальної системи є персональний комп'ютер користувача, сервер адміністратора. Вузли пов'язані суцільною лінією - наявність фізичного каналу обмінюватись інформацією, пунктирна лінія - вузли взаємодіють шляхом направлення різноманітних звернень та використання файлів обміну інформацією (рис. 6).

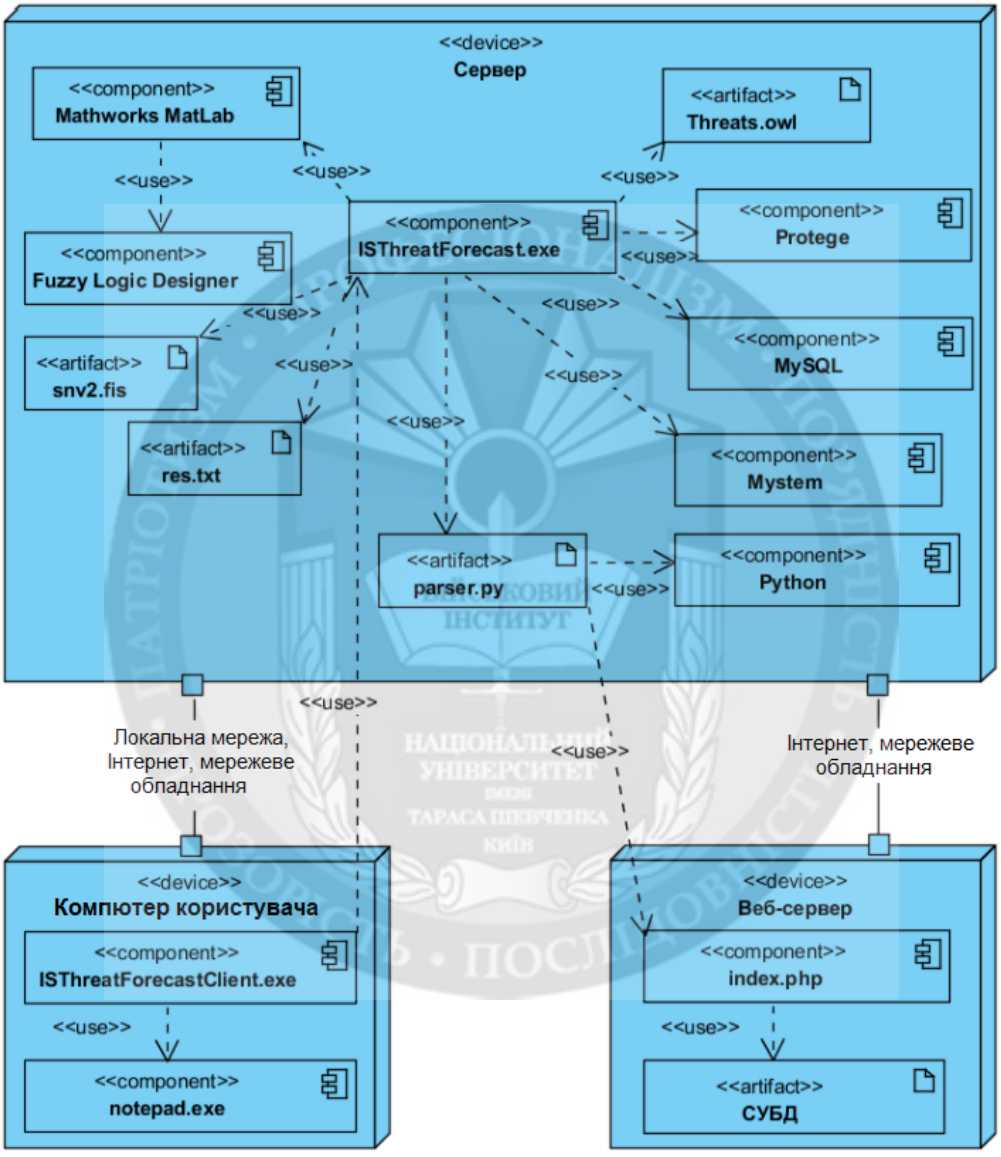

Рисунок 6 – Діаграма розгортання інформаційної системи

Інформаційно-аналітична система використовує онтологію інформаційної безпеки, побудовану на початковому етапі роботи системи, експертним шляхом на основі тезауруса. Розширення онтології здійснюється виявленням нових термінів у потоку текстових повідомлень хакерських термінів відповідно до запропонованих алгоритмів. Для реалізації онтології експертом використовується Protege, результат OWL-онтологія, яка завантажується до бази знань. Тематичні текстові повідомлення інтернет-ресурсів піддаються морфологічному аналізу з використанням Mystem, результати додаються до бази знань.

Оцінка системи прогнозування вразливостей та загроз інформаційної безпеки. Для оцінки ефективності та адекватності системи прогнозування вразливостей та загроз інформаційної безпеки, а також коректності роботи інформаційної системи, проведено експерименти, під час яких проводився аналіз потоку повідомлень форумів тематичних інтернет-ресурсів. В рамках експерименту проведені наступні дії [4,10,14,16,17]:

1. На основі тезауруса інформаційної безпеки, класифікацій вразливостей і загроз безпеки інформації, отриманої експериментальним шляхом, побудована онтологія вразливостей і загроз інформаційної безпеки конфіденційних даних.

2. Сформовано набір правил логічних нечітких продукцій - відображають закономірності залежності кількості створюваних на форумі тематичних інтернет-ресурсів текстових повідомлень та середнього рейтингу авторів від ймовірності виникнення вразливостей та загроз інформаційної безпеки конфіденційних даних. Запропоновано базу правил продукцій та визначено функції приналежності для інформаційної системи.

3. Здійснено збір текстових повідомлень, відібраних експертним шляхом тематичних форумів інтернет-ресурсів.

4. Результати отриманих ймовірностей виникнення вразливостей та загроз інформаційної безпеки співвіднесені зі значеннями кількості записів, включених до бази знань. Проведено розрахунки ефективності запропонованого алгоритму безпеки інформації.

5. Проведено обчислення показників кількості текстових повідомлень форумів тематичних інтернет-ресурсів, середнього рейтингу авторів повідомлень. Отримані результати використані в якості вхідних параметрів інформаційної системи логічного нечіткого виводу, обчисленні значення ймовірності виникнення загроз та вразливостей інформаційної безпеки;

6. Проведено статистичний та семантичний аналіз отриманих текстових повідомлень шляхом фільтрації даних, що не містять термінів заданої онтології інформаційної безпеки конфіденційних даних.

При вирішенні задачі оцінки якості прогнозування інформаційної системи, використовуються показники, що наведені в табл. 3.

Таблиця 3

Показники якості прогнозування аналітичної системи

№ п/п	Назва, формула, опис
1	*MAPE* – середня абсолютна процентна помилка системи прогнозування $$MAPE = \frac{1}{h}\sum_{i=1}^{h}\left|\frac{f_{T,i} - y_{T+i}}{y_{T+i}}\right| \cdot 100\%, \quad (1)$$ де h - довжина інтервалу, на якому проводиться прогнозування загроз; $f_{T,i}$ - прогнозне значення часового ряду, отримане в момент часу T на i кроків наперед; y_{T+i} - значення часового ряду в момент часу $T+i$
2	*MAE* - середня абсолютна помилка системи прогнозування: $$MAE = \frac{1}{h}\sum_{i=1}^{h}\left|f_{T,i} - y_{T+i}\right|, \quad (2)$$
3	*RMSE* - квадратний корінь із середньої квадратичної помилки системи прогнозування: $$RMSE = \sqrt{\frac{1}{h}\sum_{i=1}^{h}\left(f_{T,i} - y_{T+i}\right)^2}, \quad (3)$$

Для оцінки якості прогнозування загроз зручніше використання середньої абсолютної процентної помилка ($MAPE$), вимірюється у відсотках від значення прогнозованого показника. Показник може бути використаний для порівняння якості прогнозування загроз, систем побудованих із застосуванням різних моделей, також в якості прогнозування конкретних моделей, для яких визначено рівень помилки прогнозування критичний.

Для оцінки ефективності методу прогнозування вразливостей та загроз безпеки інформації на основі проведеного аналізу текстових повідомлень учасників тематичних форумів інтернет-ресурсів проведено експерименти з автоматизованого збору повідомлень інтернет-ресурсів. При обчисленні функцій приналежності вхідних параметрів використовувалися результати проведеного аналізу текстових повідомлень. На основі отриманих результатів із застосуванням запропонованого методу, проведено обчислювальні експерименти щодо формування логічного нечіткого виводу про виникнення вразливостей та загроз інформаційної безпеки конфіденційних даних.

Для оцінки отриманих результатів проведено розрахунки показників *MAPE, MAE, RMSE* (за формулами 1, 2, 3) для значень прогнозування інформаційно-аналітичної нечіткої системи про виникнення вразливостей та загроз безпеки інформації та кількості виявлених вразливостей та загроз, в період проведення аналізу, а також проведені розрахунки, згладжених часових рядів із інтервалом згладжування три та п'ять діб. Результати представлені у табл. 4.

Таблиця 4

Показники якості прогнозування загроз

Показник	Експериментальні дані	Згладжування з періодом 3 доби	Згладжування з періодом 5 діб
MAPE (%)	94,21	147,04	119,73
MAE (%)	26,87	17,82	13,14
RMSE (%)	17,14	12,27	10,01

Проведено розрахунок показника точності прогнозування загроз η, для довірчих інтервалів 10, 15, 20%. Результати обчислень наведено у табл. 5.

$$\eta = \frac{p}{p+q}, \tag{4}$$

де p – число випадків прогнозування, які підтверджені фактичними даними; q – число випадків, які не знайшли фактичного підтвердження.

Таблиця 5

Показника точності прогнозування загроз

Довірчий інтервал	Експериментальні дані	Згладжування з періодом 3 доби	Згладжування з періодом 5 діб
10%	0,517	0,189	0,114
15%	0,554	0,559	0,482
20%	0,683	0,695	0,696

Наведені показники дозволяють зробити висновок, що результати прогнозування інформаційно-аналітичної нечіткої системи в більшості випадків підтверджуються даними бази знань вразливостей та загроз.

Таким чином, спеціаліст із інформаційної безпеки, на основі отриманих результатів прогнозування вразливості або загрози, може оцінити ступінь небезпеки інформаційних ресурсів організації та вжити відповідних заходів щодо нейтралізації загроз та вразливостей.

Покращення якості прогнозування виникнення вразливостей та загроз безпеки інформації з використанням систем логічного нечіткого виводу може сприяти збільшення кількості вхідних змінних, використанню більш точних нечітких правил продукцій, також велике значення має визначення функцій приналежності вихідних та вхідних параметрів системи логічного нечіткого виводу, необхідно враховувати статистичні показники потоку текстових повідомлень форумів тематичних інтернет-ресурсів.

Висновки. Запропонованого структурну схему інформаційної системи для прогнозування вразливостей та загроз безпеки інформації. Для проведення логічного моделювання інформаційно системи побудовані UML–діаграми діяльності, послідовності дій, класів. Для фізичного моделювання системи розроблено UML-діаграми розгортання та компонентів.

Проведено аналіз систем нечіткого логічного виводу, сучасних засобів обробки великих об’ємів даних, засобів морфологічного аналізу тексту, редакторів онтологій.

Обґрунтовано можливість реалізації інформаційно-аналітичної системи прогнозування вразливостей та загроз безпеки інформації на основі аналізу текстових повідомлень тематичних інтернет-ресурсів з використанням наступних програмних продуктів: СКБД MySQL, редактора онтології – Protégé, системи нечіткого логічного виводу – Fuzzy Logic Designer, засобів морфологічного аналізу даних – Mystem.

Реалізований в інформаційно-аналітичній системі метод прогнозування вразливостей та загроз безпеки інформації на основі дослідження потоку даних тематичних ресурсів дозволяє автоматизувати інформаційний процес виявлення нових вразливостей, загроз, надає фахівцям інформаційної безпеки можливість оцінити своєчасно ступінь захищеності ресурсів та при необхідності вжити відповідних заходів щодо нейтралізації можливих загроз та вразливостей, тим самим підвищити інформаційну безпеку обчислювальних комп'ютерних систем від реалізації нових мережевих комп'ютерних атак.

Для проведення оцінки отриманих результатів обчисленні показники *MAPE, MAE, RMSE* для значень прогнозування виникнення вразливостей та загроз інформаційної безпеки, а також розраховані на їх основі згладжені часові рядки з періодом три і п’ять діб.

ЛІТЕРАТУРА:

1. Ленков, С.В. Метод прогнозування вразливостей інформаційної безпеки на основі аналізу даних тематичних інтернет-ресурсів / С.В. Лєнков, В.М. Джулій, А.М. Берназ, І.В. Муляр, І.В. Пампуха // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – К.: ВІКНУ, 2023. – Вип. №78. – С. 123-134.

2. Ленков, С.В. Метод протидії поширенню та виявлення шкідливої інформації в соціальних мережах/ С.В. Ленков, В.М. Джулій, Л.В. Солодєєва // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – К.: ВІКНУ, 2022. – Вип. №77. – С. 103-117.

3. Ленков, С.В. Модель безпеки поширення забороненої інформації в інформаційно-телекомунікаційних мережах / С.В. Ленков, В.М. Джулій, В.С. Орленко, О.В. Сєлюков, А.В. Атаманюк // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – К.: ВІКНУ, 2020. – Вип. №68. – С. 53-64.

4. Джулій, В.М. Модель потоку текстових повідомлень тематичних інтернет-ресурсів системи прогнозування інформаційної безпеки / В. Джулій, Н. Петляк, Ю. Хмельницький, О. Пахар // Вісник Хмельницького національного університету. Технічні науки. – 2022. – № 5. – С. 294-300.

5. Lienkov, S., Podlipaiev, V., Tolok, I., Lisitsky I., .Lytvynenko, N., Kuznichenko, S. The Information and Analytical Using of Non-Structured Information Resources CEUR Workshop Proceedingsthis link is disabled, 2021, 3126, pp. 81–87.

6. Соціальні мережі – реальні загрози віртуального світу. [Електронний ресурс]. – Режим доступу : http://ogo.ua/ articles/view/011–02–23/26490.htm.

7. Ленков, С.В. Методы и средства защиты информации. В 2-х томах /С.В. Ленков, Д.А. Перегудов, В.А. Хорошко –К: Арий, 2008. – 464 с.

8. Остапов С. Е. Технології захисту інформації: навчальний посібник / С.Е. Остапов, С.П. Євсеєв, О.Г. Король – Харків : Вид-во ХНЕУ, 2016. – 476 с.

9. Лєнков, С.В. Аналіз існуючих методів та алгоритмів виявлення атак в бездротових мережах передачі даних / С.В. Лєнков, В.М. Джулій, Н.М. Берназ, С.О. Божук // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – К.: ВІКНУ, 2017. – Вип. № 56. – С.124-132

10. Джулій, В.М. Інформаційно-ознакова модель шкідливої інформації в соціальних мережах/ І.В. Муляр, В.М. Джулій, В. М. Пічура, О.О Зацепіна – Вимірювальна та обчислювальна техніка в технологічних процесах № 3 (2022), - 73–78 с.

11. Джулій, В.М., Кльоц Ю.П., Муляр І.В., Жилевич М.Л., Джулій А.В. Контроль додатків інтернет-трафіка комп'ютерних мереж методами машинного навчання. Вісник Хмельницького національного університету. Технічні науки. 2021. № 5. С. 22-26.

12. Джулій, В.М. Метод класифікації додатків трафіка комп'ютерних мереж на основі машинного навчання в умовах невизначеності / В.М. Джулій, О.В. Мірошніченко, Л.В. Солодєєва // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – К.: ВІКНУ, 2022. – Вип. №74. – С. 73-82.

13. Лавров, Є. А. Математичні методи дослідження операцій : підручник / Є. А. Лавров, Л. П. Перхун, В. В. Шендрик – Суми : Сумський державний університет, 2017. – 212 с.

14. Гончар С. Ф. Оцінювання ризиків кібербезпеки інформаційних систем об'єктів критичної інфраструктури : монографія. / С. Ф. Гончар. – Київ, 2019. – 175 с.

15. Yemchuk L. Organizational Network Analysis as a Tool for Leadership Assessment in Software Development Team. Zhylinska O.; Chornyi A.; Dzhuliy V. – Institute of Electrical and Electronics Engineers (30 September 2020); INSPEC Accession Number: 20008165; DOI: 10.1109/ACIT49673.2020.

16. Сигнатура атаки. Wikipedia [Електронний ресурс] – Режим доступу до ресурсу: https://uk.wikipedia.org/wiki/Сигнатура_атаки.

17. OPWNAI: Cybercriminals Starting to Use ChatGPT, January 6, 2023 [Електронний ресурс] – Режим доступу до ресурсу: https://research.checkpoint.com/2023/opwnai-cybercriminals-starting-to-usechatgpt.

REFERENCES:

1. Lenkov, S.V.(2023), Metod prohnozuvannia vrazlyvostei informatsiinoi bezpeky na osnovi analizu danykh tematychnykh internet-resursiv / S.V. Lienkov, V.M. Dzhulii, A.M. Bernaz, I.V. Muliar, I.V. Pampukha // Zbirnyk naukovykh prats Viiskovoho instytutu Kyivskoho natsionalnoho universytetu imeni Tarasa Shevchenka. – K.: VIKNU -. №78. – pp. 123-134.

2. Lenkov, S.V.(2022) Metod protydii poshyrenniu ta vyiavlennia shkidlyvoi informatsii v sotsialnykh merezhakh/ S.V. Lenkov, V.M. Dzhulii, L.V. Solodieieva // Zbirnyk naukovykh prats Viiskovoho instytutu Kyivskoho natsionalnoho universytetu imeni Tarasa Shevchenka. – K.: VIKNU. – Vyp. №77. – pp. 103-117.

3. Lenkov, S.V. (2020), Model bezpeky poshyrennia zaboronenoi informatsii v informatsiino-telekomunikatsiinykh merezhakh / S.V. Lenkov, V.M. Dzhulii, V.S. ORLENKO, O.V. Sieliukov, A.V. Atamaniuk // Zbirnyk naukovykh prats Viiskovoho instytutu Kyivskoho natsionalnoho universytetu imeni Tarasa Shevchenka. – K.: VIKNU. – №68. – pp. 53-64.

4. Dzhulii, V.M. (2022.), Model potoku tekstovykh povidomlen tematychnykh internet-resursiv systemy prohnozuvannia informatsiinoi bezpeky / V. Dzhulii, N. Petliak, Yu. Khmelnytskyi, O. Pakhar // Visnyk Khmelnytskoho natsionalnoho universytetu. Tekhnichni nauky. – 2022. – № 5. – pp. 294-300.

5. Lienkov, S., Podlipaiev, V., Tolok, I., Lisitsky I., .Lytvynenko, N., Kuznichenko, S. (2021). The Information and Analytical Using of Non-Structured Information Resources CEUR Workshop Proceedingsthis link is disabled, 3126, pp. 81–87.

6. Cotsialni merezhi – realni zahrozy virtualnoho svitu. [Elektronnyi resurs]. – Rezhym dostupu : http://ogo.ua/ articles/view/011– 02–23/26490.htm

7. Lenkov, S.V. (2008), Metodyy sredstva zashchytы ynformatsyy. V 2-kh tomakh / S.V. Lenkov, D.A. Perehudov, V.A. Khoroshko –K: Aryi–464 p.

8. Ostapov, S. E. (2016) Tekhnolohii zakhystu informatsii: navchalnyi posibnyk / S.E. Ostapov, S.P. Yevseiev, O.H. Korol–Kharkiv : Vyd-vo KhNEU. – 476 p.

9. Lenkov, S.V. (2017), AnalIz Isnuyuchih metodiv ta algoritmiv viyavlennya atak v bezdrotovih merezhah peredachI danih / S.V. Lenkov, V.M. Dzhuliy, N.M. Bernaz, S.O. Bozhuk // Zbirnik naukovih prats Viyskovogo Institutu Kiyivskogo natsionalnogo universitetu imeni Tarasa Shevchenka. – K.: VIKNU. – Vip. No 56. – pp.124-132.

10. Dzhulii, V.M. (2022). Informatsiino-oznakova model shkidlyvoi informatsii v sotsialnykh merezhakh/ I.V. Muliar, V.M. Dzhulii, V. M. Pichura, O.O Zatsepina – Vymiriuvalna ta obchysliuvalna tekhnika v tekhnolohichnykh protsesakh - pp. 373–78.

11. Dzhulii V.M., Klots Yu.P., Muliar I.V., Zhylevych M.L., Dzhulii A.V. (2021), Kontrol dodatkiv internet-trafika kompiuternykh merezh metodamy mashynnoho navchannia. Visnyk Khmelnytskoho natsionalnoho universytetu. Tekhnichni nauky. – Khmelnytskyi. – No 5. – pp. 22–26.

12. Dzhulii, V.M. (2022), Metod klasyfikatsii dodatkiv trafika kompiuternykh merezh na osnovi mashynnoho navchannia v umovakh nevyznachenosti / V.M. Dzhulii, O.V. Miroshnichenko, L.V. Solodieieva // Zbirnyk naukovykh prats Viiskovoho instytutu Kyivskoho natsionalnoho universytetu imeni Tarasa Shevchenka. – K.: VIKNU. – Vyp. No 74. – pp. 73-82.

13. Lavrov, Ye. A. (2017.), Matematychni metody doslidzhennia operatsii : pidruchnyk / Ye. A. Lavrov, L. P. Perkhun, V. V. Shendryk – Sumy : Sumskyi derzhavnyi universytet, – 212 p.

14. Informatsiino-komunikatyvni tekhnolohii v humanitarnii sferi Zbroinykh Syl Ukrainy: dosvid, problemy, perspektyvy: Pidruchnyk. – Kyiv: NAOU, 2007.

15. Yemchuk L. Organizational Network Analysis as a Tool for Leadership Assessment in Software Development Team. Zhylinska O.; Chornyi A.; Dzhuliy V. – Institute of Electrical and Electronics Engineers (30 September 2020); INSPEC Accession Number: 20008165; DOI: 10.1109/ACIT49673.2020.

16. Syhnatura ataky. Wikipedia [Elektronnyi resurs] – Rezhym dostupu do resursu: https://uk.wikipedia.org/wiki/Syhnatura_ataky.

17. OPWNAI: Cybercriminals Starting to Use ChatGPT, January 6, 2023 [Elektronnyi resurs] – Rezhym dostupu do resursu: https://research.checkpoint.com/2023/opwnai-cybercriminals-starting-to-usechatgpt.

D.Sci.Tech., prof. **Lienkov S.V.,** Ph.D. **Dzhuliy V.M.**
Ph.D. **Miroshnichenko O.V.**, Ph.D. **Brayn V.O., Prokhorskyi S.I.**

INFORMATION AND ANALYTICAL FORECASTING SYSTEMS INFORMATION SECURITY VULNERABILITIES AND THREATS

The paper proposes a block diagram of an information-analytical system for predicting vulnerabilities and threats to information security. The analysis of the conducted research allows us to conclude that in order to solve the problem of research and development of an information-analytical fuzzy system for a logical fuzzy conclusion about the emergence of vulnerabilities and threats to information security, automating the analysis of the message flow of thematic Internet resources, it is advisable to use expert forecasting systems. To solve the problems of predicting vulnerabilities and threats to information security of confidential data based on the flow of thematic messages of Internet resources using the proposed algorithms and method, hybrid-type expert forecasting systems designed for use on general-purpose information and computer technology can be used.

The algorithm for predicting vulnerabilities and threats to information security implemented in the information and analytical system based on the analysis of the data flow of thematic Internet resources allows automating the information process of detecting new vulnerabilities and threats, provides information security specialists with the opportunity to assess the degree of security of resources in a timely manner and, if necessary, take appropriate measures for neutralization possible threats and vulnerabilities, thereby increasing the information security of computing computer systems against the implementation of new network computer attacks. An analysis of fuzzy logical inference systems, modern tools for processing large volumes of data, tools for morphological text analysis, and ontology editors was conducted. UML diagrams of activities, sequences of actions, and classes were built to carry out logical modeling of the information system for forecasting vulnerabilities and threats to information security. For the physical modeling of the system, UML-diagrams of deployment and components have been developed. The possibility of implementing an information-analytical system for predicting vulnerabilities and threats to information security based on the analysis of text messages of thematic Internet resources using the following software products is substantiated: DBMS MySQL, ontology editor - Protégé, fuzzy logic inference system - Fuzzy Logic Designer, morphological data analysis tools - Mystem. To evaluate the obtained results, the indicators MAPE, MAE, RMSE for the values of forecasting the occurrence of vulnerabilities and information security threats, as well as smoothed time series calculated on their basis with a period of three and five.

Key words: information security, thematic Internet resources, social networks, sources of messages, vulnerabilities, attacks, information system.