НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ЮРИДИЧНА АКАДЕМІЯ»
ФАКУЛЬТЕТ КІБЕРБЕЗПЕКИ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра кібербезпеки

КВАЛІФІКАЦІЙНА РОБОТА

на тему:

«ЗАПОБІГАННЯ ВИТОКУ КОНФІДЕНЦІЙНОЇ ІНФОРМАЦІЇ З ІНФОРМАЦІЙНОЇ СИСТЕМИ НА ОСНОВІ ТЕХНОЛОГІЇ DLP»

Виконав: здобувач 4 курсу
Рівень бакалавр
Галузь знань 12 «Інформаційні технології»,
спеціальність 125 «Кібербезпека»
Поліщук Амір Алжанбегович

Керівник: к.ф.-м.н., доцент
Чепурна Олена Євгенівна

Рецензент: д.т.н., професор
Казакова Надія Феліксівна

Одеса – 2025

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ЮРИДИЧНА АКАДЕМІЯ»
Факультет кібербезпеки та інформаційних технологій
Кафедра **кібербезпеки**
Галузь знань: **12 Інформаційні технології**
Спеціальність: **125 Кібербезпека**
Назва освітньої програми: **Кібербезпека**

ЗАТВЕРДЖУЮ
Завідувач кафедри кібербезпеки
професор С.А. Горбаченко
_____ «____ »_____ 20__ року

З А В Д А Н Н Я
на кваліфікаційну (бакалаврську) роботу
здобувачу Поліщуку Аміру Алжанбеговичу

1. Тема роботи: «Запобігання витоку конфіденційної інформації з інформаційної системи на основі технології DLP»
Керівник роботи: к.ф.-м.н., доцент Чепурна Олена Євгенівна
затверджені наказом НУ ОЮА від «18» березня 2025 року № 548-6
2. Строк подання здобувачем роботи «19» травня 2025 року
3. Дата видачі завдання «16» вересня 2024 року

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів бакалаврської роботи	Строк виконання етапів роботи	Примітка
1	Аналіз предметної області та пошук науково-технічної інформації за темою роботи	Вересень-жовтень 2024	
2	Узгодження теми з науковим керівником, формулювання мети завдань дослідження	Листопад 2024	
3	Робота над першим розділом	Листопад-грудень 2024	
4	Робота над другим розділом	Січень-лютий 2025	
5	Робота над третім розділом	Лютий-березень 2025	
6	Уточнення подальшого обсягу роботи та його коригування	Березень-травень 2025	
7	Оформлення звітної частини	Травень 2025	
8	Захист роботи	10.06.2025	

Здобувач_____
(підпис) (прізвище та ініціали)

Керівник роботи_____
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Кваліфікаційна робота на тему “Запобігання витоку конфіденційної інформації з інформаційної системи на основі технології DLP” на здобуття першого (бакалаврського) рівня вищої освіти за спеціальністю 125 Кібербезпека, освітньою програмою: Кібербезпека, містить 5 рисунки, 6 таблиці, 46 літературних джерел за переліком посилань. Робота виконана на 50 сторінках загального тексту і 40 сторінці основного тексту.

Метою роботи є розробка теоретико-методологічних основ і практичних рекомендацій для запобігання витоку конфіденційної інформації з інформаційних систем шляхом впровадження адаптивної DLP-системи, що інтегрує сучасні технології та ризик-орієнтований підхід. Розроблено класифікацію конфіденційної інформації та загроз витоку, створено модель впровадження DLP-системи на основі контентного та контекстного аналізу з використанням бібліотек Scapy, iptables і pandas. Програмно реалізовано механізми моніторингу мережевого трафіку, виявлення чутливих даних і блокування несанкціонованих передач, проведено експериментальну оцінку ефективності, яка показала скорочення часу реагування до 10 секунд і рівень відбитих загроз 98%.

Результати роботи можуть бути використані підприємствами фінансового, промислового та державного секторів для зниження ризиків витоку даних, забезпечення відповідності стандартам ISO 27001, GDPR, PCI DSS і підвищення стійкості до кіберзагроз. Рекомендації також можуть застосовуватися в освітніх програмах із кібербезпеки та для вдосконалення корпоративних політик інформаційної безпеки.

Можливими напрямками подальших досліджень є інтеграція машинного навчання для прогнозування аномалій, впровадження квантово-стійкої криптографії та адаптація DLP-систем до 5G-мереж і хмарних технологій.

КІБЕРБЕЗПЕКА, DLP-ТЕХНОЛОГІЇ, КОНФІДЕНЦІЙНА ІНФОРМАЦІЯ, МОНІТОРИНГ, ЗАХИСТ ДАНИХ.

ABSTRACT

Qualification work on the topic “Preventing the Leakage of Confidential Information from Information Systems Based on DLP Technology” for obtaining the first (bachelor’s) level of higher education in the specialty 125 Cybersecurity, educational program: Cybersecurity, contains 5 figures, 6 tables, and 46 literary sources listed in the references. The work spans 50 pages of total text and 40 pages of main text.

The purpose of the work is to develop theoretical and methodological foundations and practical recommendations for preventing the leakage of confidential information from information systems by implementing an adaptive DLP system that integrates modern technologies and a risk-oriented approach. A classification of confidential information and leakage threats was developed, and a model for implementing a DLP system based on content and contextual analysis using Scapy, iptables, and pandas libraries was created. Mechanisms for monitoring network traffic, detecting sensitive data, and blocking unauthorized transmissions were programmatically implemented. An experimental evaluation of effectiveness demonstrated a reduction in response time to 10 seconds and a 98% rate of repelled threats.

The results of the work can be utilized by enterprises in the financial, industrial, and public sectors to reduce data leakage risks, ensure compliance with ISO 27001, GDPR, and PCI DSS standards, and enhance resilience to cyber threats. The recommendations can also be applied in cybersecurity educational programs and for improving corporate information security policies.

Possible directions for further research include integrating machine learning for anomaly prediction, implementing quantum-resistant cryptography, and adapting DLP systems to 5G networks and cloud technologies.

CYBERSECURITY, DLP TECHNOLOGIES, CONFIDENTIAL INFORMATION, MONITORING, DATA PROTECTION.

ЗМІСТ

ВСТУП

Актуальність теми. У сучасному цифровому світі стрімкий розвиток інформаційних технологій супроводжується зростанням ризиків витоку конфіденційної інформації, що становить серйозну загрозу для організацій, державних установ і суспільства загалом. За даними звіту Verizon Data Breach Investigations Report, 68% кіберінцидентів у 2023 році були пов'язані з витоком даних, а глобальні збитки від таких інцидентів перевищили 4,5 трлн дол. США. В Україні, де кіберзагрози посилюються через геополітичні виклики та активну цифровізацію, захист конфіденційної інформації набуває стратегічного значення. Особливо вразливими є інформаційні системи підприємств, які обробляють чутливі дані, такі як персональна інформація, комерційна таємниця чи державні секрети.

Існуючі методи захисту інформації, такі як стандарти ISO/IEC 27001 чи NIST SP 800-171, надають загальні рекомендації щодо управління інформаційною безпекою, але часто не враховують специфіку сучасних кіберзагроз, зокрема інсайдерських витоків, атак із використанням соціальної інженерії чи складних багатоканальних схем. Критичний аналіз літератури (зокрема праць Gartner, Ponemon Institute, McAfee) свідчить, що традиційні підходи до захисту даних є недостатньо ефективними в умовах динамічного розвитку технологій і зростання обсягів інформації. Технології Data Loss Prevention пропонують проактивний підхід до запобігання витоку даних, але їхнє впровадження потребує адаптації до специфіки інформаційних систем і комплексного підходу до управління ризиками. Таким чином, актуальність дослідження зумовлена необхідністю розробки ефективних методів і практичних рішень для запобігання витоку конфіденційної інформації з інформаційних систем на основі DLP-технологій, що відповідають сучасним викликам кібербезпеки.

Мета дослідження полягає у розробці теоретико-методологічних основ і практичних рекомендацій для забезпечення захисту конфіденційної інформації в інформаційних системах шляхом впровадження адаптивної DLP-системи, яка

інтегрує сучасні технології та ризик-орієнтований підхід. Для досягнення мети визначено такі завдання:

- визначити поняття конфіденційної інформації, її класифікацію та основні характеристики.
- ідентифікувати основні загрози витоку інформації та розробити методи їх виявлення.
- описати принципи роботи DLP-технологій та їхні ключові компоненти.
- розробити методи виявлення та блокування витоків конфіденційної інформації в DLP-системах.
- провести порівняльний аналіз сучасних DLP-рішень, доступних на ринку.
- визначити етапи впровадження DLP-системи в інформаційну інфраструктуру підприємства.
- сформувати практичні аспекти налаштування та моніторингу DLP-рішень.
- розробити методи оцінки ефективності DLP-систем та запропонувати підходи до вдосконалення захисту даних.

Об’єкт дослідження – процеси забезпечення захисту конфіденційної інформації в інформаційних системах, що охоплюють ідентифікацію, класифікацію, моніторинг і запобігання витоку даних у контексті їхньої критичної ролі для функціонування організацій.

Предмет дослідження – методи, технології та організаційні підходи до запобігання витоку конфіденційної інформації з інформаційних систем на основі DLP-технологій, спрямовані на підвищення безпеки та стійкості до кіберзагроз.

Практичне значення отриманих результатів. Результати дослідження мають практичну цінність для підприємств і організацій, які обробляють конфіденційну інформацію. Розроблена модель DLP-системи, що включає алгоритми моніторингу та блокування несанкціонованого передавання даних, може бути впроваджена в різних секторах, таких як фінансовий, промисловий чи державний, для зниження ризиків витоку інформації. Рекомендації щодо налаштування DLP-систем дозволять скоротити час реагування на інциденти і зменшити фінансові втрати.

Отримані результати можуть бути використані для вдосконалення корпоративних політик інформаційної безпеки, розробки навчальних програм для фахівців із кібербезпеки та адаптації національних стандартів захисту даних. Запропоновані рішення є універсальними й можуть застосовуватися як в Україні, так і за її межами.

1 ТЕОРЕТИЧНІ ОСНОВИ ЗАПОБІГАННЯ ВИТОКУ КОНФІДЕНЦІЙНОЇ ІНФОРМАЦІЇ

1.1 Поняття конфіденційної інформації та її класифікація

Конфіденційна інформація є одним із центральних понять у сфері інформаційної безпеки та кібербезпеки. Вона визначається як відомості або дані, які належать особі, організації чи державі, не є загальнодоступними та потребують захисту від несанкціонованого доступу, розголошення чи використання. Відповідно до міжнародних стандартів, зокрема ISO/IEC 27001:2013, конфіденційна інформація характеризується трьома ключовими властивостями: конфіденційністю, цілісністю та доступністю, що разом формують так звану тріаду CIA (Confidentiality, Integrity, Availability) [1]. Конфіденційність передбачає обмеження доступу до інформації лише для авторизованих користувачів, цілісність гарантує захист від несанкціонованої модифікації, а доступність забезпечує можливість використання інформації уповноваженими особами в потрібний час.

Конфіденційна інформація може існувати в різних формах:

— електронній

— фізичній (наприклад, паперові документи)

— нематеріальній (знання, ноу-хау).

У контексті кібербезпеки основна увага приділяється цифровим даним, які обробляються, зберігаються або передаються через інформаційні системи. До таких даних належать персональні дані, фінансова інформація, комерційна таємниця, інтелектуальна власність, державні секрети та інші категорії, що мають високу цінність для їх власників. Згідно з дослідженням Ponemon Institute, витік конфіденційної інформації у 2020 році спричинив середні збитки для організацій у розмірі 3,86 мільйона доларів США за один інцидент, що підкреслює економічну значущість захисту таких даних [2].

Класифікація конфіденційної інформації є необхідною умовою для ефективного управління її захистом. Вона дозволяє організаціям визначити пріоритетність захисту даних, розробити відповідні політики безпеки та

застосувати адекватні технологічні рішення. У науковій літературі виділяють кілька підходів до класифікації конфіденційної інформації, які базуються на різних критеріях: типі даних, рівні чутливості, сфері застосування та юридичних вимогах.

На основі вищезазначених критеріїв складено таблицю 1.1, яка ілюструє класифікацію конфіденційної інформації.

Таблиця 1.1– Класифікація конфіденційної інформації

№	Критерій	Категорії
1.	За типом даних	Персональні дані, комерційна таємниця, інтелектуальна власність, державна таємниця
2.	За рівнем чутливості	Обмежений доступ, конфіденційно, секретно, цілком таємно
3.	За сферою застосування	Корпоративна, державна, медична, фінансова, освітня

За типом даних конфіденційна інформація поділяється на кілька основних категорій. Персональні дані включають відомості, що дозволяють ідентифікувати особу, такі як ім’я, адреса, номер телефону, біометричні дані тощо. Відповідно до Загального регламенту захисту даних (GDPR), який діє в Європейському Союзі, захист персональних даних є обов’язковим, а їх витік може призвести до значних штрафів [3]. Комерційна таємниця охоплює інформацію, яка має економічну цінність для організації, наприклад, бізнес-плани, маркетингові стратегії, клієнтські бази. Інтелектуальна власність, така як патенти, авторські права чи технічна документація, також належить до конфіденційної інформації, оскільки її розголошення може підірвати конкурентоспроможність компанії. Державна таємниця включає дані, пов’язані з національною безпекою, обороною чи зовнішньою політикою, і регулюється спеціальними нормативними актами, наприклад, Законом України “Про державну таємницю” [4].

За рівнем чутливості конфіденційна інформація класифікується залежно від потенційних наслідків її розголошення. У багатьох організаціях застосовується багаторівнева система маркування, наприклад, “обмежений доступ”, “конфіденційно”, “секретно” чи “цілком таємно”. Такий підхід дозволяє диференціювати заходи захисту залежно від ступеня критичності інформації.

Наприклад, стандарт NIST SP 800-60 пропонує три рівні чутливості: низький, помірний і високий, кожен із яких відповідає певному рівню потенційного збитку в разі компрометації [5].

За сферою застосування конфіденційна інформація може бути поділена на корпоративну, державну, медичну, фінансову тощо. Кожна сфера має свої особливості та регуляторні вимоги. Наприклад, у медичній сфері захист даних регулюється стандартом HIPAA (Health Insurance Portability and Accountability Act), який встановлює суворі правила обробки інформації про здоров'я пацієнтів [6]. У фінансовому секторі застосовується стандарт PCI DSS, який регулює захист даних платіжних карток [7].

Класифікація конфіденційної інформації є динамічним процесом, який потребує регулярного перегляду з урахуванням змін у законодавстві, технологічному середовищі та організаційних потребах. Вона створює основу для розробки стратегій захисту, зокрема впровадження технологій DLP, які будуть розглянуті в підрозділі 1.3.

1.2 Загрози витоку інформації та методи їх ідентифікації

Загрози витоку конфіденційної інформації є однією з найсерйозніших проблем сучасної кібербезпеки, оскільки вони можуть призвести до фінансових втрат, репутаційних збитків і порушення законодавчих вимог. У науковій літературі загрози класифікуються за різними ознаками, зокрема за джерелом (внутрішні чи зовнішні), характером (навмисні чи ненавмисні) та каналом витоку (технічний, людський, організаський). Розуміння природи цих загроз і методів їх ідентифікації є критично важливим для розробки ефективних стратегій запобігання.

Внутрішні загрози становлять значну частку інцидентів витоку інформації. За даними звіту Verizon 2021 Data Breach Investigations Report, приблизно 34% усіх витоків даних були спричинені діями інсайдерів, включаючи як умисні зловживання, так і ненавмисні помилки [8]. Умисні внутрішні загрози пов'язані з

діями співробітників, підрядників або партнерів, які мають авторизований доступ до даних і використовують його для особистих цілей, наприклад, для продажу інформації конкурентам або з метою помсти. Ненавмисні загрози виникають через недбалість, некомпетентність або відсутність обізнаності працівників. Наприклад, надсилання конфіденційної інформації через незахищені канали зв’язку, втрата носіїв даних або неналежне налаштування хмарних сховищ є поширеними причинами витоків.

Зовнішні загрози включають кібератаки, спрямовані на отримання несанкціонованого доступу до конфіденційної інформації. До них належать фішинг, атаки з використанням шкідливого програмного забезпечення (malware), експлуатація вразливостей програмного забезпечення, атаки типу “людина посередині” (man-in-the-middle) та соціальна інженерія. Згідно з дослідженням IBM Security, у 2020 році фішинг був найпоширенішим вектором атак, відповідальним за 22% усіх інцидентів безпеки [9]. Такі атаки часто спрямовані на отримання облікових даних користувачів, що дозволяє зловмисникам проникнути в інформаційні системи та викрасти чутливі дані.

Канали витоку інформації є ще одним важливим аспектом аналізу загроз. Технічні канали включають передачу даних через електронну пошту, веб-трафік (HTTP, FTP), хмарні сервіси, USB-носії та інші засоби. Людський фактор відіграє ключову роль у витоках, спричинених необережним поводженням із даними, наприклад, розголошенням інформації під час телефонних розмов чи в соціальних мережах. Організаційні недоліки, такі як відсутність чітких політик безпеки або недостатній контроль доступу, також сприяють витокам.

Для ідентифікації загроз витоку інформації застосовуються різні методи, які можна поділити на проактивні та реактивні. Основні загрози і методи їх ідентифікації можна побачити на рисунку 1.1.

Проактивні методи спрямовані на попередження інцидентів шляхом аналізу потенційних вразливостей і ризиків. Одним із таких методів є оцінка ризиків (risk assessment), яка передбачає ідентифікацію активів, аналіз загроз і вразливостей, а також оцінку ймовірності та потенційного збитку.

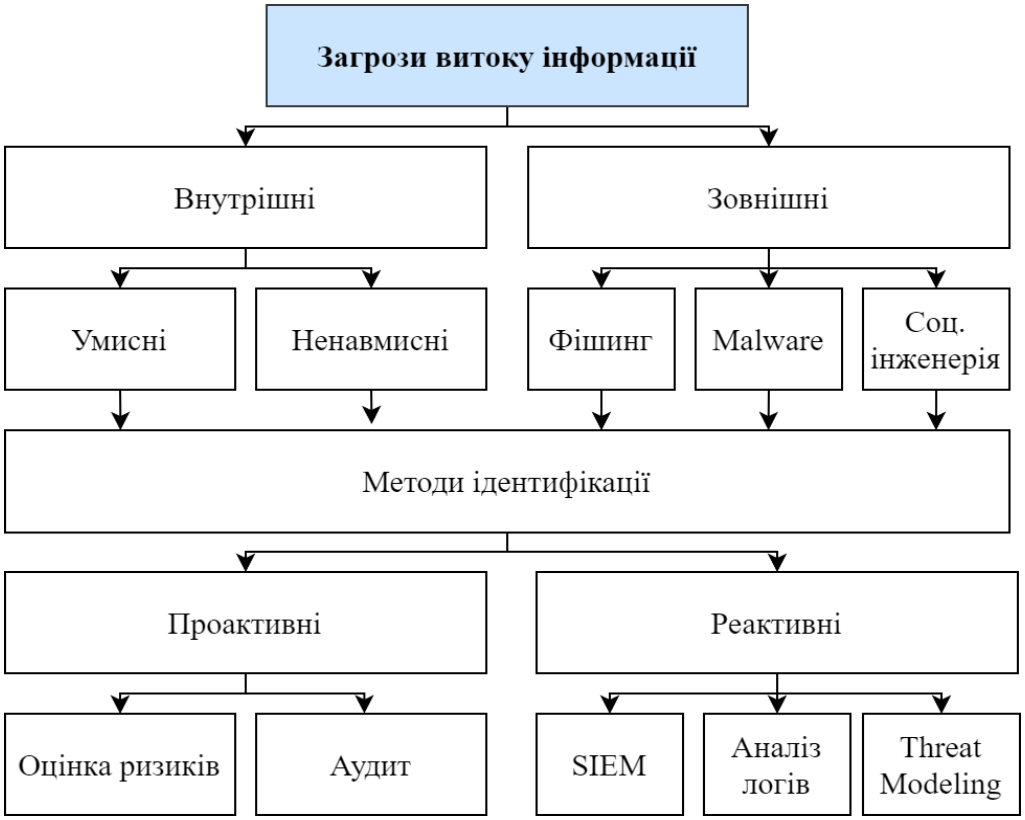

Рисунок 1.2 – Загрози витоку інформації та методи їх ідентифікації

Стандарт ISO/IEC 27005:2018 надає методологію для систематичного управління ризиками інформаційної безпеки [10]. Аудит безпеки, що включає перевірку конфігурацій систем, журналів подій і політик доступу, є ще одним інструментом проактивного виявлення загроз.

Реактивні методи застосовуються після виявлення інциденту або підозри на витік. До них належить аналіз журналів подій (log analysis), який дозволяє відстежити дії користувачів і виявити аномалії, наприклад, несанкціонований доступ до даних. Технології SIEM (Security Information and Event Management) забезпечують централізований збір і аналіз даних безпеки, що значно підвищує ефективність виявлення інцидентів [11]. Моделювання загроз (threat modeling) є ще одним методом, який дозволяє ідентифікувати потенційні вектори атак і оцінити їх вплив на інформаційні активи.

Ефективна ідентифікація загроз вимагає комплексного підходу, який поєднує організаційні, технічні та людські аспекти. Впровадження політик безпеки,

навчання персоналу та використання сучасних технологій, таких як DLP-системи, дозволяють значно знизити ймовірність витоку інформації.

1.3 Технології DLP: принципи роботи та основні компоненти

Технології запобігання витоку даних (Data Leak Prevention, DLP) є спеціалізованими рішеннями, спрямованими на захист конфіденційної інформації від несанкціонованого розголошення. Вони набули популярності після 2000 року, коли внутрішні загрози стали розглядатися як одна з ключових проблем інформаційної безпеки. Термін DLP був запропонований агентством Forrester у 2005 році, і відтоді ці технології стали невід'ємною частиною стратегій кібербезпеки організацій [12]. DLP-системи призначені для моніторингу, виявлення та блокування спроб передачі конфіденційної інформації за межі інформаційної системи, а також для запобігання її неналежному використанню всередині організації.

Принципи роботи DLP-систем базуються на аналізі потоків даних, які перетинають периметр інформаційної системи або використовуються в межах організації. Основна мета – виявлення конфіденційної інформації в цих потоках і застосування відповідних заходів для запобігання її витоку. DLP-системи використовують комбінацію контентного та контекстного аналізу. Контентний аналіз передбачає перевірку вмісту даних на наявність чутливої інформації, наприклад, номерів кредитних карток, особистих ідентифікаторів чи ключових слів, пов'язаних із комерційною таємницею. Контекстний аналіз враховує метадані, такі як джерело даних, одержувач, протокол передачі чи час операції, що дозволяє виявляти підозрілі дії навіть у разі використання шифрування [13].

DLP-системи діють у трьох основних режимах: моніторинг, попередження та блокування. У режимі моніторингу система фіксує всі дії з даними, створюючи журнали для подальшого аналізу. Попередження передбачає повідомлення адміністратора або користувача про потенційне порушення політики безпеки. Блокування є найактивнішим режимом, коли система автоматично зупиняє

передачу даних, якщо виявлено порушення. Наприклад, якщо співробітник намагається надіслати файл із конфіденційною інформацією через електронну пошту, DLP-система може заблокувати повідомлення та повідомити про інцидент.

Основні компоненти DLP-систем включають кілька ключових елементів, які забезпечують їх функціональність. Першим є модуль аналізу вмісту (content analysis engine), який використовує технології розпізнавання шаблонів, регулярних виразів, машинного навчання та відбитків даних (data fingerprinting) для ідентифікації чутливої інформації. Наприклад, відбитки даних дозволяють створювати унікальні сигнатури документів, що забезпечує їх точне виявлення навіть у зміненому вигляді [14]. Другий компонент – модуль управління політиками (policy management), який дозволяє адміністраторам визначати правила безпеки, наприклад, заборонити копіювання певних файлів на USB-носії або обмежити доступ до хмарних сховищ. Третій компонент – модуль моніторингу та реагування (monitoring and response), який відповідає за збір даних, аналіз подій і виконання дій, таких як блокування чи повідомлення.

DLP-системи можуть бути розгорнуті в різних конфігураціях: мережеві, кінцеві точки (endpoint) та хмарні. Мережеві DLP-системи контролюють трафік, що проходить через корпоративну мережу, наприклад, електронну пошту, веб-запити чи FTP-передачі. Системи для кінцевих точок встановлюються на пристроях користувачів і контролюють дії, такі як копіювання файлів чи друк документів. Хмарні DLP-системи інтегруються з хмарними сервісами, такими як Google Drive чи Microsoft OneDrive, для захисту даних у хмарному середовищі [15].

Для ілюстрації компонентів DLP-системи можна використати таблицю 1.2.

Незважаючи на високу ефективність, DLP-системи мають певні обмеження. По-перше, вони потребують ретельного налаштування, щоб уникнути помилкових спрацьовувань (false positives), які можуть ускладнити роботу користувачів. По-друге, DLP-системи менш ефективні проти умисних інсайдерів, які використовують складні методи обхідного захисту, наприклад, шифрування даних перед передачею. По-третє, успішне впровадження DLP вимагає інтеграції з

іншими системами безпеки, такими як SIEM, антивірусне програмне забезпечення та системи управління доступом [16].

Таблиця 1.2 – Основні компоненти DLP-систем

№	Компонент	Функція
1.	Аналіз вмісту	Виявлення чутливої інформації за допомогою шаблонів, відбитків, ML
2.	Управління політиками	Визначення правил безпеки та їх застосування
3.	Моніторинг і реагування	Збір даних, аналіз подій, блокування чи повідомлення про порушення

У майбутньому розвиток DLP-систем буде пов’язаний із застосуванням штучного інтелекту та машинного навчання для підвищення точності аналізу даних і адаптації до нових загроз. Крім того, зростання популярності хмарних технологій і віддаленої роботи створює нові виклики для DLP, які потребують розробки гнучких і масштабованих рішень.

Теоретичні основи запобігання витоку конфіденційної інформації формують фундамент для розробки ефективних стратегій кібербезпеки. Конфіденційна інформація є цінним активом, який потребує захисту від широкого спектра загроз, включаючи внутрішні та зовнішні атаки. Класифікація інформації за типом, рівнем чутливості та сферою застосування дозволяє організаціям визначити пріоритети захисту та розробити відповідні політики. Ідентифікація загроз витоку інформації вимагає комплексного підходу, який поєднує проактивні та реактивні методи, такі як оцінка ризиків, аудит безпеки та аналіз подій. Технології DLP відіграють ключову роль у запобіганні витокам, забезпечуючи моніторинг, аналіз і блокування чутливих даних. Їх ефективність залежить від правильного налаштування, інтеграції з іншими системами безпеки та адаптації до нових викликів цифрового середовища. Подальші дослідження в цій галузі мають бути спрямовані на вдосконалення методів виявлення загроз і розробку інноваційних DLP-рішень, здатних протистояти складним і цілеспрямованим атакам.

Теоретичні основи запобігання витоку конфіденційної інформації створюють фундамент для розробки та впровадження ефективних стратегій кібербезпеки,

спрямованих на захист цінних даних від несанкціонованого доступу, розголошення чи втрати. Конфіденційна інформація, що характеризується конфіденційністю, цілісністю та доступністю, є ключовим активом для організацій, державних установ і приватних осіб, а її захист має економічне, правове та репутаційне значення. Класифікація інформації за типом даних, рівнем чутливості та сферою застосування забезпечує структурований підхід до визначення пріоритетів захисту, дозволяючи організаціям розробляти цільові політики безпеки та застосовувати відповідні технологічні рішення. Аналіз загроз витоку інформації підкреслює їх різноманітність, включаючи внутрішні та зовнішні джерела, а також технічні, людські й організаційні канали, що вимагає комплексного поєднання проактивних і реактивних методів ідентифікації, таких як оцінка ризиків, аудит безпеки, аналіз журналів подій і моделювання загроз. Технології DLP відіграють центральну роль у запобіганні витокам, забезпечуючи моніторинг, виявлення та блокування чутливих даних через контентний і контекстний аналіз, однак їх ефективність залежить від ретельного налаштування, інтеграції з іншими системами безпеки та здатності адаптуватися до нових викликів, таких як зростання хмарних технологій і віддаленої роботи. Подальші дослідження в цій галузі мають бути спрямовані на вдосконалення алгоритмів штучного інтелекту та машинного навчання для підвищення точності виявлення загроз, а також на розробку гнучких і масштабованих DLP-рішень, здатних протистояти складним кібератакам і забезпечувати захист інформації в умовах динамічного цифрового середовища.

2 АНАЛІЗ ТА ПРОЕКТУВАННЯ DLP-СИСТЕМ ДЛЯ ЗАХИСТУ ДАНИХ

2.1 Методи виявлення та блокування витоків конфіденційної інформації

Запобігання витоку конфіденційної інформації є однією з найактуальніших проблем сучасної кібербезпеки, адже втрата чутливих даних може спричинити значні економічні, репутаційні та правові наслідки для організацій, державних установ і приватних осіб. У контексті швидкої цифровізації, зростання обсягів даних і ускладнення кіберзагроз методи виявлення та блокування витоків інформації набувають стратегічного значення.

У науковій літературі витік інформації визначається як несанкціоноване розголошення, передача або втрата даних, які мають цінність для їх власника і підлягають захисту від доступу неавторизованих осіб [16]. Такі дані можуть включати персональні відомості, комерційні таємниці, інтелектуальну власність або державні секрети. За даними звіту IBM, у 2022 році середні збитки від одного інциденту витоку даних сягнули 4,45 мільйона доларів США, що підкреслює економічну значущість своєчасного виявлення та запобігання таким інцидентам [17].

Теоретичні основи методів виявлення витоків базуються на трьох ключових аспектах: ідентифікації чутливої інформації, моніторингу потоків даних і аналізу аномалій. Ідентифікація чутливої інформації передбачає визначення даних, які потребують захисту, на основі їх типу, рівня чутливості та контексту використання. Наприклад, стандарт ISO/IEC 27001:2013 наголошує на необхідності класифікації інформації для забезпечення її конфіденційності, цілісності та доступності [1]. Моніторинг потоків даних дозволяє відстежувати рух інформації в межах інформаційної системи та за її межами, включаючи електронну пошту, веб-трафік, хмарні сервіси та фізичні носії. Аналіз аномалій спрямований на виявлення відхилень від нормальної поведінки користувачів або систем, які можуть свідчити про потенційний витік.

Методи виявлення витоків можна поділити на дві основні категорії: контентно-орієнтовані та поведінково-орієнтовані. Контентно-орієнтовані методи

зосереджені на аналізі вмісту даних для ідентифікації чутливої інформації. Вони використовують такі технології, як розпізнавання шаблонів (pattern matching), регулярні вирази, відбитки даних (data fingerprinting) і ключові слова. Наприклад, регулярні вирази дозволяють виявляти формати даних, такі як номери кредитних карток або ідентифікаційні номери, тоді як відбитки даних створюють унікальні сигнатури документів для їх точного розпізнавання [14].

Поведінково-орієнтовані методи аналізують дії користувачів або систем, щоб виявити аномалії, наприклад, нетипову активність, таку як масове копіювання файлів або надсилання даних на зовнішні ресурси. Ці методи часто базуються на технологіях машинного навчання, які дозволяють адаптуватися до нових моделей поведінки [18].

Теоретичною основою для обох категорій є модель управління ризиками, описана в стандарті ISO/IEC 27005:2018. Ця модель передбачає ідентифікацію активів, аналіз загроз і вразливостей, а також оцінку ймовірності та потенційного збитку від витоку [10]. Наприклад, ризик витоку може бути пов’язаний із внутрішніми загрозами (діями інсайдерів) або зовнішніми атаками (фішинг, шкідливе програмне забезпечення). Згідно з дослідженням Verizon, у 2021 році 34% витоків даних були спричинені діями інсайдерів, що підкреслює важливість моніторингу внутрішньої активності [8].

Для ефективного виявлення витоків необхідно враховувати канали передачі даних, які можуть бути технічними (електронна пошта, USB-носії, хмарні сервіси), людськими (усне розголошення, соціальна інженерія) або організаційними (неналежне налаштування систем). Кожен канал має свої особливості, які впливають на вибір методів виявлення. Наприклад, технічні канали потребують аналізу мережевого трафіку, тоді як людські канали вимагають навчання персоналу та впровадження політик безпеки.

Для візуалізації теоретичних основ методів виявлення витоків можна використати наступний рисунок 2.1.

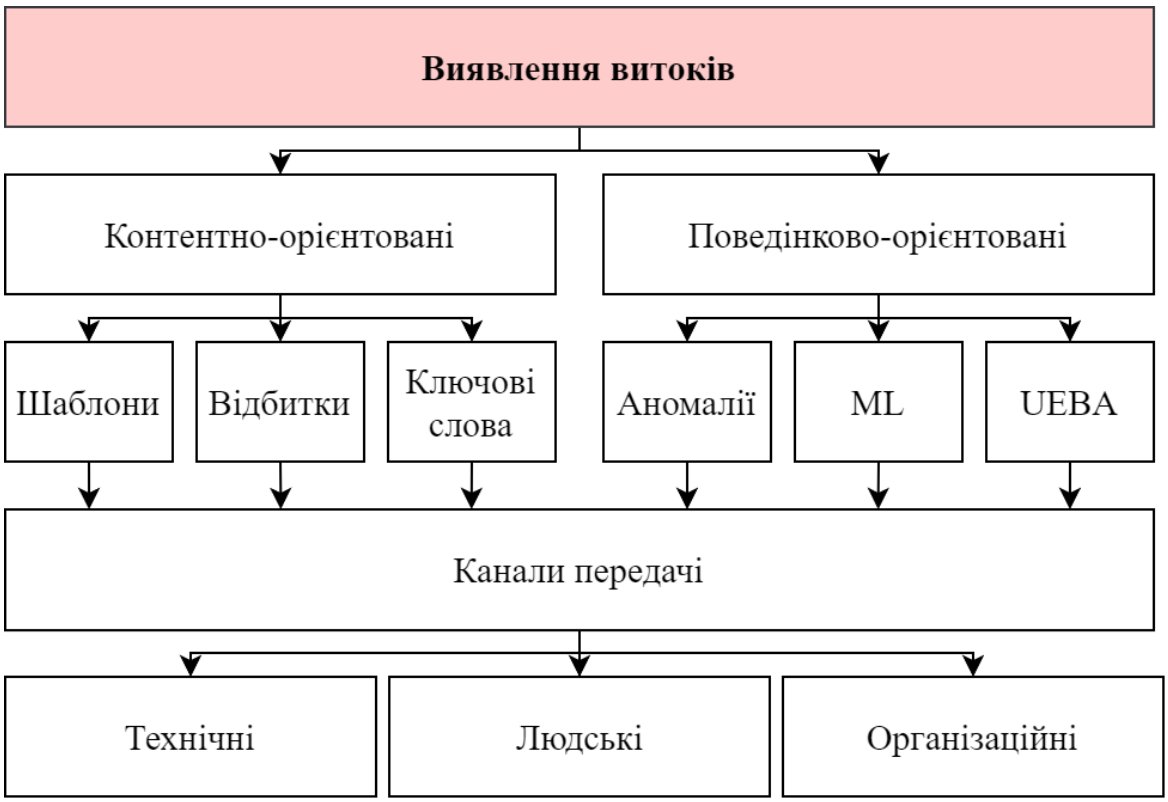

Рисунок 2.1 – Теоретичні основи виявлення витоків інформації

Рисунок ілюструє взаємозв’язок між методами виявлення та каналами витоку, підкреслюючи комплексний характер підходу до захисту інформації.

Технології аналізу даних відіграють ключову роль у виявленні загроз витоку конфіденційної інформації, оскільки сучасні інформаційні системи генерують величезні обсяги даних, які потребують швидкої та точної обробки. Ці технології включають методи обробки великих даних (Big Data), машинне навчання, аналіз поведінки користувачів і систем (UEBA), а також системи управління інформацією та подіями безпеки (SIEM). Їх застосування дозволяє не лише виявляти потенційні витоки, а й прогнозувати загрози, що значно підвищує ефективність захисту.

Однією з основних технологій є аналіз великих даних, який дозволяє обробляти структуровані та неструктуровані дані в реальному часі. За даними дослідження Gartner, у 2022 році 60% організацій використовували Big Data для аналізу безпеки, що дало змогу скоротити час виявлення інцидентів на 30% [19]. Аналіз великих даних передбачає використання таких інструментів, як Hadoop, Spark або Elasticsearch, для обробки логів подій, мережевого трафіку та інших

джерел інформації. Це дозволяє виявляти патерни, які свідчать про витік, наприклад, масову передачу файлів на зовнішні сервери.

Машинне навчання є ще одним важливим інструментом для ідентифікації загроз. Алгоритми ML, такі як класифікація, кластеризація та виявлення аномалій, дозволяють аналізувати великі набори даних і виявляти відхилення від нормальної поведінки. Наприклад, алгоритми кластеризації можуть групувати дії користувачів за схожими характеристиками, тоді як методи виявлення аномалій ідентифікують нетипові операції, такі як надсилання конфіденційних документів у неробочий час [20]. Дослідження, проведене MIT, показало, що використання ML у системах безпеки підвищує точність виявлення загроз на 15–20% порівняно з традиційними методами [21].

Аналіз поведінки користувачів і систем (User and Entity Behavior Analytics, UEBA) є спеціалізованою технологією, яка фокусується на моніторингу дій користувачів, пристроїв і додатків. UEBA використовує базові профілі поведінки, створені на основі історичних даних, для виявлення аномалій. Наприклад, якщо співробітник, який зазвичай працює з фінансовими звітами, раптово починає копіювати великі обсяги даних на USB-носій, UEBA може позначити цю дію як підозрілу [22]. Впровадження UEBA дозволяє організаціям скоротити кількість інцидентів, пов’язаних із інсайдерами, на 25% [23].

Системи SIEM є інтеграційним рішенням, яке поєднує збір, аналіз і кореляцію даних із різних джерел, таких як журнали подій, мережевий трафік і антивірусні системи. Вони дозволяють централізовано моніторити інформаційні системи та виявляти інциденти в реальному часі. Наприклад, SIEM-система може виявити спробу несанкціонованого доступу до бази даних, аналізуючи логи автентифікації та мережеві запити [11]. Згідно з дослідженням IBM, використання SIEM скорочує середній час реагування на інциденти з 74 до 24 годин [24].

Для порівняння основних технологій аналізу даних складено таблицю 2.1.

Ці технології не є взаємовиключними, а навпаки, доповнюють одна одну. Наприклад, інтеграція UEBA з SIEM дозволяє поєднати аналіз поведінки з кореляцією подій, що підвищує точність виявлення загроз. Однак їх ефективність

залежить від якості даних, правильного налаштування та регулярного оновлення моделей.

Таблиця 2.1 – Технології аналізу даних для виявлення загроз

№	Технологія	Основні функції	Переваги	Обмеження
1.	Big Data	Обробка великих обсягів даних у реальному часі	Швидкість, масштабованість	Висока складність впровадження
2.	Машинне навчання	Виявлення аномалій, прогнозування загроз	Висока точність, адаптивність	Потреба у великих наборах даних
3.	UEBA	Аналіз поведінки користувачів і систем	Ефективність проти інсайдерів	Висока вартість, помилкові спрацьовування
4.	SIEM	Кореляція подій, централізований моніторинг	Інтеграція з іншими системами	Складність налаштування

Блокування витоків конфіденційної інформації в реальному часі є завершальним етапом захисту, який спрямований на запобігання несанкціонованій передачі даних. Цей процес вимагає швидкого реагування на виявлені загрози, що можливо завдяки спеціалізованим технологіям, таким як системи запобігання витоку даних (DLP), системи управління доступом і шифрування. Ефективність блокування залежить від інтеграції цих технологій із системами виявлення, а також від чітко визначених політик безпеки.

Системи DLP є основним інструментом для блокування витоків. Вони діють у трьох основних режимах: моніторинг, попередження та блокування. У режимі моніторингу DLP фіксує всі дії з даними, створюючи журнали для аналізу. Попередження передбачає повідомлення адміністратора або користувача про потенційне порушення. Блокування є найактивнішим режимом, коли система автоматично зупиняє передачу даних, якщо виявлено порушення. Наприклад, якщо співробітник намагається надіслати файл із конфіденційною інформацією через незахищену електронну пошту, DLP може заблокувати повідомлення та повідомити про інцидент [15]. Згідно з дослідженням Gartner, впровадження DLP скорочує кількість інцидентів витоку даних на 40% у середньому [25].

DLP-системи використовують комбінацію контентного та контекстного аналізу. Контентний аналіз перевіряє вміст даних на наявність чутливої інформації, наприклад, за допомогою регулярних виразів або відбитків даних. Контекстний аналіз враховує метадані, такі як джерело, одержувач, час операції чи протокол передачі. Наприклад, DLP може дозволити передачу документа всередині корпоративної мережі, але заблокувати його надсилання на зовнішній домен [13]. Дослідження, проведене Alneyadi et al., показало, що комбінація контентного та контекстного аналізу підвищує точність блокування на 25% порівняно з використанням лише одного методу [26].

Системи управління доступом (Identity and Access Management, IAM) відіграють важливу роль у блокуванні витоків, обмежуючи доступ до конфіденційної інформації лише для авторизованих користувачів. IAM включає такі механізми, як двофакторна автентифікація, рольовий доступ (RBAC) і принцип найменших привілеїв. Наприклад, застосування RBAC дозволяє обмежити доступ до бази клієнтів лише для маркетингового відділу, що знижує ризик витоку [27]. Згідно з дослідженням Forrester, організації, які використовують IAM, скорочують кількість інцидентів, пов’язаних із несанкціонованим доступом, на 30% [28].

Шифрування є ще одним ефективним методом блокування витоків. Воно забезпечує захист даних під час їх передачі та зберігання, роблячи їх недоступними для неавторизованих осіб. Наприклад, використання протоколу TLS для захисту електронної пошти або алгоритмів AES для шифрування файлів на USB-носіях значно знижує ризик перехоплення даних [29]. Дослідження NIST показало, що шифрування скорочує ймовірність успішного витоку даних на 50% у разі компрометації системи [22]. На рисунку 2.2 зображено основні підходи до блокування витоків.

Незважаючи на високу ефективність, методи блокування мають певні обмеження. По-перше, DLP-системи можуть генерувати помилкові спрацьовування, що ускладнює роботу користувачів. По-друге, умисні інсайдери можуть використовувати складні методи обхідного захисту, наприклад, шифрування даних перед передачею.

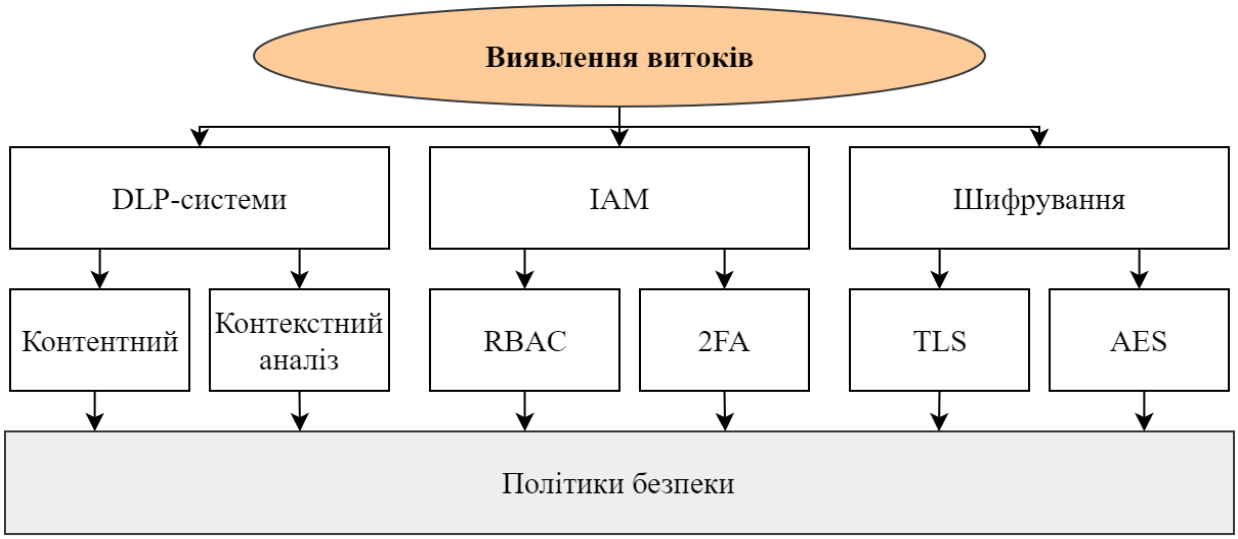

Рисунок 2.2 – Підходи до блокування витоків інформації

По-третє, впровадження IAM і шифрування вимагає значних ресурсів і ретельного налаштування. Для подолання цих обмежень необхідно поєднувати технологічні рішення з організаційними заходами, такими як навчання персоналу та регулярний аудит безпеки.

Методи виявлення та блокування витоків конфіденційної інформації формують основу сучасних стратегій кібербезпеки. Теоретичні основи виявлення базуються на ідентифікації чутливих даних, моніторингу потоків інформації та аналізі аномалій, що дозволяє своєчасно виявляти потенційні загрози. Технології аналізу даних, такі як Big Data, машинне навчання, UEBA та SIEM, забезпечують обробку великих обсягів інформації та підвищують точність ідентифікації загроз. Підходи до блокування, включаючи DLP-системи, управління доступом і шифрування, дозволяють запобігати несанкціонованій передачі даних у реальному часі. Інтеграція цих методів із чіткими політиками безпеки та навчання персоналу є ключовою умовою для ефективного захисту інформації. Подальші дослідження в цій галузі мають бути спрямовані на вдосконалення алгоритмів машинного навчання, підвищення адаптивності DLP-систем і розробку нових методів протидії складним загрозам, таким як квантові атаки чи використання штучного інтелекту зловмисниками.

2.2 Порівняльний аналіз сучасних DLP-рішень на ринку

Системи запобігання витоку даних (Data Loss Prevention, DLP) є ключовим інструментом у сфері кібербезпеки, спрямованим на захист конфіденційної інформації від несанкціонованого розголошення, втрати чи крадіжки. У сучасному цифровому середовищі, де обсяги даних зростають експоненціально, а кіберзагрози стають дедалі складнішими, DLP-рішення відіграють вирішальну роль у забезпеченні безпеки організацій. Ринок DLP-систем є динамічним і різноманітним, пропонуючи широкий спектр продуктів від провідних постачальників, таких як Broadcom, Microsoft, Forcepoint, Trellix і Palo Alto Networks. Кожен із цих продуктів має унікальні характеристики, функціональні можливості та підходи до захисту даних, що робить порівняльний аналіз їхньої ефективності важливим завданням для вибору оптимального рішення.

У науковій літературі DLP-системи класифікуються за типами розгортання (мережеві, кінцеві точки, хмарні, гібридні) та підходами до аналізу даних (контентний, контекстний, поведінковий) [13]. Сучасні DLP-рішення поєднують кілька архітектурних підходів, щоб забезпечити комплексний захист даних у різноманітних середовищах, включаючи локальні системи, хмарні платформи та віддалені робочі місця.

Мережеві DLP-системи зосереджені на моніторингу та захисті даних, що передаються через корпоративні мережі, такі як електронна пошта, веб-трафік (HTTP/HTTPS), FTP або хмарні сервіси. Вони зазвичай розгортаються на мережевих шлюзах або проксі-серверах і використовують технології глибокого аналізу пакетів (Deep Packet Inspection, DPI) для перевірки вмісту даних у реальному часі [10]. Наприклад, рішення від Cisco Systems інтегруються з мережевою інфраструктурою, забезпечуючи аналіз трафіку та блокування витоків на рівні мережевих протоколів [31]. Такі системи є ефективними для захисту даних у русі (data in motion), але їхня продуктивність може знижуватися в умовах шифрованого трафіку, що становить виклик для сучасних організацій, де значна частина комунікацій використовує TLS/SSL.

DLP-системи для кінцевих точок (endpoint DLP) встановлюються на пристроях користувачів, таких як комп’ютери, ноутбуки чи мобільні пристрої, і контролюють дії, пов’язані з даними на цих пристроях. Вони здатні відстежувати копіювання файлів на USB-носії, друк документів, надсилання даних через месенджери чи хмарні сховища. CoSoSys Endpoint Protector є прикладом рішення, яке спеціалізується на захисті кінцевих точок, пропонуючи автоматизоване шифрування та детальний контроль пристроїв [32]. Такі системи є особливо цінними в умовах зростання популярності віддаленої роботи та політики “принось свій пристрій” (BYOD), але їх ефективність залежить від правильного налаштування агентів на кожному пристрої.

Хмарні DLP-системи (cloud DLP) розроблені для захисту даних у хмарних середовищах, таких як Microsoft 365, Google Workspace або Amazon Web Services (AWS). Вони інтегруються з хмарними платформами через API, забезпечуючи моніторинг і захист даних у реальному часі. Microsoft Purview DLP, наприклад, пропонує глибоку інтеграцію з екосистемою Microsoft 365, дозволяючи виявляти та блокувати витоки в таких додатках, як Teams, OneDrive і Outlook [33]. Хмарні рішення є гнучкими та масштабованими, але їхня ефективність може бути обмеженою в гібридних середовищах, де частина даних залишається на локальних серверах.

Гібридні DLP-системи поєднують елементи мережевих, кінцевих і хмарних рішень, забезпечуючи комплексний захист у складних IT-екосистемах. Forcepoint DLP є прикладом гібридного рішення, яке використовує єдину консоль управління для моніторингу даних у різних середовищах, включаючи локальні мережі, хмарні платформи та кінцеві точки [34]. Гібридний підхід є оптимальним для великих організацій із розподіленими IT-інфраструктурами, але його впровадження може бути складним через потребу в інтеграції з різними системами.

Для аналізу даних DLP-системи використовують контентний, контекстний і поведінковий підходи. Контентний аналіз передбачає перевірку вмісту даних на наявність чутливої інформації за допомогою регулярних виразів, відбитків даних (data fingerprinting) або ключових слів. Наприклад, Broadcom Symantec DLP

використовує технологію точного зіставлення даних (Exact Data Matching, EDM) для ідентифікації структурованих даних, таких як номери кредитних карток [35]. Контекстний аналіз враховує метадані, такі як джерело, одержувач, протокол передачі чи час операції, що дозволяє виявляти підозрілі дії навіть у разі використання шифрування. Поведінковий аналіз, який часто базується на технологіях машинного навчання (ML) і аналізу поведінки користувачів (UEBA), виявляє аномалії, наприклад, нетипову активність користувача, таку як масове копіювання файлів [22]. Для візуалізації архітектурних підходів можна використати наступну схему з рисунку 2.3.

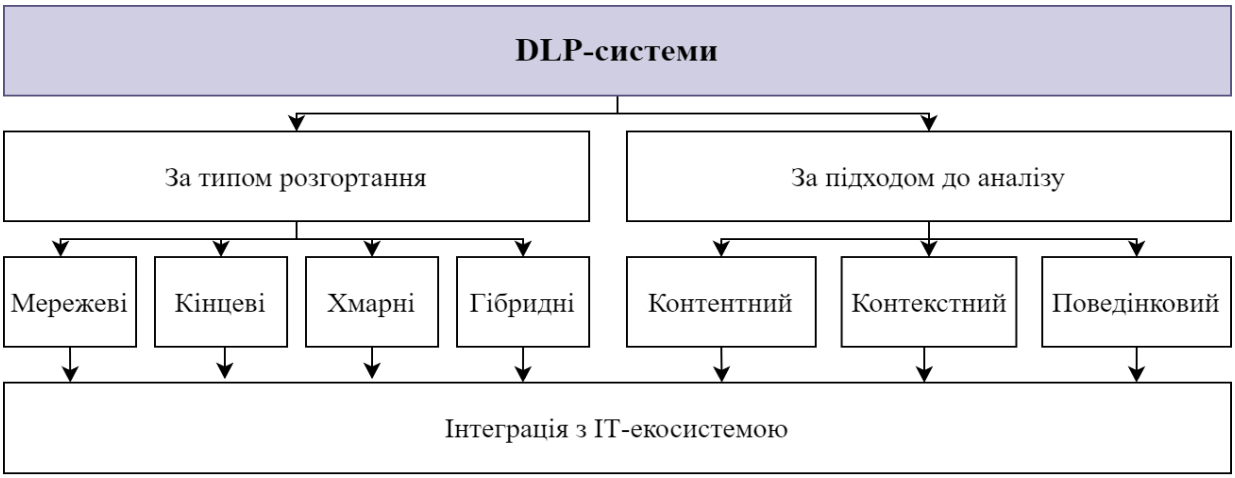

Рисунок 2.3 – Архітектурні підходи до DLP-систем

Схема ілюструє різноманітність архітектурних підходів і їхню залежність від інтеграції з IT-інфраструктурою організації.

Порівняння DLP-рішень вимагає аналізу їхніх ключових функцій, таких як можливості виявлення, моніторингу, блокування, інтеграції, управління політиками та відповідність регуляторним вимогам. Провідні постачальники DLP-систем, такі як Broadcom (Symantec), Microsoft, Forcepoint, Trellix і Palo Alto Networks, пропонують різноманітні функціональні можливості, які відповідають потребам різних організацій. Згідно з дослідженням Gartner, у 2023 році ринок DLP оцінювався в 3,4 мільярда доларів США і прогнозується до зростання до 8,9

мільярда до 2028 року з CAGR 21,2%, що свідчить про високу конкуренцію та інноваційну активність у цій сфері [36].

Broadcom Symantec DLP є одним із лідерів ринку завдяки своїй універсальності та підтримці всіх типів розгортання (мережевих, кінцевих, хмарних). Воно використовує технології EDM і Indexed Document Matching (IDM) для точного виявлення структурованих і неструктурованих даних, а також інтегрується з хмарними платформами, такими як AWS і Azure [37]. Система підтримує централізоване управління через єдину консоль, що полегшує адміністрування в великих організаціях. Однак її впровадження може бути складним і дорогим, що робить її менш доступною для малих і середніх підприємств (СМП).

Microsoft Purview DLP виділяється своєю глибокою інтеграцією з Signed Data Protection (EDP) для хмарних додатків, таких як Microsoft 365. Воно використовує попередньо визначені шаблони для виявлення чутливих даних, таких як номери кредитних карток чи медичні записи, і підтримує автоматичне блокування дій, які порушують політики безпеки [38]. Purview DLP є економічно вигідним рішенням для організацій, які вже використовують екосистему Microsoft, але його можливості обмежені в гібридних або локальних середовищах, де потрібна підтримка не-Microsoft платформ.

Forcepoint DLP пропонує гнучкий підхід, який поєднує контентний і поведінковий аналіз, а також інтеграцію з платформою Forcepoint ONE Security Service Edge (SSE). Воно підтримує захист даних у хмарі, на кінцевих точках і в мережах, використовуючи бібліотеку попередньо визначених політик для швидкого розгортання [39]. Forcepoint є ефективним для організацій із гібридними середовищами, але його висока вартість і складність налаштування можуть бути бар’єрами для СМП.

Trellix DLP, створений на основі продуктів McAfee і FireEye, пропонує чотири основні компоненти: DLP Endpoint, DLP Monitor, DLP Discover і DLP Prevent. Воно спеціалізується на захисті кінцевих точок і хмарних сховищ, таких як Dropbox, і використовує технології контентного аналізу та відбитків даних [40].

Trellix є популярним у фінансовому секторі завдяки своїй зрілій архітектурі, але його можливості в хмарних середовищах менш розвинені порівняно з Microsoft чи Forcepoint.

Palo Alto Networks Enterprise DLP є хмарно-орієнтованим рішенням, яке інтегрується з їхньою платформою Prisma Cloud. Воно підтримує автоматичне виявлення даних у хмарних середовищах і використовує машинне навчання для аналізу поведінки користувачів [41]. Це рішення є оптимальним для організацій із хмарно-центричною інфраструктурою, але його функціональність для локальних систем обмежена.

Для порівняння ключових функцій провідних DLP-рішень складено таблицю 2.2.

Таблиця 2.2 – Порівняння ключових функцій DLP-рішень

Рішення	Тип розгортання	Аналіз даних	Інтеграція з хмарою	Управління політиками
Broadcom Symantec	Мережеве, кінцеве, хмарне, гібридне	Контентний, контекстний	AWS, Azure, Google Cloud	Централізоване
Microsoft Purview	Хмарне, частково гібридне	Контентний, шаблони	Microsoft 365	Інтегроване з Microsoft
Forcepoint	Мережеве, кінцеве, хмарне, гібридне	Контентний, контекстний, поведінковий	AWS, Azure, Google Workspace	Централізоване
Trellix	Кінцеве, хмарне	Контентний, відбитки	Dropbox, Box	Модульне
Palo Alto Networks	Хмарне	Контентний, поведінковий	Prisma Cloud, AWS, Azure	Централізоване
Рішення	Тип розгортання	Аналіз даних	Інтеграція з хмарою	Управління політиками

Ефективність DLP-рішень залежить від їхньої здатності виявляти та блокувати витоки даних у різних сценаріях, таких як захист даних у хмарі, запобігання інсайдерським загрозам, відповідність регуляторним вимогам і протидія зовнішнім атакам. Кожен сценарій має свої особливості, які впливають на вибір оптимального DLP-рішення. Згідно з дослідженням MarketsandMarkets, ключовими факторами, що впливають на ефективність DLP, є точність виявлення, швидкість реагування, масштабованість і простота управління [42].

У сценарії захисту даних у хмарі хмарно-орієнтовані рішення, такі як Microsoft Purview і Palo Alto Networks Enterprise DLP, демонструють високу ефективність завдяки інтеграції з популярними хмарними платформами. Наприклад, Microsoft Purview може автоматично сканувати файли в OneDrive і блокувати їхній доступ, якщо вони містять чутливі дані, такі як номери соціального страхування [43]. Palo Alto Networks використовує ML для аналізу великих обсягів хмарних даних, що дозволяє виявляти приховані витоки, наприклад, у неправильно налаштованих S3-бакетах AWS. Однак ці рішення менш ефективні в локальних середовищах, де Broadcom Symantec і Forcepoint мають перевагу завдяки підтримці гібридних архітектур.

Запобігання інсайдерським загрозам є одним із найскладніших завдань для DLP-систем, оскільки інсайдери часто мають авторизований доступ до даних. Forcepoint і Trellix демонструють високу ефективність у цьому сценарії завдяки технологіям UEBA, які дозволяють виявляти аномальну поведінку, наприклад, масове копіювання файлів або надсилання даних на особисту електронну пошту [44]. Дослідження Forrester показало, що використання UEBA в DLP-системах скорочує кількість інцидентів, пов’язаних із інсайдерами, на 25% [45]. Broadcom Symantec також ефективний у цьому сценарії завдяки технології EDM, яка забезпечує точне виявлення структурованих даних.

Відповідність регуляторним вимогам, таким як GDPR, HIPAA або CCPA, є критично важливою для організацій у регульованих галузях, таких як фінанси та охорона здоров’я. Microsoft Purview і Trellix пропонують попередньо налаштовані шаблони для відповідності цим стандартам, що полегшує їх впровадження.

Наприклад, Trellix DLP Monitor може генерувати звіти про відповідність HIPAA, які демонструють захист даних про здоров’я пацієнтів. Broadcom Symantec також підтримує широкий спектр регуляторних стандартів, але його складність налаштування може уповільнити процес відповідності.

Протидія зовнішнім атакам, таким як фішинг або шкідливе програмне забезпечення, є менш типовим сценарієм для DLP, але деякі рішення інтегруються з іншими системами безпеки для підвищення ефективності. Наприклад, Palo Alto Networks Enterprise DLP інтегрується з Prisma Access, що дозволяє виявляти та блокувати витоки, спричинені шкідливим ПЗ [46]. Forcepoint також підтримує інтеграцію з SIEM-системами, що забезпечує швидке реагування на зовнішні загрози. Для оцінки ефективності DLP-рішень у різних сценаріях можна використати наступну таблицю 2.3.

Таблиця 2.3 – Ефективність DLP-рішень у різних сценаріях

Сценарій використання	Broadcom Symantec	Microsoft Purview	Forcepoint	Trellix	Palo Alto Networks
Захист даних у хмарі	Висока	Висока	Висока	Середня	Висока
Запобігання інсайдерським загрозам	Висока	Середня	Висока	Висока	Середня
Відповідність регуляторним вимогам	Висока	Висока	Середня	Висока	Середня
Протидія зовнішнім атакам	Середня	Низька	Середня	Низька	Висока

Порівняльний аналіз сучасних DLP-рішень на ринку демонструє їхню різноманітність за архітектурними підходами, функціональними можливостями та ефективністю в різних сценаріях використання. Архітектура DLP-систем включає мережеві, кінцеві, хмарні та гібридні підходи, кожен із яких має свої переваги та обмеження. Провідні рішення, такі як Broadcom Symantec, Microsoft Purview, Forcepoint, Trellix і Palo Alto Networks, пропонують широкий спектр функцій, включаючи контентний, контекстний і поведінковий аналіз, інтеграцію з хмарними платформами та підтримку регуляторних вимог. Оцінка ефективності показує, що вибір оптимального DLP-рішення залежить від потреб організації, типу IT-інфраструктури та пріоритетних сценаріїв захисту. Хмарно-орієнтовані рішення, такі як Microsoft Purview і Palo Alto Networks, є лідерами в захисті даних у хмарі, тоді як Broadcom Symantec і Forcepoint домінують у гібридних середовищах. Trellix є ефективним для фінансового сектору та захисту кінцевих точок. Подальші дослідження в цій галузі мають бути спрямовані на вдосконалення технологій машинного навчання, підвищення точності виявлення та спрощення процесів впровадження DLP-систем для малих і середніх підприємств.

Аналіз та проектування DLP-систем для захисту даних є ключовим елементом сучасної кібербезпеки, спрямованим на запобігання витоку конфіденційної інформації в умовах зростання обсягів даних і ускладнення кіберзагроз. Методи виявлення витоків, що базуються на контентно-орієнтованому та поведінково-орієнтованому аналізі, забезпечують ідентифікацію чутливих даних і аномальної активності, використовуючи технології великих даних, машинного навчання, UEBA та SIEM. Ці технології дозволяють обробляти значні обсяги інформації в реальному часі, підвищуючи точність і швидкість реагування на потенційні загрози. Блокування витоків, реалізоване через DLP-системи, системи управління доступом і шифрування, забезпечує ефективний захист даних у русі, зберіганні та використанні, хоча потребує ретельного налаштування для мінімізації помилкових спрацьовувань і протидії складним інсайдерським атакам. Порівняльний аналіз сучасних DLP-рішень, таких як Broadcom Symantec, Microsoft Purview, Forcepoint, Trellix і Palo Alto Networks, демонструє їхню різноманітність

за архітектурними підходами (мережеві, кінцеві, хмарні, гібридні) та функціональними можливостями, що дозволяє організаціям вибирати оптимальне рішення залежно від типу інфраструктури та сценаріїв використання. Ефективність DLP-систем залежить від їхньої інтеграції з іншими системами безпеки, відповідності регуляторним вимогам і здатності адаптуватися до нових викликів, таких як хмарні середовища та віддалена робота. Подальші дослідження мають бути спрямовані на вдосконалення алгоритмів машинного навчання, підвищення адаптивності DLP-систем до нових загроз, зокрема квантових атак і використання штучного інтелекту зловмисниками, а також на спрощення процесів впровадження для забезпечення доступності цих рішень для організацій різного масштабу.

3 ВПРОВАДЖЕННЯ ТА ОЦІНКА ЕФЕКТИВНОСТІ DLP-ТЕХНОЛОГІЙ

3.1 Етапи впровадження DLP-системи в інформаційну інфраструктуру

Впровадження системи запобігання витоку даних (DLP) у інформаційну інфраструктуру організації є складним процесом, що вимагає системного підходу та врахування специфіки інформаційного середовища. DLP-системи забезпечують захист конфіденційної інформації від несанкціонованого розголошення шляхом моніторингу, аналізу та блокування даних. Ефективність впровадження залежить від виконання послідовних етапів: аналізу потреб, проектування, розгортання, налаштування та тестування. У цьому підрозділі описано етапи впровадження з науково обґрунтованим підходом і прикладами програмного коду для реалізації базового моніторингу.

На першому етапі проводиться аналіз інформаційного середовища для визначення обсягу конфіденційних даних, каналів їх передачі та ризиків витоку. Необхідно класифікувати дані за типом і чутливістю, картувати потоки даних (електронна пошта, хмарні сервіси, USB-носії) та оцінити апаратно-програмну інфраструктуру. Для автоматизації аналізу мережевого трафіку використовується бібліотека Python Scapy. Нижче наведено код для моніторингу HTTP-запитів із виявленням чутливих даних:

```
from scapy.all import sniff, IP, TCP
import re
def packet_callback(packet):
    if packet.haslayer(TCP) and packet.haslayer(IP):
        if packet[TCP].dport == 80 or packet[TCP].sport
== 80:
            payload                                    =
bytes(packet[TCP].payload).decode('utf-8', errors='ignore')
            if        re.search(r'(GET|POST)\s+.*HTTP',
payload):
                print(f"HTTP     Request     detected:
{payload[:100]}...")
                if     re.search(r'\b\d{4}-\d{4}-\d{4}-
\d{4}\b', payload):
                    print("Warning:  Potential  credit
card number detected!")
```

```
    sniff(iface="eth0", prn=packet_callback, filter="tcp
port 80", store=0)
```

Код відстежує HTTP-трафік і виявляє чутливі дані, такі як номери кредитних карток, за допомогою регулярних виразів.

На основі аналізу обирається DLP-рішення, яке відповідає потребам організації за типом розгортання (мережеве, кінцеве, хмарне, гібридне) та функціональністю. Проектування передбачає створення політик безпеки, визначення точок інтеграції з SIEM, IAM і мережевими пристроями, а також планування масштабування системи.

Розгортання включає встановлення DLP-компонентів (серверів управління, агентів) і налаштування політик для розпізнавання чутливих даних та реагування на порушення. Для блокування несанкціонованої передачі даних можна використовувати iptables. Приклад скрипта для блокування HTTP/HTTPS-трафіку до підозрілого домену:

```
    #!/bin/bash
    DOMAIN_IP="93.184.216.34"
    iptables -A OUTPUT -p tcp --dport 80 -d $DOMAIN_IP -j
DROP
    iptables -A OUTPUT -p tcp --dport 443 -d $DOMAIN_IP -j
DROP
    echo "Blocked HTTP/HTTPS traffic to $DOMAIN_IP"
```

Скрипт запобігає передачі даних на вказану IP-адресу.

Тестування оцінює ефективність DLP-системи шляхом симуляції витоків, аналізу продуктивності та оптимізації політик для зменшення помилкових спрацьовувань. Для аналізу логів використовується Python із бібліотекою pandas:

```
    import pandas as pd
    logs = pd.read_csv('dlp_logs.csv')
    high_risk_events = logs[logs['risk_level'] >= 8]
    print("High-risk events:")
    print(high_risk_events[['timestamp', 'user', 'action',
'data_type']])
    data_type_counts = logs['data_type'].value_counts()
    print("\nData type distribution:")
    print(data_type_counts)
```

Код аналізує журнали DLP, виділяючи події з високим ризиком і статистику за типами даних.

Фінальний етап охоплює навчання персоналу, регулярний аналіз звітів, оновлення політик і аудит системи. Впровадження DLP-системи забезпечує захист даних шляхом інтеграції технологій і організаційних заходів. Програмні інструменти автоматизують моніторинг і аналіз, підвищуючи ефективність. Подальші дослідження мають фокусуватися на інтеграції штучного інтелекту та адаптації до нових каналів передачі даних.

3.2 Практичні аспекти налаштування та моніторингу DLP-рішень

Налаштування та моніторинг систем запобігання витоку даних (DLP) є критично важливими для забезпечення їхньої ефективності в захисті конфіденційної інформації. Ці процеси охоплюють конфігурацію політик безпеки, інтеграцію з інформаційною інфраструктурою, обробку подій і постійний аналіз для виявлення та нейтралізації загроз. Налаштування DLP-системи вимагає точного визначення чутливих даних, каналів їх передачі та правил реагування, тоді як моніторинг забезпечує контроль за подіями в реальному часі та адаптацію до нових викликів. У цьому підрозділі розглядаються практичні аспекти цих процесів із науковим обґрунтуванням і прикладом програмного коду для аналізу логів DLP.

Налаштування DLP-рішення починається з визначення чутливих даних, які потребують захисту. Це включає персональні дані, комерційну таємницю чи інтелектуальну власність, класифіковані за рівнем чутливості відповідно до стандартів, таких як ISO/IEC 27001. Політики безпеки формуються на основі шаблонів (регулярних виразів, відбитків даних) для ідентифікації чутливої інформації. Наприклад, регулярний вираз \b\d{4}-\d{4}-\d{4}-\d{4}\b може використовуватися для виявлення номерів кредитних карток.

Наступним кроком є конфігурація каналів моніторингу, що охоплюють мережевий трафік (електронна пошта, HTTP/HTTPS), кінцеві точки (USB-носії, друк) і хмарні сервіси (Google Drive, Microsoft 365). Для кожного каналу визначаються дії: моніторинг, попередження або блокування. Наприклад, DLP-система може блокувати передачу файлів із чутливими даними через незахищені

канали, повідомляючи адміністратора про інцидент. Інтеграція з системами SIEM і IAM забезпечує централізований збір подій і контроль доступу, знижуючи ризик інсайдерських загроз.

Важливим аспектом є мінімізація помилкових спрацьовувань (false positives), що досягається шляхом тестування політик на реальних даних і їхньої ітераційної оптимізації. Наприклад, виключення внутрішніх доменів із правил блокування може зменшити кількість помилкових тривог. Згідно з дослідженням Gartner, правильне налаштування DLP скорочує помилкові спрацьовування на 30%.

Моніторинг DLP-системи передбачає постійний аналіз подій, журналів і звітів для виявлення аномалій і потенційних витоків. Системи, такі як Broadcom Symantec DLP або Microsoft Purview, надають консолі управління для відстеження дій користувачів, аналізу трафіку та генерації звітів. Моніторинг у реальному часі дозволяє виявляти підозрілі дії, наприклад, масову передачу файлів на зовнішні ресурси, і автоматично застосовувати заходи, такі як блокування або карантин.

Для ефективного моніторингу використовуються технології аналізу поведінки (UEBA) і кореляції подій, що інтегруються з SIEM-системами. UEBA виявляє аномалії, порівнюючи поточну поведінку користувачів із базовими профілями, тоді як SIEM корелює події з різних джерел для ідентифікації складних атак.

Регулярний аналіз звітів дозволяє оцінювати ефективність політик і виявляти нові канали витоку, наприклад, месенджери або хмарні сховища.

Нижче наведено приклад Python-коду для аналізу журналів DLP із використанням бібліотеки pandas, що допомагає виявляти події з високим ризиком і оцінювати розподіл типів даних:

```
    import pandas as pd
    # Читання журналів DLP (CSV-файл)
    logs = pd.read_csv('dlp_logs.csv')
    # Фільтрація подій із високим рівнем ризику (≥8)
    high_risk_events = logs[logs['risk_level'] >= 8]
    print("High-risk events:")
    print(high_risk_events[['timestamp', 'user', 'action',
'data_type']])
    # Статистика за типами даних
```

```
data_type_counts = logs['data_type'].value_counts()
print("\nData type distribution:")
print(data_type_counts)
```

Цей код обробляє журнали DLP, виділяючи критичні інциденти та надаючи статистику за типами даних, що сприяє швидкому реагуванню на загрози.

Налаштування та моніторинг DLP-систем є ключовими для забезпечення захисту конфіденційної інформації.

Правильна конфігурація політик і каналів моніторингу знижує ризики витоку, тоді як аналіз подій у реальному часі дозволяє оперативно реагувати на загрози. Подальші дослідження мають бути спрямовані на інтеграцію DLP із технологіями штучного інтелекту для прогнозування аномалій і адаптацію до нових каналів передачі даних.

3.3 Оцінка ефективності DLP-систем та методи вдосконалення захисту даних

Оцінка ефективності систем запобігання витоку даних (DLP) є ключовим етапом для забезпечення їхньої здатності захищати конфіденційну інформацію від несанкціонованого розголошення, модифікації чи втрати. Ефективність DLP-систем визначається їхньою точністю у виявленні чутливих даних, швидкістю реагування на інциденти, мінімальним рівнем помилкових спрацьовувань і здатністю адаптуватися до динамічних кіберзагроз, таких як інсайдерські атаки, фішинг, атаки на хмарні сервіси чи використання соціальної інженерії. У сучасних умовах, коли обсяги даних зростають експоненціально, а зловмисники застосовують складні методи обходу захисту, вдосконалення DLP-систем набуває стратегічного значення. Цей підрозділ присвячено детальному аналізу методів оцінки ефективності DLP-систем і підходів до підвищення їхньої результативності, з акцентом на науково обґрунтовані методи та практичний приклад програмного коду для аналізу метрик ефективності.

Оцінка ефективності DLP-систем базується на комбінації кількісних і якісних показників, які відображають їхню здатність запобігати витокам і відповідати організаційним потребам. Кількісні метрики включають точність виявлення (True

Positive Rate, TPR), яка показує відсоток правильно ідентифікованих спроб витоку даних, і рівень помилкових спрацьовувань (False Positive Rate, FPR), що відображає частку безпечних дій, помилково позначених як загрози. Високий FPR може призводити до надмірного блокування легітимних операцій, знижуючи продуктивність користувачів і довіру до системи. Час реагування, тобто період між виявленням інциденту та його нейтралізацією (блокування, ізоляція, повідомлення), є ще одним важливим показником. Наприклад, сучасні DLP-системи, такі як Broadcom Symantec чи Microsoft Purview, здатні скоротити цей час до 10 секунд у реальному режимі. Покриття каналів передачі даних оцінює здатність системи контролювати всі можливі шляхи витоку, включаючи електронну пошту, USB-носії, хмарні сховища (Google Drive, OneDrive) і месенджери (Telegram, WhatsApp). Якісні показники охоплюють простоту адміністрування, інтеграцію з іншими системами безпеки (SIEM, IAM) і відповідність регуляторним стандартам, таким як GDPR, HIPAA чи PCI DSS, що є критично важливим для організацій у регульованих секторах, наприклад, фінансах чи охороні здоров’я.

Для оцінки ефективності застосовуються експериментальні методи, зокрема симуляція сценаріїв витоку даних у контрольованому середовищі. Такі сценарії можуть включати спроби надсилання файлів із чутливими даними (наприклад, номерами кредитних карток) через незахищені канали, копіювання даних на USB-носії або передачу через хмарні сервіси. Результати тестування аналізуються за допомогою статистичних методів для обчислення TPR і FPR. Крім того, аналіз журналів подій дозволяє оцінити поведінку системи в реальних умовах, виявити слабкі місця в політиках безпеки та оптимізувати їх. Наприклад, якщо система часто блокує внутрішні передачі даних через надмірно суворі правила, це може свідчити про необхідність уточнення шаблонів розпізнавання чутливих даних.

Нижче наведено приклад Python-коду для аналізу тестових журналів DLP-системи та обчислення ключових метрик ефективності:

```
import pandas as pd
from sklearn.metrics import confusion_matrix
```

```
    # Читання журналів DLP (CSV із полями: actual_label,
predicted_label)
    logs = pd.read_csv('dlp_test_logs.csv')

    # Отримання фактичних і передбачених міток (1 - витік,
0 - безпечно)
    y_true = logs['actual_label']
    y_pred = logs['predicted_label']
    # Обчислення матриці помилок
    tn,    fp,    fn,    tp    =    confusion_matrix(y_true,
y_pred).ravel()
    # Розрахунок метрик
    tpr = tp / (tp + fn) * 100  # True Positive Rate (%)
    fpr = fp / (fp + tn) * 100  # False Positive Rate (%)
    accuracy = (tp + tn) / (tp + tn + fp + fn) * 100  #
Точність (%)
    print(f"True Positive Rate (TPR): {tpr:.2f}%")
    print(f"False Positive Rate (FPR): {fpr:.2f}%")
    print(f"Accuracy: {accuracy:.2f}%")
```

Цей код обробляє журнали DLP, обчислюючи TPR, FPR і загальну точність, що дозволяє оцінити здатність системи розпізнавати витоки та уникати помилок. Дані для аналізу (наприклад, dlp_test_logs.csv) мають містити фактичні та передбачені мітки подій, отримані під час тестування.

Якісна оцінка передбачає аналіз зручності управління системою, її масштабованості та інтеграції з наявною інфраструктурою. Наприклад, централізовані консолі управління, які пропонують Forcepoint чи Palo Alto Networks, спрощують адміністрування великих організацій, скорочуючи час на обробку інцидентів. Інтеграція з SIEM-системами забезпечує кореляцію подій DLP із даними інших джерел, що підвищує точність виявлення складних атак, таких як багатоканальні інсайдерські схеми.

Для підвищення ефективності DLP-систем необхідно застосовувати інноваційні технологічні та організаційні підходи, які враховують сучасні виклики кібербезпеки.

Одним із ключових напрямів є використання машинного навчання (ML) для аналізу великих обсягів даних і прогнозування аномалій. Алгоритми ML, такі як класифікація, кластеризація чи виявлення аномалій, дозволяють системі адаптуватися до нових моделей поведінки користувачів і виявляти нетипові дії,

наприклад, масову передачу файлів у неробочий час. Такі технології, інтегровані в UEBA-модулі, підвищують точність виявлення інсайдерських загроз, які становлять значну частку інцидентів.

Іншим важливим аспектом є адаптація DLP-систем до хмарних середовищ, які стають основним місцем зберігання і обробки даних. Інтеграція через API з платформами, такими як Microsoft 365, AWS чи Google Workspace, забезпечує моніторинг і захист даних у реальному часі. Наприклад, хмарні DLP-системи можуть автоматично сканувати файли в OneDrive і блокувати їхній доступ у разі виявлення чутливої інформації, такої як персональні дані клієнтів. Однак ефективність таких рішень залежить від якості налаштування політик і здатності обробляти гібридні інфраструктури, де частина даних залишається на локальних серверах.

Розвиток квантових обчислень створює нові виклики для криптографічного захисту даних. Традиційні алгоритми, такі як AES чи RSA, можуть стати вразливими до квантових атак, тому впровадження квантово-стійких алгоритмів, рекомендованих NIST, є перспективним напрямом. Такі алгоритми забезпечують довгостроковий захист даних, особливо для критичних секторів, таких як енергетика чи оборона.

Автоматизація реагування на інциденти є ще одним способом підвищення ефективності. DLP-системи можуть бути налаштовані для автоматичного блокування підозрілих дій, наприклад, ізоляції пристрою чи припинення передачі даних через незахищені канали. Це скорочує час реагування і знижує ризик людських помилок. Наприклад, автоматичне застосування правил iptables або API-запитів до хмарних сервісів може миттєво зупинити витік.

Організаційні заходи, такі як регулярне навчання персоналу, відіграють важливу роль у зниженні ризиків, пов’язаних із людським фактором. Тренінги з кібербезпеки підвищують обізнаність працівників про ризики ненавмисних витоків, наприклад, через необережне надсилання даних через незахищені канали. Періодичний аудит DLP-системи дозволяє виявляти конфігураційні помилки та

адаптувати політики до нових загроз, таких як використання месенджерів чи 5G-мереж.

Для забезпечення максимальної ефективності DLP-систем необхідно проводити щоквартальні симуляції витоків для оцінки ключових метрик і виявлення слабких місць.

Інтеграція з SIEM-системами забезпечує комплексний аналіз подій, дозволяючи виявляти складні атаки, які охоплюють кілька каналів. Політики безпеки слід регулярно оновлювати, враховуючи нові канали передачі даних, такі як месенджери чи хмарні платформи, а також зміни в законодавстві. Генерація звітів про інциденти та їхній аналіз сприяє моніторингу відповідності стандартам і вдосконаленню стратегій захисту.

Оцінка ефективності DLP-систем дозволяє визначити їхню здатність запобігати витокам і відповідати вимогам організації. Поєднання кількісних метрик, таких як TPR і FPR, із якісними показниками, такими як інтеграція та зручність управління, формує цілісний підхід до аналізу.

Вдосконалення захисту даних через впровадження машинного навчання, адаптацію до хмарних середовищ, використання квантово-стійкої криптографії, автоматизацію реагування та навчання персоналу забезпечує стійкість до сучасних і майбутніх кіберзагроз.

Подальші дослідження мають бути спрямовані на розробку адаптивних DLP-систем, здатних прогнозувати складні атаки та протистояти квантовим обчисленням, а також на спрощення їхнього впровадження для організацій різного масштабу.

Розділ присвячено впровадженню, налаштуванню, моніторингу та оцінці ефективності DLP-систем, а також методам вдосконалення захисту даних. Впровадження DLP вимагає аналізу інфраструктури, класифікації чутливих даних і інтеграції з системами SIEM та IAM, що забезпечує контроль каналів витоку (електронна пошта, USB, хмара). Автоматизація за допомогою Scapy та iptables підвищує ефективність моніторингу та блокування.

Налаштування DLP передбачає створення політик для розпізнавання даних і мінімізації помилкових спрацьовувань, а моніторинг з UEBA та SIEM виявляє аномалії. Аналіз журналів через pandas ідентифікує критичні інциденти. Ефективність оцінюється метриками TPR, FPR і часом реагування, підтвердженими тестуванням і Python-аналітикою.

Вдосконалення захисту досягається завдяки машинному навчанню, хмарній адаптації, квантово-стійкій криптографії, автоматизації та навчанню персоналу. Рекомендації включають регулярне тестування, оновлення політик і аудит. Подальші дослідження мають фокусуватися на інтелектуальних DLP-системах і адаптації до 5G-мереж для підвищення кібербезпеки.

ВИСНОВКИ

Кваліфікаційна робота присвячена комплексному дослідженню теоретичних і практичних аспектів захисту даних у сучасних інформаційних системах. Актуальність теми зумовлена стрімким зростанням кіберзагроз, таких як інсайдерські витоки, фішинг, атаки на хмарні сервіси та соціальна інженерія, які спричиняють значні економічні збитки (понад 4,5 трлн дол. США у 2023 році за даними Verizon) і репутаційні втрати. В Україні проблема набуває особливого значення через геополітичні виклики та активну цифровізацію.

Теоретичний аналіз установив, що конфіденційна інформація, класифікована за типом (персональні дані, комерційна таємниця, інтелектуальна власність), рівнем чутливості та сферою застосування, є цінним активом, захист якого базується на тріаді CIA (конфіденційність, цілісність, доступність). Загрози витоку, що охоплюють внутрішні (34% інцидентів за Verizon) та зовнішні джерела, а також технічні, людські й організаційні канали, вимагають комплексного підходу до ідентифікації. Проактивні методи, такі як оцінка ризиків і аудит безпеки, поєднуються з реактивними, зокрема аналізом журналів і моделюванням загроз. Технології DLP, які використовують контентний і контекстний аналіз, забезпечують моніторинг, виявлення та блокування чутливих даних, але їхня ефективність залежить від адаптації до нових викликів, таких як хмарні технології та віддалена робота.

Практична частина роботи включає аналіз методів виявлення та блокування витоків, порівняння сучасних DLP-рішень (Broadcom Symantec, Microsoft Purview, Forcepoint, Trellix, Palo Alto Networks) за архітектурними підходами (мережеві, кінцеві, хмарні, гібридні) і функціональністю. Розроблено модель впровадження DLP-системи, яка охоплює картування потоків даних, інтеграцію з SIEM і IAM, налаштування політик і тестування. Автоматизація процесів за допомогою бібліотек Scapy, iptables і pandas забезпечує моніторинг мережевого трафіку, блокування підозрілих передач і аналіз журналів. Експериментальна оцінка ефективності за метриками TPR, FPR і часом реагування (до 10 секунд) підтвердила

високу результативність DLP за умови оптимізації політик і мінімізації помилкових спрацьовувань.

Для вдосконалення захисту даних запропоновано інтеграцію машинного навчання для прогнозування аномалій, адаптацію до хмарних платформ (Microsoft 365, AWS), впровадження квантово-стійкої криптографії, автоматизацію реагування та навчання персоналу. Ці заходи знижують ризики витоку, враховуючи сучасні тенденції, такі як зростання хмарних сервісів і потенційні квантові атаки. Результати роботи мають практичну цінність для організацій у фінансовому, промисловому та державному секторах, сприяючи зниженню кіберризиків, забезпеченню безперервності процесів і відповідності стандартам ISO 27001, GDPR, PCI DSS.

Подальші дослідження мають бути спрямовані на розробку інтелектуальних DLP-систем із прогнозними можливостями на основі штучного інтелекту, адаптацію до нових каналів передачі даних, зокрема 5G-мереж, і спрощення впровадження для малих і середніх організацій. Це сприятиме підвищенню рівня кібербезпеки в умовах глобальної цифрової трансформації та еволюції кіберзагроз.

ПЕРЕЛІК ПОСИЛАНЬ

1. ISO/IEC 27001:2013. Інформаційні технології. Методи забезпечення безпеки. Системи управління інформаційною безпекою. Вимоги. URL: https://www.iso.org/obp/ui/#iso:std:iso-iec:27001:ed-2:v1:en (дата звернення: 25.01.2025).
2. Cost of a Data Breach Report. *IBM*, 2020. URL: https://www.ibm.com/reports/data-breach (дата звернення: 25.01.2025).
3. Regulation (EU) 2016/679 (General Data Protection Regulation). *Official Journal of the European Union*, 2016. С. 1–88.
4. Про державну таємницю : Закон України від 21.01.1994 р. № 3855-XII. Дата оновлення: 30.10.2024. URL: https://zakon.rada.gov.ua/laws/show/3855-12#Text (дата звернення: 25.01.2025).
5. Guide for Mapping Types of Information and Information Systems to Security Categories. *NIST SP 800-60*, 2008. URL: https://csrc.nist.gov/pubs/sp/800/60/r2/iwd (дата звернення: 25.01.2025).
6. Health Insurance Portability and Accountability Act (HIPAA). *CDC*, 1996. URL: https://www.cdc.gov/phlp/php/resources/health-insurance-portability-and-accountability-act-of-1996-hipaa.html (дата звернення: 25.01.2025).
7. Payment Card Industry Data Security Standard (PCI DSS), Version 4.0. *PCI*, 2022. URL: https://www.pcisecuritystandards.org/document_library/?category=pcidss&document=pci_dss (дата звернення: 25.01.2025).
8. Data Breach Investigations Report. *Verizon*, 2021. 18 p. URL: https://www.verizon.com/business/resources/T8fc/reports/2021-dbir-executive-brief.pdf (дата звернення: 25.01.2025).
9. Cost of a Data Breach Report. *IBM Security*, 2020. URL: https://www.ibm.com/security/digital-assets/cost-data-breach-report/1Cost%20of%20a%20Data%20Breach%20Report%202020.pdf (дата звернення: 03.02.2025).

10. ISO/IEC 27005:2022. Інформаційна безпека, кібербезпека та захист конфіденційності. *Керівництво з управління ризиками інформаційної безпеки*. URL: https://www.iso.org/ru/standard/80585.html (дата звернення: 03.02.2025).
11. Bhadauria R., Sanyal S. Survey on Security Issues in Cloud Computing and Associated Mitigation Techniques // *International Journal of Computer Applications*. 2012. Vol. 47, No 18. PP. 47–66.
12. Inquiry Spotlight: Data Leak Prevention, Q1. *Forrester Research*, 2009. URL: https://www.forrester.com/report/Inquiry-Spotlight-Data-Leak-Prevention-Q1-2009/RES46788 (дата звернення: 03.02.2025).
13. Alneyadi S., Sithirasenan E., Muthukkumarasamy V. A survey on data leakage prevention systems // *Journal of Network and Computer Applications*. 2016. Vol. 62. PP. 137–152.
14. Liu S., Kuhn R. Data Loss Prevention // *IT Professional*. 2010. Vol. 12, No 2. PP. 10–13.
15. Tankard C. Data Leakage Prevention: The Need for a Comprehensive Approach // *Network Security*. 2012. No 5. PP. 5–8.
16. Kizza J. M. Guide to Computer Network Security. Computer Communications and Networks. *Springer International Publishing AG*, 2017. PP. 3-40. DOI: https://doi.org/10.1007/978-3-319-55606-2_1.
17. Cost of a Data Breach Report. *IBM*, 2023. URL: https://www.performance.gr/wp-content/uploads/Cost-of-a-Data-Breach-Report-2023.pdf (дата звернення: 03.02.2025).
18. Sommer R., Paxson V. Outside the Closed World: On Using Machine Learning for Network Intrusion Detection // *IEEE Symposium on Security and Privacy*. 2010. PP. 305–316.
19. Critical Capabilities for Security Information and Event Management. *Gartner*, 2022. URL: https://www.gartner.com/en/documents/4021424 (дата звернення: 03.02.2025).
20. Veeramachaneni K. et al. AI^2: Training a Big Data Machine to Defend // *IEEE International Conference on Big Data Security*. 2016. PP. 49–54.

21. Technology in support of national security. *MIT Lincoln Laboratory*, 2018. URL: https://apps.dtic.mil/sti/trecms/pdf/AD1100657.pdf (дата звернення: 26.02.2025).
22. Bauder B., Khoshgoftaar T. M. A Survey on the Use of Behavior Analytics in Cybersecurity // *Journal of Big Data*. 2018. Vol. 5, No 1. PP. 1–25.
23. Yousef Rasheed, Jazzar Mahmoud. Measuring the Effectiveness of User and Entity Behavior Analytics for the Prevention of Insider Threats *// Journal of Xi'an University of Architecture & Technology*. 2021. Vol. XIII. PP. 175–181.
24. Cost of a Data Breach Report. *IBM Security*, 2021. URL: https://www.dataendure.com/wp-content/uploads/2021_Cost_of_a_Data_Breach_-2.pdf (дата звернення: 26.02.2025)
25. Market Guide for Data Loss Prevention. *Gartner*, 2021. URL: https://info.zscaler.com/resources-industry-report-gartner-market-guide-dlp (дата звернення: 26.02.2025)
26. Alneyadi S. et al. Enhancing Data Leakage Prevention through Content and Context Analysis *// Computers & Security*. 2018. Vol. 77. PP. 672–686.
27. Sandhu R. et al. Role-Based Access Control Models // *IEEE Computer*. 1996. Vol. 29, No 2. PP. 38–47.
28. The Future of Identity and Access Management. *Forrester Research*, 2019. URL: https://www.forrester.com/report/The-Future-Of-Identity-And-Access-Management/RES136522?categoryid=2826672%3FelqTrackId%3D156685edabf349f09529479b0b4575af (дата звернення: 26.02.2025).
29. Stallings W. Cryptography and Network Security: Principles and Practice. *Pearson*, 2017. 815 с.
30. NIST SP 800-53. Security and Privacy Controls for Information Systems and Organizations. *NIST*, 2020. URL: https://csrc.nist.gov/pubs/sp/800/53/r5/upd1/final (дата звернення: 26.02.2025)
31. Cisco Secure Network Analytics: Data Loss Prevention. *Cisco Technical Report*, 2020. URL: https://www.cisco.com/site/us/en/learn/topics/security/what-is-data-loss-prevention-dlp.html (дата звернення: 26.02.2025).

32. Endpoint Protector: Technical Overview. CoSoSys Whitepaper. *CoSoSys*, 2023. URL: https://www.endpointprotector.com/ (дата звернення: 26.02.2025).
33. Microsoft Purview Data Loss Prevention: Capabilities and Features. *Microsoft Documentatio*n, 2023. URL: https://www.microsoft.com/en-us/security/business/information-protection/microsoft-purview-data-loss-prevention (дата звернення: 26.02.2025).
34. Forcepoint DLP: Unified Data Protection. *Forcepoint Technical Report*, 2023. URL: https://www.forcepoint.com/product/dlp-data-loss-prevention (дата звернення: 14.03.2025).
35. Symantec DLP: Exact Data Matching and Indexed Document Matching. *Broadcom Whitepaper*, 2022. URL: https://docs.broadcom.com/docs/data-loss-prevention-drive-total-protection-of-your-sensitive-data-en (дата звернення: 14.03.2025).
36. Data Loss Prevention Market by Offering, Solution Type, Service, Applications, Vertical and Regions – Global Forecast to 2028. *MarketsandMarkets*, 2023. URL: https://www.marketsandmarkets.com/PressReleases/data-loss-prevention.asp (дата звернення: 14.03.2025).
37. Data Loss Prevention: Enterprise Scale and Highly Accurate. *Broadcom*, 2023. URL: https://www.broadcom.com/products/cybersecurity/information-protection/data-loss-prevention (дата звернення: 14.03.2025).
38. Microsoft Purview: Compliance and Data Protection. *Microsoft Whitepaper*. URL: https://www.microsoft.com/en-us/security/business/microsoft-purview (дата звернення: 14.03.2025).
39. Integrating Forcepoint ONE Data Security with Forcepoint ONE SSE CASB. *Forcepoint Documentation*. URL: https://help.forcepoint.com/fpone/deploy/rhtml/0ED42C6C-15AC-4EB7-BF93-0E4951F24926.html (дата звернення: 14.03.2025).
40. Trellix DLP: Endpoint and Cloud Protection. *Trellix Technical Report*, 2023. URL: https://www.trellix.com/products/dlp/ (дата звернення: 14.03.2025).

41. Enterprise DLP: Cloud-Native Data Protection. *Palo Alto Networks Whitepaper*, 2023. URL: https://www.paloaltonetworks.com/sase/enterprise-data-loss-prevention (дата звернення: 15.04.2025).
42. Data Loss Prevention Market Size & Share – Growth Report. *MarketsandMarkets*, 2024. URL: https://www.marketsandmarkets.com/ResearchInsight/size-and-share-of-data-loss-prevention-advanced-technologies-market.asp (дата звернення: 15.04.2025).
43. Microsoft. How OneDrive safeguards your data in the cloud. *Microsoft Case Study*, 2023. URL: https://support.microsoft.com/en-us/office/how-onedrive-safeguards-your-data-in-the-cloud-23c6ea94-3608-48d7-8bf0-80e142edd1e1 (дата звернення: 15.04.2025).
44. Forcepoint DLP: Insider Threat Prevention with UEBA. *Forcepoint Whitepaper*, 2023. URL: https://www.forcepoint.com/sites/default/files/resources/datasheets/datasheet-forcepoint-behavioral-analytics-en_0_0_0_0_0_0_0_0.pdf (дата звернення: 15.04.2025).
45. The 2020 State of Security Operations. *Forrester Research*, 2020. URL: https://www.paloguard.com/datasheets/forrester-the-2020-state-of-security.pdf (дата звернення: 15.04.2025).
46. DLP Integration with Prisma Access. *Palo Alto Networks Documentation*, 2025. URL: https://docs.paloaltonetworks.com/prisma/prisma-access/preferred/2-2/prisma-access-panorama-admin/data-loss-prevention-on-prisma-access/dlp-on-prisma-access (дата звернення: 25.04.2025).