Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя
(повне найменування вищого навчального закладу)
Факультет комп'ютерно-інформаційних систем і програмної інженерії
(назва факультету)
Кафедра кібербезпеки
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

магістр
(освітній рівень)

на тему: "Оцінка ефективності заходів протидії фішингу в Microsoft Defender XDR засобами Advanced Hunting"

Виконав: студент VI курсу, групи СБм-61

Спеціальності:
125 «Кібербезпека та захист інформації»
(шифр і назва напряму підготовки, спеціальності)

Маслянка Т. В.
підпис (прізвище та ініціали)

Керівник Загородна Н. В.
підпис (прізвище та ініціали)

Нормоконтроль Стадник М. А.
підпис (прізвище та ініціали)

Завідувач кафедри Загородна Н. В.
підпис (прізвище та ініціали)

Рецензент Марценко С. В.
підпис (прізвище та ініціали)

м. Тернопіль – 2025

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра кібербезпеки
(повна назва кафедри)

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ Загородна Н.В.
(підпис) (прізвище та ініціали)
«__» _____ 2025 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня Магістр
(назва освітнього ступеня)

за спеціальністю 125 Кібербезпека та захист інформації
(шифр і назва спеціальності)

Студенту Маслянці Тарасу Володимировичу
(прізвище, ім'я, по батькові)

1. Тема роботи Оцінка ефективності заходів протидії фішингу в Microsoft Defender XDR засобами Advanced Hunting

Керівник роботи Загородна Наталія Володимирівна, кандидат технічний наук, зав. кафедри КБ
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «24» 11 2025 року № 4/7-1024

2. Термін подання студентом завершеної роботи 12.12.2025

3. Вихідні дані до роботи

4. Зміст роботи (перелік питань, які потрібно розробити)
Проаналізувати сучасні методи виявлення фішингу та засоби моніторингу безпеки
Визначити теоретико-методологічні засади оцінювання ефективності протидії фішингу
Розробити систему оцінювання ефективності заходів протидії фішингу
Охорона праці та безпека в надзвичайних ситуаціях
Висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)
Тема кваліфікаційної роботи. Проблематика та предмет дослідження. Мета роботи. Наукова новизна одержаних результатів. Джерела даних та реалізація. Метод оцінювання. Результати експерименту: якість детектування. Результати експерименту: час реагування. Рекомендації щодо оптимізації політик безпеки. Висновки та практична цінність.

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	Осухівська Г.М., к.т.н., доцент		
Безпека в надзвичайних ситуаціях	Стручок В.С., ст. викладач кафедри ОХ		

7. Дата видачі завдання 19.09.2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	19.09 – 30.09	Виконано
2.	Підбір джерел для аналізу методів виявлення фішингу та засобів моніторингу безпеки	01.10 – 03.10	Виконано
3.	Проведення аналізу методів виявлення фішингу та засобів моніторингу безпеки	04.10 – 08.10	Виконано
4.	Оформлення першого розділу	09.10 – 15.10	Виконано
5.	Підбір джерел для визначення теоретико-методологічних засад оцінювання ефективності протидії фішингу	16.10 – 18.10	Виконано
6.	Визначення теоретико-методологічних засад оцінювання ефективності протидії фішингу	19.10 – 24.10	Виконано
7.	Оформлення другого розділу	25.10 – 31.10	Виконано
8.	Практична реалізація методу оцінювання ефективності заходів протидії фішингу	01.11 – 20.11	Виконано
9.	Оформлення третього розділу	21.11 – 30.11	Виконано
10.	Виконання завдання до підрозділу «Охорона праці та безпека в надзвичайних ситуаціях»	01.12 – 02.12	Виконано
11.	Оформлення кваліфікаційної роботи	03.12 – 11.12	Виконано
12.	Нормоконтроль	12.12 – 14.12	Виконано
13.	Перевірка на плагіат	15.12 – 18.12	Виконано
14.	Попередній захист кваліфікаційної роботи	19.12 – 21.12	Виконано
15.	Захист кваліфікаційної роботи	22.12.2025	

Студент _____ Маслянка Т. В.
(підпис) (прізвище та ініціали)

Керівник роботи _____ Загородна Н. В.
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Оцінка ефективності заходів протидії фішингу в Microsoft Defender XDR засобами Advanced Hunting // ОР «Магістр» // Маслянка Тарас Володимирович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем і програмної інженерії, кафедра кібербезпеки, група СБм-61 // Тернопіль, 2025 // С. 60, рис. – 12, табл. – 0 , кресл. – 10, додат. – 1.

КЛЮЧОВІ СЛОВА: Microsoft Defender XDR, Advanced Hunting, KQL, фішинг.

Кваліфікаційну роботу присвячено розробці методу оцінювання ефективності заходів протидії фішингу в середовищі Microsoft Defender XDR. У роботі проаналізовано сучасний ландшафт фішингових загроз та обмеження стандартних засобів звітності систем захисту електронної пошти. Запропоновано метод розрахунку ключових показників ефективності на основі ретроспективного аналізу подій у журналах Advanced Hunting. Розроблено та протестовано набір унікальних KQL-запитів, які дозволяють автоматизувати процес аудиту якості роботи політик безпеки та виявляти приховані загрози, що оминули первинні фільтри. Результати дослідження можуть бути використані SOC-аналітиками для підвищення рівня захищеності корпоративної інфраструктури.

ABSTRACT

Evaluation of the Effectiveness of Anti-Phishing Measures in Microsoft Defender XDR using Advanced Hunting // Thesis of educational level «Master» // Taras Maslianka // Ternopil Ivan Puluj National Technical University, Faculty of Computer Information Systems and Software Engineering, Department of Cybersecurity, group SBm-61 // Ternopil, 2025 // p. 60, figs. 12, tbls. 0, drws. 10, apps. 1.

KEYWORDS: Microsoft Defender XDR, Advanced Hunting, KQL, phishing.

This thesis is devoted to the development of a method for evaluating the effectiveness of anti-phishing measures within the Microsoft Defender XDR environment. The paper analyzes the modern phishing threat landscape and the limitations of standard reporting tools in email security systems. A method for calculating key performance indicators based on a retrospective analysis of events in Advanced Hunting logs is proposed. A set of unique KQL queries has been developed and tested to automate the audit of security policy performance and identify hidden threats that bypassed primary filters. The research results can be utilized by SOC analysts to enhance the security posture of corporate infrastructure.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

XDR – Extended Detection and Response.

KQL – Kusto Query Language.

BEC – Business Email Compromise.

AiTM – Adversary-in-the-Middle.

URL – Uniform Resource Locator.

SOC – Security Operations Center.

ZAP – Zero-hour Auto Purge.

MTTD – Mean Time to Detect.

MTTR – Mean Time to Respond.

ВСТУП

Актуальність теми. В умовах стрімкої цифровізації бізнес-процесів електронна пошта залишається основним вектором кібератак. Сучасні фішингові кампанії еволюціонували від масових розсилок до складних цільових атак, які використовують методи соціальної інженерії та обфускації контенту для обходу традиційних сигнатурних фільтрів. Системи класу XDR (Extended Detection and Response), такі як Microsoft Defender, надають потужні інструменти для захисту, проте стандартні звіти часто відображають лише статичну статистику заблокованих загроз, не даючи об’єктивної оцінки реальної ефективності захисту в динаміці. Використання мов запитів для глибокого аналізу телеметрії дозволяє створити більш гнучкі метрики, які можна використовувати для всестороннього аналізу ефективності впроваджених заходів протидії фішингу. Тому розробка методу оцінки ефективності протидії фішингу засобами Advanced Hunting та мови KQL є актуальним науково-прикладним завданням, що дозволяє підвищити кіберстійкість організацій.

Мета і задачі дослідження. Метою роботи є розробка методу для автоматизованого оцінювання ефективності механізмів протидії фішингу в середовищі Microsoft Defender XDR.

Завданнями дослідження є:

– проведення аналізу архітектури Microsoft Defender XDR та структури даних Advanced Hunting для визначення ключових індикаторів компрометації та ефективності;

– розробка метрик та методу оцінювання якості ідентифікації фішингу, що базується специфіці роботи Microsoft Defender XDR та результатах рестроспективного аналізу;

– створення алгоритмів та програмних запитів мовою KQL для розрахунку динамічних метрик, зокрема коефіцієнта помилкових рішень та латентності виявлення загроз;

– експериментальна перевірка запропонованих методів на реальних даних;

– розробка рекомендації щодо оптимізації політик безпеки на основі отриманих метрик.

Об’єкт дослідження. Процес виявлення та нейтралізації фішингових атак у корпоративних поштових системах, захищених хмарними платформами класу XDR.

Предмет дослідження. Методи, моделі та програмні засоби оцінювання ефективності систем захисту від фішингу з використанням інструментарію розширеного пошуку загроз (Advanced Hunting) та мови запитів KQL.

Наукова новизна одержаних результатів кваліфікаційної роботи. Удосконалено методи оцінки надійності поштової системи захисту у платформі Microsoft Defender XDR шляхом застосування ретроспективного аналізу подій засобами Advanced Hunting, що дозволяє виявляти приховані загрози без розгортання додаткових агентів моніторингу та забезпечує верифікацію автоматичних вердиктів системи безпеки на основі результатів ретроспективного аналізу.

Практичне значення одержаних результатів. Практична цінність роботи полягає у створенні бібліотеки адаптивних KQL-запитів та аналітичних правил, які можуть бути безпосередньо імплементовані в робочі процеси SOC для моніторингу якості роботи Microsoft Defender. Запропоновані рішення дозволяють скоротити час на підготовку звітів про ефективність системи безпеки та надають обґрунтовані дані для налаштування антифішингових політик, що підтверджується можливістю їх використання в реальній інфраструктурі підприємства.

Апробація результатів магістерської роботи. Основні результати проведених досліджень обговорювались на: XIII науково-технічній конференції «Інформаційні моделі, системи та технології» (м.Тернопіль, Україна).

Публікації. Основні результати кваліфікаційної роботи опубліковано у працях конференції (див. Додаток А).

РОЗДІЛ 1 АНАЛІЗ СУЧАСНИХ МЕТОДІВ ВИЯВЛЕННЯ ФІШИНГУ ТА ЗАСОБІВ МОНІТОРИНГУ БЕЗПЕКИ

1.1 Еволюція фішингових атак та проблематика їх виявлення у корпоративних мережах

В умовах глобальної цифровізації суспільства та переходу бізнес-процесів у віртуальний простір забезпечення конфіденційності, цілісності та доступності інформації стає критично важливим завданням. Серед чисельних загроз інформаційній безпеці особливе місце посідає фішинг, який за останні десятиліття трансформувався з примітивних спроб обману користувачів у високотехнологічну індустрію кіберзлочинності. Електронна пошта залишається основним каналом ділової комунікації, і саме тому вона є головним вектором атак для зловмисників, які намагаються отримати несанкціонований доступ до корпоративних мереж. Статистичні дані провідних аналітичних агенцій свідчать, що переважна більшість інцидентів безпеки, які призводять до значних фінансових збитків, починається саме з фішингового листа [1]. Розуміння еволюції цих загроз та аналіз обмежень існуючих методів захисту є необхідною умовою для побудови ефективної системи кібербезпеки.

Історичний розвиток фішингу демонструє чітку тенденцію до ускладнення методів соціальної інженерії та технічних засобів доставки шкідливого контенту. На початкових етапах розвитку інтернету фішингові кампанії мали масовий характер і були спрямовані на широке коло користувачів без урахування їхніх індивідуальних особливостей. Такі атаки, відомі як «спрей-енд-прей», покладалися на закон великих чисел, де навіть незначний відсоток успішних спроб приносив зловмисникам прибуток. Однак з розвитком спам-фільтрів та підвищенням обізнаності користувачів ефективність таких методів почала стрімко знижуватися. Це спонукало кіберзлочинців до зміни тактики та переходу до більш цілеспрямованих атак.

Сучасний етап розвитку фішингових загроз характеризується переходом до персоналізованого впливу на жертву. Зловмисники проводять ретельну розвідку,

використовуючи відкриті джерела інформації, соціальні мережі та корпоративні вебсайти для збору даних про співробітників цільової організації [2]. Це дозволяє створювати переконливі повідомлення, які імітують стиль спілкування колег, керівництва або партнерів. Такий підхід значно ускладнює автоматичне виявлення загроз, оскільки листи можуть не містити класичних ознак фішингу, таких як граматичні помилки, підозрілі вкладення або посилання на відомі шкідливі ресурси. Замість цього атака будується на довірі та авторитеті, експлуатуючи людський фактор, який залишається найслабшою ланкою в системі захисту інформації [3].

Однією з найбільш небезпечних форм сучасного фішингу є компрометація ділової електронної пошти, відома як BEC-атаки. Цей вид шахрайства спрямований на компанії, які проводять електронні платежі та працюють з іноземними постачальниками. Особливістю BEC-атак є відсутність шкідливого програмного забезпечення в тілі листа. Атака базується виключно на методах соціальної інженерії та маніпуляції [4]. Зловмисники можуть зламати або підробити поштову скриньку керівника компанії та надіслати вказівку фінансовому відділу здійснити терміновий переказ коштів на підконтрольний рахунок. Оскільки лист надходить з легітимної адреси або візуально ідентичної до неї, а текст повідомлення не містить підозрілих посилань, традиційні засоби захисту, такі як антивіруси та шлюзи безпеки, часто виявляються безсилими. Виявлення таких атак вимагає аналізу контексту листування, поведінкових патернів користувачів та аномалій у фінансових операціях, що є складним завданням для класичних систем моніторингу.

Проблематика виявлення фішингу ще більше ускладнюється з появою атак типу «супротивник посередині» (AiTM). Ця технологія дозволяє зловмисникам обходити багатофакторну автентифікацію, яка тривалий час вважалася надійним захистом від крадіжки облікових записів [5]. У класичній схемі фішингу користувач вводить свої дані на підробленому сайті, але якщо у нього увімкнена двофакторна автентифікація, зловмисник не зможе увійти в систему без одноразового коду. Атаки AiTM вирішують цю проблему шляхом використання проксі-сервера, який розташовується між жертвою та легітимним сервісом. Коли

користувач переходить за посиланням, він потрапляє на сервер зловмисника, який у реальному часі транслює запити на справжній сайт. Жертва вводить логін, пароль та код підтвердження, які проксі-сервер передає на сервіс, успішно проходячи автентифікацію. У результаті зловмисник перехоплює не лише облікові дані, а й сесійний токен, що дозволяє отримати повний доступ до акаунту користувача без необхідності повторного введення пароля. Такі атаки становлять значну загрозу, оскільки для користувача процес входу виглядає абсолютно нормальним, а система безпеки не фіксує помилок входу.

Еволюція технічних засобів захисту змушує зловмисників постійно вдосконалювати методи доставки шкідливого контенту. Одним із нових трендів стало використання легітимних хмарних сервісів для розміщення фішингових сторінок. Ця тактика, відома як «життя за рахунок землі», передбачає використання довіреної інфраструктури, такої як файлові сховища або сервіси для створення онлайн-форм, для обходу репутаційних фільтрів [6]. Посилання на документ, розміщений на відомому хмарному ресурсі, зазвичай не блокується системами безпеки, оскільки домен має високу репутацію. Користувач, довіряючи знайомому інтерфейсу, з більшою ймовірністю перейде за посиланням та виконає дії, до яких його спонукає зловмисник. Це створює серйозну проблему для систем фільтрації контенту, які змушені балансувати між блокуванням загроз та забезпеченням доступу до необхідних для роботи ресурсів.

Ще одним викликом для сучасних систем виявлення є використання QR-кодів у фішингових кампаніях, що отримало назву «квішинг». Зловмисники вбудовують шкідливі посилання у графічні коди, які розміщуються в тілі електронного листа або у вкладених документах. Оскільки більшість традиційних засобів захисту електронної пошти аналізують текстовий вміст та URL-адреси, зображення з QR-кодом може пройти перевірку безперешкодно. Крім того, сканування коду зазвичай відбувається за допомогою мобільного пристрою, який часто знаходиться поза периметром корпоративного захисту та не має встановлених засобів моніторингу. Це призводить до того, що атака переноситься з захищеного корпоративного середовища на особистий пристрій користувача, де можливості контролю з боку служби безпеки є обмеженими.

Зростання популярності квішингу вимагає впровадження нових технологій розпізнавання зображень та аналізу вмісту кодів у реальному часі.

Окремої уваги заслуговує вплив штучного інтелекту на еволюцію фішингових атак. Доступність великих мовних моделей дозволила кіберзлочинцям автоматизувати процес створення фішингових повідомлень. Якщо раніше однією з головних ознак фішингу були помилки в тексті та неприродна мова, то сучасні інструменти штучного інтелекту дозволяють створювати бездоганні листи будь-якою мовою, адаптуючи тон та стиль повідомлення під конкретну жертву [7]. Це дозволяє проводити масовані атаки з високим рівнем персоналізації, що раніше було доступно лише висококваліфікованим групам зловмисників. Крім того, штучний інтелект використовується для створення поліморфного коду шкідливих програм, який змінюється при кожному завантаженні, ускладнюючи його виявлення сигнатурними методами антивірусного захисту.

Аналіз фінансових наслідків фішингових атак демонструє їх руйнівний потенціал для економіки. Згідно зі звітом Центру скарг на інтернет-злочини ФБР, збитки від кіберзлочинності продовжують зростати з кожним роком, досягаючи мільярдів доларів. Левова частка цих збитків припадає на інциденти, пов’язані з компрометацією ділової пошти та шахрайством з інвестиціями. Вартість одного успішного інциденту для організації може вимірюватися сотнями тисяч доларів, не враховуючи репутаційних втрат та витрат на відновлення систем. Це підкреслює економічну доцільність інвестицій у сучасні засоби захисту та навчання персоналу. Проте, як показує практика, покладання виключно на технічні засоби блокування є недостатнім, оскільки жодна система не може гарантувати стовідсотковий захист від загроз, що базуються на людській психології.

Суттєвою проблемою у боротьбі з фішингом є часовий лаг між появою нової загрози та оновленням баз даних засобів захисту. Традиційні системи безпеки часто працюють у реактивному режимі, блокуючи загрози лише після того, як вони були ідентифіковані та додані до чорних списків. У проміжок часу між початком атаки та оновленням сигнатур інфраструктура залишається вразливою.

Дослідження показують, що значна частина користувачів відкриває фішингові листи в перші хвилини після їх отримання. Це явище, відоме як «нульовий пацієнт», створює ризик швидкого поширення загрози мережею до моменту спрацювання автоматичних систем захисту. Для вирішення цієї проблеми необхідні механізми ретроспективного аналізу та автоматичного вилучення листів, які вже були доставлені до поштових скриньок користувачів, але згодом були визнані шкідливими.

У контексті корпоративної безпеки важливу роль відіграє видимість подій та можливість кореляції даних з різних джерел. Фішингова атака рідко обмежується одним листом; вона часто є частиною складнішого ланцюжка дій, що включає перехід на вебсайт, завантаження файлу, запуск процесу на кінцевій точці та спробу горизонтального переміщення мережею. Традиційні засоби захисту часто працюють ізольовано, аналізуючи лише окремі аспекти атаки. Поштовий шлюз бачить лист, веб-фільтр бачить URL-адресу, а антивірус на комп’ютері бачить файл. Відсутність єдиної картини подій ускладнює виявлення комплексних атак та уповільнює реагування на інциденти. Сучасні підходи до безпеки, такі як розширене виявлення та реагування (XDR), спрямовані на об'єднання телеметрії з усіх компонентів інфраструктури в єдину платформу для аналізу. Це дозволяє виявляти складні патерни атак, які залишаються непоміченими для окремих засобів захисту.

Обмеження існуючих метрик оцінки ефективності систем захисту також є важливою частиною проблематики. Більшість організацій оцінюють свою захищеність на основі кількості заблокованих листів або результатів навчальних симуляцій фішингу. Однак ці показники не дають повної картини реальної ефективності. Високий відсоток блокування може свідчити про велику кількість спаму, а не про якість захисту від цільових атак. Результати симуляцій часто не корелюють з реальною поведінкою користувачів під час справжньої атаки. Критично важливими метриками, які часто ігноруються, є час виявлення загрози, рівень хибнопозитивних та хибнонегативних спрацювань, а також час, витрачений на розслідування та нейтралізацію інциденту. Відсутність

інструментів для автоматизованого розрахунку цих показників ускладнює прийняття обґрунтованих рішень щодо налаштування політик безпеки.

Стандартні консолі управління засобами безпеки часто надають лише узагальнену статистику, яка не дозволяє глибоко аналізувати інциденти. Аналітикам SOC (Security Operations Center) доводиться вручну збирати інформацію з різних журналів подій, що є часомістким та неефективним процесом. В умовах великого потоку сповіщень це призводить до так званої «втоми від тривог», коли аналітики можуть пропустити реальну загрозу серед великої кількості хибних спрацювань. Для вирішення цієї проблеми необхідні інструменти, які дозволяють автоматизувати процес пошуку загроз та аналізу даних. Використання спеціалізованих мов запитів, таких як KQL, дозволяє створювати гнучкі правила детектування, адаптовані до специфіки конкретної інфраструктури, та проводити глибокий аналіз телеметрії для виявлення прихованих загроз.

Особливої актуальності набуває питання захисту від внутрішніх загроз, які можуть виникати внаслідок компрометації облікових записів співробітників. Якщо зловмисник отримує доступ до легітимного акаунту, він може використовувати його для розсилки фішингових листів колегам або партнерам. Оскільки такі листи надходять з довіреного джерела всередині периметра безпеки, вони часто не проходять повної перевірки на шлюзах, налаштованих переважно на фільтрацію зовнішнього трафіку. Виявлення внутрішнього фішингу вимагає аналізу поведінкових аномалій, таких як нетиповий час відправки повідомлень, масова розсилка або нехарактерний зміст листів. Це підкреслює необхідність впровадження систем поведінкової аналітики та моніторингу внутрішнього поштового трафіку.

Узагальнюючи викладене, можна стверджувати, що еволюція фішингових атак призвела до появи загроз, які ефективно обходять традиційні засоби захисту. Використання методів соціальної інженерії, обфускації контенту, легітимних хмарних сервісів та технологій штучного інтелекту робить сучасний фішинг надзвичайно небезпечним інструментом у руках кіберзлочинців. Статичні правила фільтрації та сигнатурний аналіз вже не можуть забезпечити достатній

рівень безпеки. Ефективна протидія вимагає комплексного підходу, який поєднує передові технології детектування, глибокий аналіз даних, автоматизацію реагування та постійне навчання персоналу. Перехід від реактивного захисту до проактивного пошуку загроз засобами Advanced Hunting та використання детальної телеметрії є логічним кроком у розвитку систем корпоративної безпеки, що дозволить виявляти та нейтралізувати атаки на ранніх стадіях, мінімізуючи потенційні збитки.

Зростання складності інфраструктури та розмиття периметра безпеки через віддалену роботу створюють додаткові вектори атак. Мобільні пристрої, домашні мережі та особисті комп'ютери співробітників стають частиною корпоративного середовища, збільшуючи поверхню атаки. Фішингові кампанії все частіше спрямовані на викрадення облікових даних для доступу до хмарних ресурсів та SaaS-додатків. Це вимагає від організацій впровадження моделі «нульової довіри» (Zero Trust), де кожен запит на доступ підлягає перевірці, незалежно від його походження. У такій моделі аналіз контексту автентифікації та поведінки користувача стає ключовим елементом захисту, а інструменти XDR відіграють роль центрального вузла для збору та аналізу інформації про загрози.

Таким чином, проблема оцінки ефективності заходів протидії фішингу є багатогранною та актуальною. Вона охоплює технічні, організаційні та людські аспекти кібербезпеки. Розробка методів, які дозволили б об'єктивно вимірювати якість роботи систем захисту в умовах постійної зміни ландшафту загроз, є важливим науковим та практичним завданням. Використання можливостей платформи Microsoft Defender XDR та мови запитів KQL відкриває нові перспективи для створення адаптивних та ефективних механізмів виявлення та реагування на фішингові атаки, що і буде предметом подальшого дослідження в даній роботі. Зокрема, реалізація таких підходів дозволить не лише оптимізувати наявні інструменти моніторингу, але й розробити нові сценарії автоматичного реагування на інциденти. Це, у свою чергу, забезпечить мінімізацію операційних ризиків та підвищення загальної кіберстійкості інформаційної інфраструктури підприємства.

1.2 Огляд архітектури та можливостей Microsoft Defender XDR для захисту поштового трафіку

Для успішної реалізації завдань кваліфікаційної роботи та побудови адекватної моделі оцінки ефективності необхідно провести глибокий аналіз архітектури платформи Microsoft Defender XDR, яка виступає середовищем для проведення дослідження. Ця система представляє собою комплексне рішення класу Extended Detection and Response, яке виходить за межі традиційних засобів захисту, об'єднує захист кінцевих точок, ідентичностей, хмарних додатків та електронної пошти в єдину інтегровану екосистему безпеки. В основі архітектури лежить принцип наскрізної видимості, який, на відміну від ізольованих рішень, збирає телеметрію з різних доменів інфраструктури, нормалізує її та застосовує алгоритми машинного навчання для виявлення складних, багатоступеневих атак.

Загальну архітектуру платформи Microsoft Defender XDR та взаємодії її компонентів зображено на рисунку 1.1.

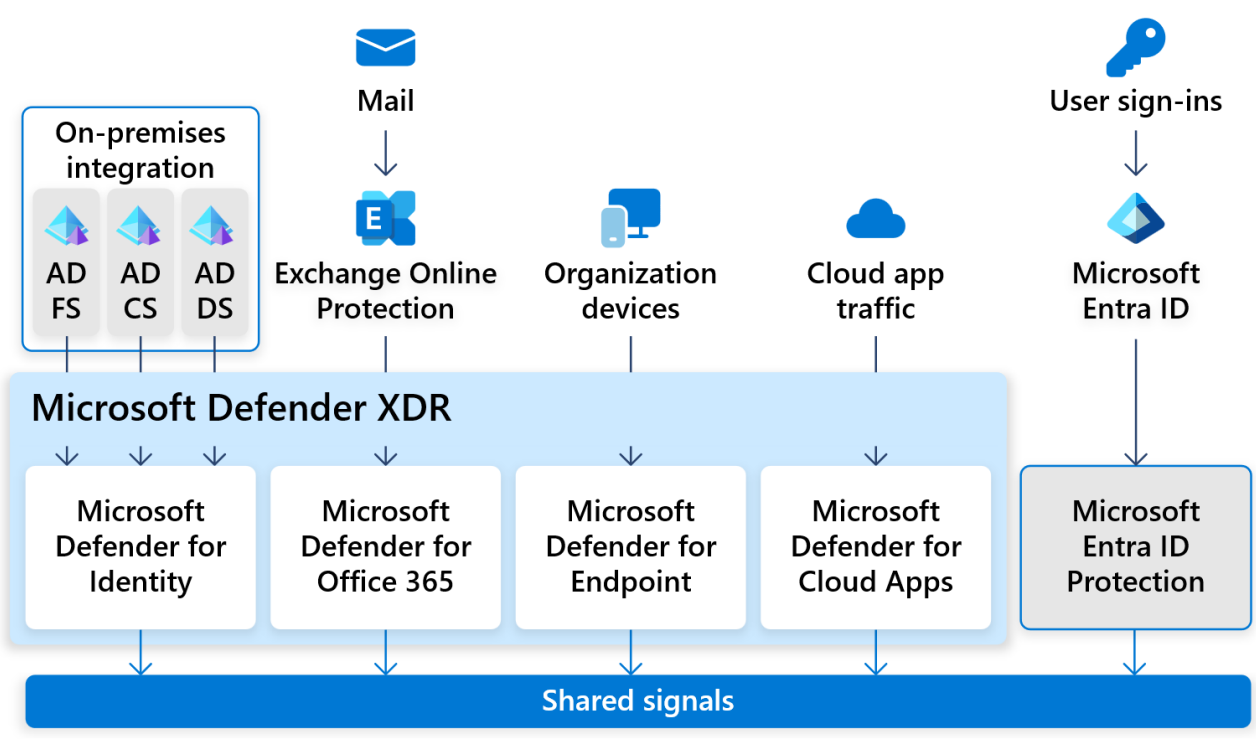


Рисунок 1.1 – Загальна архітектура платформи Microsoft Defender XDR

Центральним елементом платформи є єдиний портал Microsoft Defender Portal, який надає аналітикам консолідований інтерфейс для моніторингу інцидентів. Архітектурно рішення складається з кількох ключових стовпів безпеки, що функціонують у тісній взаємодії. Зокрема, компонент Defender for Endpoint відповідає за захист робочих станцій та серверів, використовуючи поведінковий аналіз процесів, тоді як Defender for Identity аналізує сигнали Active Directory для виявлення атак на ідентичність та горизонтального переміщення зловмисників мережею. Контроль над використанням хмарних сервісів та тіньовим IT забезпечує Defender for Cloud Apps. Проте ключовим компонентом для даного дослідження є модуль Defender for Office 365, який відповідає за захист корпоративних комунікацій та функціонує поверх базового захисту Exchange Online Protection [8]. Розуміння логіки роботи цього модуля та ієрархії шарів фільтрації є критично важливим для коректної інтерпретації отриманих даних.

Процес обробки вхідної кореспонденції в досліджуваній системі відбувається багатоетапно, формуючи складний фільтраційний стек, як це зображено на рисунку 1.2.

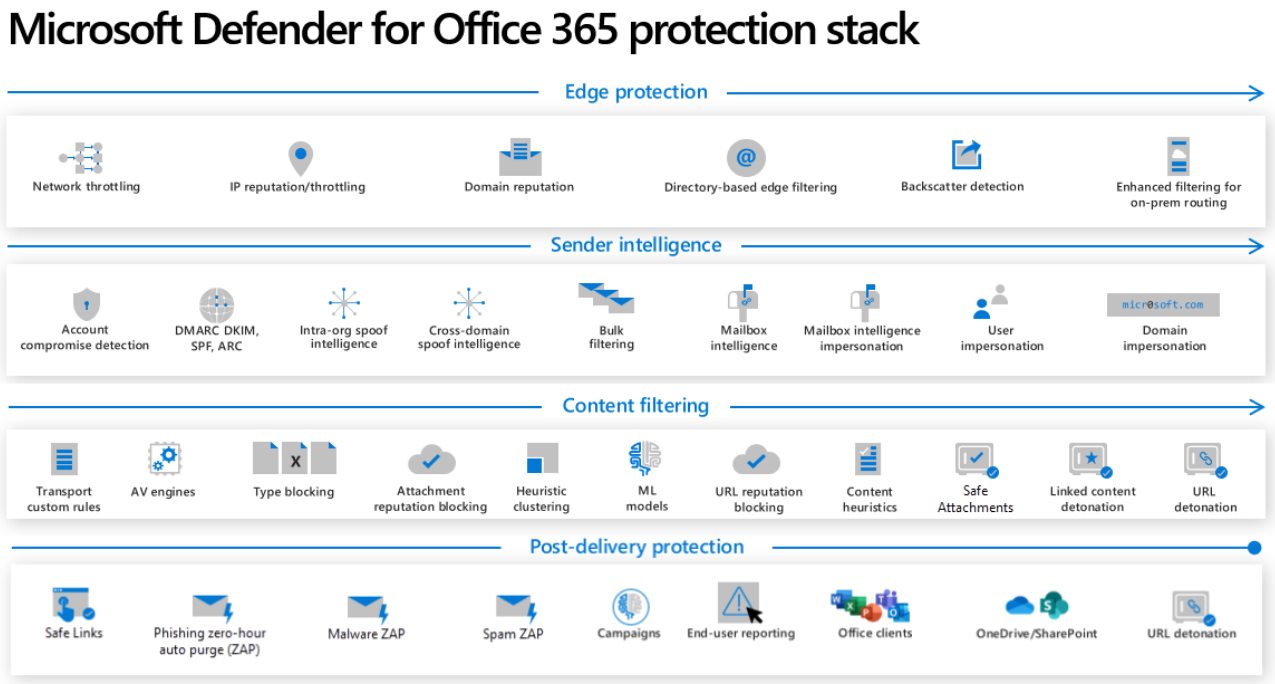


Рисунок 1.2 – Стек фільтрації Microsoft Defender for Office 365

На першому, мережевому рівні, застосовується фільтрація за репутацією відправника та IP-адреси, що дозволяє відсіяти значну частину масового спаму та відомих шкідливих розсилок ще до моменту завантаження повного вмісту листа, що суттєво економить обчислювальні ресурси системи.

Після проходження мережевого бар'єра вступають в дію механізми перевірки автентичності відправника. Система послідовно аналізує заголовки листа на відповідність низці галузевих стандартів. Насамперед перевіряється SPF (Sender Policy Framework) для підтвердження права конкретної IP-адреси надсилати пошту від імені домену. Далі аналізується цифровий підпис листа за протоколом DKIM (DomainKeys Identified Mail), що гарантує цілісність повідомлення. Завершує цей процес перевірка політики DMARC, яка визначає алгоритм дій поштового сервера у випадку, якщо попередні перевірки не були успішними. Зазначені протоколи формують фундамент для запобігання спуфінгу доменів, що є типовим вектором для фішингових кампаній.

Втім, найбільш інтелектуально місткі процеси відбуваються на етапі глибокого контентного аналізу. Для обробки вкладень використовується технологія Safe Attachments, яка базується на методі динамічного аналізу, відомому як «детонація». Вкладення, які не мають відомих сигнатур загроз, відкриваються в ізольованому віртуальному середовищі Microsoft, де емулюється поведінка користувача, включаючи рух мишкою та кліки, а система моніторить будь-які зміни в реєстрі, спроби мережевих з’єднань та запуску сторонніх процесів.

Такий підхід дозволяє виявляти нові, раніше невідомі шкідливі програми та експлойти нульового дня. Важливою особливістю реалізації є функція динамічної доставки, завдяки якій користувач отримує тіло листа майже миттєво, тоді як вкладення продовжує скануватися. У цей проміжок часу замість файлу відображається повідомлення про перевірку, що мінімізує вплив заходів безпеки на швидкість бізнес-процесів.

Захист від шкідливих посилань реалізується через технологію Safe Links. Її відмінність від традиційних фільтрів полягає в тому, що перевірка URL-адреси відбувається не лише статично в момент отримання листа, а безпосередньо в

момент переходу користувача за посиланням. Система автоматично переписує оригінальні посилання в листі, спрямовуючи трафік через захищений проксі-сервер Microsoft. Коли користувач натискає на таке посилання, відбувається миттєва перевірка цільового ресурсу за актуальними базами загроз. Це критично важливо для протидії атакам, де зловмисники спочатку надсилають посилання на легітимний ресурс, а вже після доставки листа підміняють вміст сторінки на фішинговий, намагаючись обійти первинну фільтрацію.

Окремої уваги в контексті дослідження заслуговує здатність системи до ретроспективного аналізу, що реалізується через механізм Zero-hour Auto Purge (ZAP). Ця технологія вирішує проблему латентності виявлення, коли лист початково визнається безпечним і доставляється користувачу, але згодом, на основі нових даних розвідки загроз, його статус змінюється на шкідливий.

Вигляд інтерфейсу налаштувань технології ZAP зображено на рисунку 1.3.

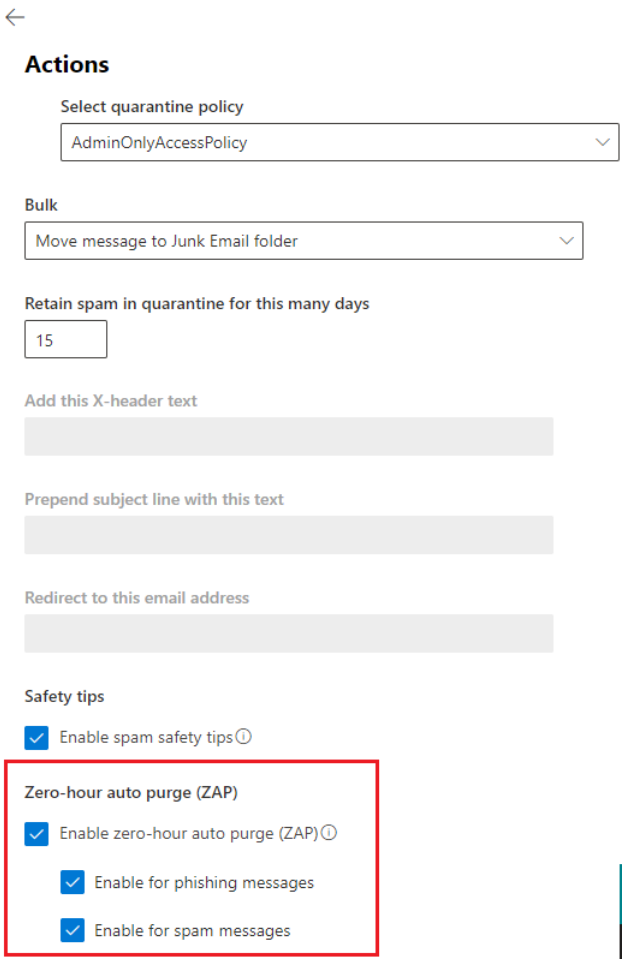


Рисунок 1.3 – Налаштування ZAP в Microsoft Defender XDR

Механізм ZAP постійно моніторить скриньки користувачів і у разі ретроспективного виявлення загрози автоматично вилучає лист та переміщує його до карантину [9], навіть якщо користувач вже прочитав повідомлення. Для дослідника ця функція є надзвичайно цінним джерелом даних, оскільки події спрацювання ZAP фактично є маркерами помилок першого ешелону захисту, тобто помилково негативними спрацюваннями на момент доставки.

Ще однією важливою складовою архітектури є підсистема автоматизованого розслідування та реагування. Вона імітує дії аналітика безпеки, автоматично запускаючи сценарії розслідування при виявленні підозрілого листа. Система самостійно шукає схожі листи в організації, перевіряє активність користувачів та пропонує адміністратору консолідовані дії для усунення загрози. Приклад графу розслідування інциденту в Microsoft Defender XDR зображено на рисунку 1.4.

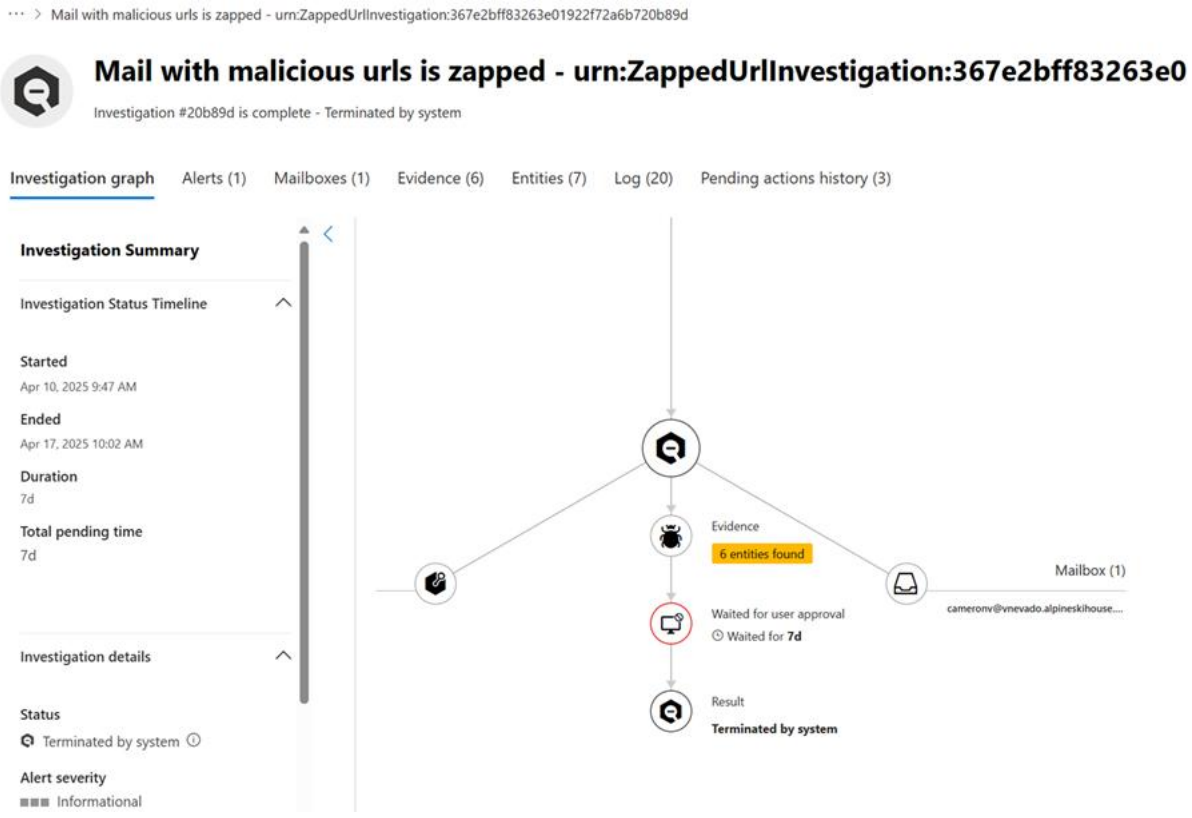


Рисунок 1.4 – Граф розслідування інциденту в Microsoft Defender XDR

Усі події, пов’язані з обробкою пошти, детектуванням загроз, роботою автоматизованих систем та діями користувачів, агрегуються в єдиному сховищі даних, доступ до якого надається через інструмент Advanced Hunting. Робота з

цими даними здійснюється за допомогою мови запитів Kusto Query Language (KQL), яка оптимізована для обробки великих масивів логів та дозволяє проводити глибокий аналіз і кореляцію подій. Схема даних включає кілька ключових таблиць, що становлять безпосередній інтерес для даного дослідження. Зокрема, таблиця EmailEvents містить вичерпні метадані про всі оброблені листи, включаючи інформацію про відправника, отримувача, початковий вердикт фільтра та застосовані політики. Дії, виконані системою після доставки, такі як спрацювання ZAP або ручне видалення адміністратором, фіксуються в таблиці EmailPostDeliveryEvents. Інформація про взаємодію користувачів із посиланнями, захищеними технологією Safe Links, зберігається в таблиці UrlClickEvents, що дозволяє відстежити переходи за шкідливими ресурсами [10]. Додатково використовуються таблиці AlertInfo та AlertEvidence, які містять дані про інциденти, автоматично сформовані системою безпеки на основі кореляції подій.

Інтерфейс Advanced Hunting зображено на рисунку 1.5.

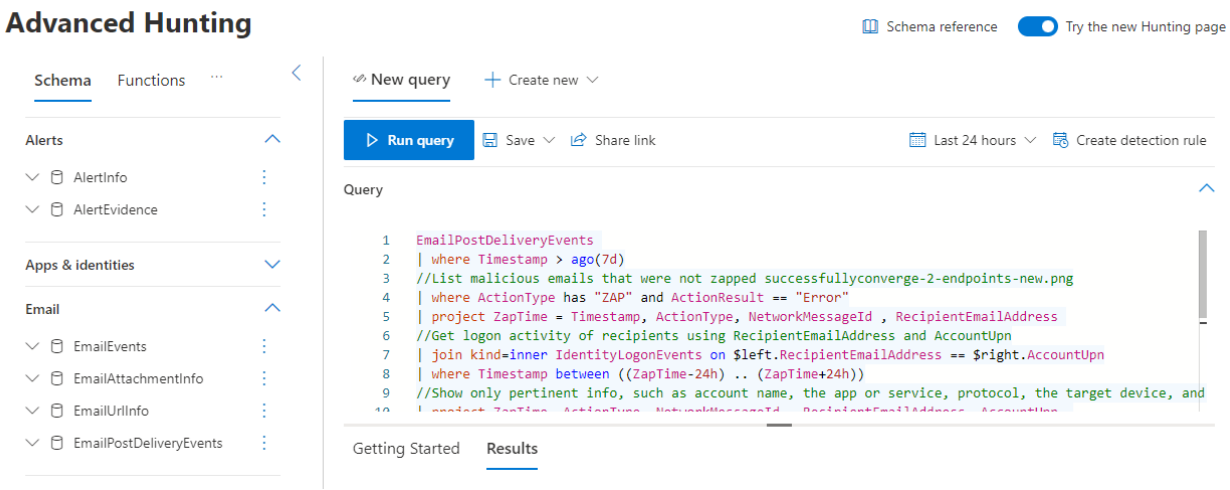


Рисунок 1.5 – Інтерфейс Advanced Hunting з введеним прикладом запиту

Інтеграція цих масивів даних дозволяє досліднику не просто констатувати факт атаки, а й відновити повний контекст інциденту, побудувати часову шкалу подій та визначити слабкі місця в налаштуваннях політик безпеки. Завершуючи огляд архітектури, необхідно підкреслити важливість інтеграції сигналів від Defender for Endpoint та Defender for Identity в межах єдиної платформи XDR. Це дає можливість корелювати подію отримання фішингового листа з подальшою

підозрілою активністю на робочій станції, наприклад, запуском невідомого процесу або аномаліями в аутентифікації. Така крос-доменна видимість є ключовою перевагою досліджуваного рішення порівняно з традиційними системами, оскільки вона забезпечує контекстуальне збагачення даних без необхідності складного налаштування правил нормалізації.

1.3 Аналіз існуючих підходів та метрик оцінки ефективності систем протидії фішингу

Традиційні підходи до оцінки роботи SOC (Security Operations Center) часто базуються на базових метриках, таких як кількість оброблених інцидентів, кількість заблокованих атак або середній час реагування. Проте в контексті протидії сучасному фішингу ці показники часто виявляються недостатніми або навіть такими, що вводять в оману. Наприклад, висока кількість заблокованих листів може свідчити не про ефективність захисту, а про масовану спам-атаку низької складності, тоді як один пропущений цільовий фішинговий лист може призвести до компрометації всієї інфраструктури.

Існуючі методи оцінювання часто ігнорують динамічну природу загроз. Класичні метрики, такі як Mean Time to Detect (MTTD) та Mean Time to Respond (MTTR), зазвичай розраховуються від моменту генерації алерту системою безпеки. Однак такий підхід не враховує час, протягом якого загроза перебувала в інфраструктурі до моменту її виявлення. У випадку з фішингом критично важливим є інтервал між доставкою листа та першою взаємодією користувача з ним. Якщо система виявляє та видаляє лист через 5 хвилин після доставки, але користувач встиг відкрити його через 2 хвилини, формально система спрацювала, але фактично захист не було забезпечено. Тому виникає необхідність у розробці більш гранулярних метрик, які б враховували часову шкалу подій з точністю до секунд і корелювали дії системи з поведінкою користувачів.

Ще однією проблемою існуючих підходів є те, що важко оцінити, скільки атак система не виявила – тобто випадків, коли атака відбулася, але механізм захисту не спрацював. У класичній теорії виявлення сигналів точність і повнота

розраховуються на основі відомої вибірки. У реальних умовах функціонування SOC ми часто не знаємо точної кількості пропущених загроз, доки не станеться інцидент. Стандартні звіти поштових шлюзів показують лише те, що вони заблокували, створюючи ілюзію повної видимості. Для отримання об'єктивної оцінки необхідно використовувати методи ретроспективного пошуку та кореляційний аналіз, який дозволяє виявити пропущені загрози за непрямими ознаками, такими як створення підозрілих правил у поштовій скриньці або аномальні входи в систему відразу після отримання листа [11].

Сучасні дослідження в галузі метрик кібербезпеки пропонують використовувати інтегральні показники, які поєднують технічну ефективність з впливом на бізнес. Зокрема, розглядається метрика «вартості ігнорування», яка оцінює потенційні збитки від пропущених атак. Однак практична реалізація таких метрик ускладнена відсутністю інструментарію для автоматизованого збору необхідних даних. Більшість доступних на ринку рішень надають статичні дашборди, які важко адаптувати під специфічні потреби організації. Використання мови запитів KQL у середовищі Microsoft Defender XDR відкриває нові можливості для створення кастомізованих метрик у реальному часі. Це дозволяє перейти від пасивного споглядання статистики до проактивного управління ризиками, базуючись на даних про реальну ефективність політик безпеки.

Важливим аспектом оцінки є також рівень «шуму» або хибнопозитивних спрацювань. Надмірно агресивні політики фільтрації можуть призводити до блокування легітимної ділової кореспонденції, що негативно впливає на бізнес-процеси та призводить до так званої «втоми від безпеки» у користувачів та адміністраторів. Оптимальна система захисту повинна балансувати між мінімізацією ризиків та забезпеченням доступності комунікацій. Для вимірювання цього балансу доцільно використовувати F-міру (F_1-score), яка є гармонічним середнім між точністю та повнотою, адаптувавши її розрахунок до специфіки лог-файлів поштових систем [12].

Отже, проведений аналіз свідчить про те, що існуючі засоби звітності не надають повної картини ефективності захисту від складних фішингових атак.

Виникає об’єктивна потреба у розробці методів, яка б дозволяла автоматизувати розрахунок розширених метрик ефективності, враховуючи як успішні блокування, так і виявлені постфактум пропуски, а також часові затримки реакції. Використання інструментарію Advanced Hunting та мови KQL є перспективним напрямком для вирішення цієї науково-прикладної задачі.

РОЗДІЛ 2 ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ЗАСАДИ ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ ПРОТИДІЇ ФІШИНГУ

2.1 Метрики оцінювання якості детектування загроз

Для об’єктивного вимірювання ефективності засобів захисту в середовищі Microsoft Defender XDR необхідно формалізувати процес обробки електронної кореспонденції, представивши його як задачу бінарної класифікації, де система приймає рішення щодо належності кожного вхідного повідомлення до одного з двох класів: легітимного листа або фішингової загрози. Теоретичним підґрунтям для такої формалізації виступає теорія бінарної класифікації в машинному навчанні, яка дозволяє математично описати здатність системи розрізняти два класи об’єктів (легітимні листи та фішингові атаки) в умовах невизначеності [13]. Застосування цього підходу дозволяє перейти від суб’єктивних оцінок надійності до кількісних метрик (precision, recall, F1-score, accuracy), що базуються на аналізі результатів класифікації та матриці помилок.

У контексті функціонування поштового шлюзу та XDR-системи простір подій можна описати за допомогою матриці помилок, яка визначає чотири можливі результати класифікації, як це зображено на рисунку 2.1.

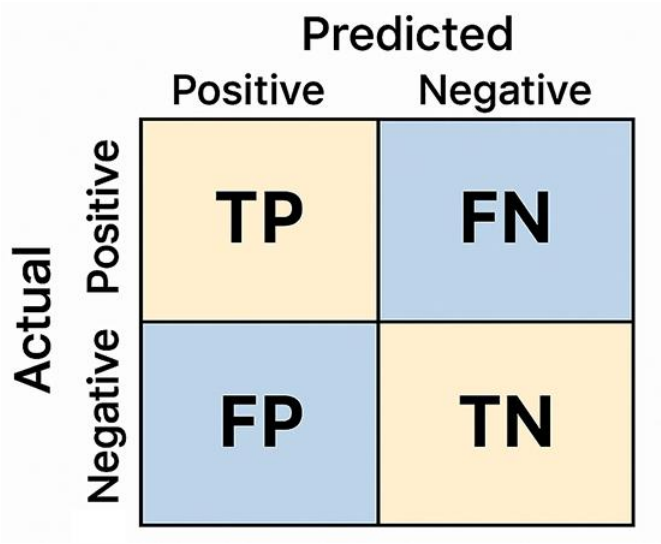


Рисунок 2.1 – Матриця помилок

Перший стан – істинно-позитивне рішення (True Positive, TP), виникає, коли система коректно ідентифікує фішинговий лист та блокує його. У логах Advanced Hunting цьому відповідають події з таблиці EmailEvents, де поле ThreatTypes містить значення Phish, а DeliveryAction має статус Block або Quarantine.

Другий стан – істинно-негативне рішення (True Negative, TN), описує ситуацію, коли легітимний лист успішно доставляється користувачу. Це наймасовіша категорія подій, яка характеризує нормальний режим роботи бізнес-процесів.

Третій стан – хибно-позитивне рішення (False Positive, FP), або помилка першого роду. Ця ситуація виникає, коли система помилково ідентифікує легітимне повідомлення як шкідливе (наприклад, класифікує ділову переписку як фішинг або спам) і застосовує до нього обмежувальні заходи, такі як карантин або переміщення у папку «Небажана пошта».

Технічно процес виникнення FP у середовищі Microsoft Defender XDR виглядає як послідовність зміни станів. Спочатку система детектує лист за сигнатурами, евристикою або моделями машинного навчання і присвоює йому вердикт, що фіксується в логах як успішне спрацювання захисту. Однак згодом відбувається процес рекласифікації. Це може трапитися внаслідок дії адміністратора (Admin Submission), який звільняє лист з карантину, або звіту користувача (User Reported as Safe), який позначає лист як помилково заблокований. Після надсилання такого звіту Microsoft проводить повторний аналіз (rescan) і, у разі підтвердження помилки, змінює вердикт на Clean. Саме наявність у журналі аудиту запису про те, що первинний вердикт «Шкідливо» було змінено на «Безпечно», дозволяє ретроспективно кваліфікувати цю подію як False Positive.

Четвертий стан – хибно-негативне рішення (False Negative, FN), або помилка другого роду, є найбільш критичним для кібербезпеки. Це ситуація, коли фішинговий лист проходить через фільтри захисту та потрапляє до скриньки користувача.

Важливо зауважити фундаментальну часову асиметрію процесу класифікації в автоматизованих системах захисту. У момент первинної обробки

кореспонденції (час t_0) для алгоритмів Microsoft Defender XDR не існує понять «помилка першого роду» або «помилка другого роду». Система апріорі вважає всі свої рішення вірними. Тобто, з «точки зору» машини, у момент t_0 існують лише два стани: True Positive (загрозу виявлено і заблоковано) та True Negative (лист визнано чистим і пропущено).

Стани помилок (False Positive та False Negative) є латентними (прихованими) у момент надходження листа. Вони проявляються лише ретроспективно, у момент часу t_1 ($t_1 > t_0$), коли зовнішній відносно алгоритму арбітр (адміністратор безпеки, користувач або інший механізм верифікації) спростовує первинний вердикт системи. Таким чином, аналіз ефективності є по суті пошуком розбіжностей між первинним автоматичним рішенням та остаточним статусом листа, який було скориговано в процесі життєвого циклу інциденту.

Зазвичай оцінка якості класифікаційних моделей може проводитись через метрики точності, коли ми рахуємо відношення правильно класифікованих листів (легітимних та фішингових) до загальної кількості листів. Математично цей показник розраховується за формулою (2.1).

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \tag{2.1}$$

Іншою характеристикою є влучність або достовірність (Precision), яка визначає частку дійсно шкідливих листів серед усіх повідомлень, заблокованих системою. Математично цей показник розраховується за формулою (2.2).

$$Precision = \frac{TP}{TP + FP} \tag{2.2}$$

Проте для нашої задачі важливою є саме мінімізація ймовірності виникнення подій FN, що і є основною метою налаштування політик безпеки.

Специфіка застосування даного методу до аналізу лог-файлів Microsoft Defender XDR полягає в тому, що на момент доставки листа (час t_0) стан False Negative (коли система визнала лист легітимним, але насправді він є фішинговим)

є прихованим. Система вважає, що прийняла правильне рішення (True Negative), і лише згодом, у момент часу t_1, коли спрацьовує механізм Zero-hour Auto Purge (ZAP) або надходить звіт від користувача, статус події ретроспективно змінюється на False Negative. Таким чином, множину помилок другого роду можна представити як суму подій, виявлених автоматично після доставки (FN_{ZAP}), та подій, про які повідомили користувачі (FN_{User}). Математично цей показник розраховується за формулою (2.3).

$$FN=FN_{ZAP}+FN_{User} \tag{2.3}$$

Така декомпозиція дозволяє точніше оцінити ефективність різних компонентів системи захисту.

Високе значення точності свідчить про низький рівень помилкових спрацювань, що зменшує навантаження на адміністраторів SOC. Другою критично важливою метрикою є повнота (Recall), або чутливість системи, яка показує здатність системи виявити всі наявні загрози. Повнота визначається як відношення кількості заблокованих загроз до загальної кількості реальних атак, включаючи ті, що були пропущені, згідно з формулою (2.4).

$$Recall=\frac{TP}{TP+FN} \tag{2.4}$$

У контексті даного дослідження знаменник формули (2.4) є динамічною величиною, оскільки кількість пропущених атак (FN) може зростати з часом по мірі виявлення нових інцидентів або отримання даних про компрометацію. Для комплексної оцінки якості роботи системи, яка балансує між необхідністю блокувати загрози та забезпеченням безперервності бізнесу, використовується F-міра (F_1-score). Цей показник є гармонічним середнім між точністю та повнотою і розраховується за формулою (2.5).

$$F_1=2*\frac{Precision*Recall}{Precision+Recall} \tag{2.5}$$

Використання F_1-score дозволяє порівнювати ефективність різних конфігурацій політик безпеки за єдиним скалярним значенням. Крім того, для оцінки ризику доцільно ввести показник коефіцієнта пропуску (Miss Rate або False Negative Rate – FNR), який відображає ймовірність того, що реальна атака не буде заблокована первинними механізмами детектування та досягне поштової скриньки користувача. З урахуванням того, що множина FN у даній роботі формується ретроспективно (за рахунок спрацювань ZAP та/або підтверджень користувачів відповідно до (2.3)), значення FNR слід інтерпретувати як частку пропущених загроз серед усіх наявних загроз за обраний часовий інтервал спостереження. Математично цей показник визначається як доповнення до повноти (Recall) та розраховується за формулою (2.6). У практичній частині роботи наведений показник буде імплементовано мовою KQL для автоматичного обчислення на основі телеметрії, що дозволить відстежувати тренди зміни ефективності захисту в динаміці [14].

$$FNR=\frac{FN}{TP+FN} \tag{2.6}$$

Важливо зазначити, що запропонований метод враховує часовий фактор. Оскільки вердикт щодо шкідливості листа може змінитися після оновлення баз загроз, поняття "істинності" класифікації є функцією від часу. Тому розрахунок метрик повинен проводитися у фіксованих часових вікнах (наприклад, 7 або 30 днів), що дозволяє стабілізувати значення FN за рахунок завершення циклу пост-аналізу та реакції ZAP. Такий підхід забезпечує наукову строгість та відтворюваність результатів оцінювання ефективності системи протидії фішингу.

2.2 Методи розрахунку інтегральних показників ефективності

Для отримання репрезентативної оцінки ефективності протидії фішингу недостатньо обмежуватися лише метриками якості класифікації з підрозділу 2.1, оскільки вони описують коректність вердиктів, але не відображають тривалість “вікна експозиції”, протягом якого лист уже перебуває в скриньці й користувач

потенційно може взаємодіяти зі шкідливим контентом. У середовищі Microsoft Defender XDR значна частина критично важливих подій проявляється постфактум: лист може бути доставлений із первинним вердиктом Clean, а згодом вилучений механізмом ZAP або вручну в рамках реагування. Саме тому методологічним ядром цього підрозділу є кореляційний підхід, за якого подія доставки з таблиці EmailEvents розглядається як стартова точка життєвого циклу повідомлення, а пост-доставочні дії з таблиці EmailPostDeliveryEvents – як спостережувані завершальні стани, що дозволяють вимірювати оперативність нейтралізації та навантаження на механізми реагування.

Формально кожен лист у вибірці доцільно описувати як сутність, що має унікальний ідентифікатор повідомлення (наприклад, NetworkMessageId або InternetMessageId) та часову мітку доставки $T_{delivery}$. Якщо після доставки над листом виконано дію нейтралізації, їй відповідає час $T_{remediation}$, отриманий із пост-доставочних подій. У практичних запитах важливо фіксувати саме перший момент нейтралізації, оскільки одна й та сама сутність може породжувати кілька пост-доставочних записів (наприклад, повторні переміщення або додаткові сервісні дії). Тому в процедурі кореляції слід застосовувати правило мінімального часу $T_{remediation}$ для кожного ідентифікатора, щоб уникнути штучного “роздування” вибірки та викривлення агрегованих статистик.

Окремим показником, який доповнює FNR і дозволяє оцінити проникнення загроз у робочий контур комунікацій, є коефіцієнт просочування в абсолютному вираженні щодо всього вхідного потоку. На відміну від FNR, що нормується відносно множини атак, коефіцієнт просочування доцільно трактувати як частку листів, які були доставлені користувачам і потребували пост-доставочного вилучення через підтвердження загрози. Нехай N_{in} – загальна кількість вхідних листів у часовому вікні спостереження, а N_{leak} – кількість листів, для яких одночасно виконується факт доставки та факт подальшої нейтралізації (ZAP або ручне реагування) за тим самим ідентифікатором. Тоді коефіцієнт просочування визначається формулою (2.7).

$$LeakageRate = \frac{N_{leak}}{N_{in}} \quad (2.7)$$

Практична цінність LeakageRate полягає в тому, що він нормує пост-доставочні інциденти відносно всього потоку листів і тим самим дозволяє порівнювати періоди з різним навантаженням на поштову інфраструктуру, а також відстежувати тренди при зміні політик фільтрації чи налаштувань антифішингового профілю.

Окремим виміром ефективності, релевантним до ризиків користувацької взаємодії, є результативність блокування переходів за посиланнями. Навіть якщо лист було доставлено, механізми на кшталт Safe Links здатні зупинити шкідливу дію в момент кліку, що суттєво знижує ймовірність компрометації. Для цього вводиться коефіцієнт блокування кліків, який відображає частку переходів, що були заблоковані засобами захисту, відносно всіх зафіксованих спроб переходу за аналізований період. Цей показник визначається формулою (2.8) та обчислюється на основі подій переходів користувачів із відповідною ознакою блокування.

$$ClickBlockRate = \frac{N_{blocked}}{N_{total}} \quad (2.8)$$

У комплексі LeakageRate та ClickBlockRate описують дві взаємодоповнювальні площини ризику – частоту проникнення шкідливих повідомлень у скриньки користувачів і здатність системи зупиняти атаку в момент взаємодії з її елементами. У практичній частині ці показники доцільно розраховувати у прив’язці до однакового часового інтервалу та з однаковими правилами фільтрації, щоб забезпечити відтворюваність результатів і коректність порівнянь між вибірками.

РОЗДІЛ 3 ПРАКТИЧНА РЕАЛІЗАЦІЯ МЕТОДУ ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ ЗАХОДІВ ПРОТИДІЇ ФІШИНГУ

3.1 Розробка пошукових запитів для збору телеметрії

Практична реалізація запропонованої методу оцінювання ефективності системи захисту від фішингу базується на використанні інструментарію Advanced Hunting у середовищі Microsoft Defender XDR. Першочерговим завданням є формування масиву даних, придатного для подальшого статистичного аналізу та розрахунку метрик, визначених у другому розділі. Оскільки сирі дані в журналах подій мають гетерогенну структуру, де одне поле може містити кілька значень або нестандартизовані формати, необхідно розробити спеціалізовані KQL-запити для їх нормалізації [16]. Основним джерелом даних для аналізу вхідного потоку загроз виступає таблиця EmailEvents, яка містить детальну інформацію про кожен оброблений електронний лист, включаючи вердикти фільтрації, політики та дії системи.

Для забезпечення репрезентативності вибірки необхідно визначити часовий діапазон аналізу. У даному дослідженні обрано період у 30 днів, що є стандартом для ретроспективного аналізу в Advanced Hunting і дозволяє охопити повний життєвий цикл більшості фішингових кампаній, включаючи відкладені детектування. Критично важливим етапом попередньої обробки даних є нормалізація поля ThreatTypes. У вихідній схемі даних це поле може містити комбінацію значень (наприклад, «Phish, Spam»), що ускладнює пряму агрегацію. Для вирішення цієї проблеми розроблено алгоритм пріоритезації загроз, реалізований через оператор case мови KQL. Логіка нормалізації полягає у присвоєнні кожному листу єдиного статусу за пріоритетом небезпеки: шкідливе програмне забезпечення (Malware) має найвищий пріоритет, за ним слідує фішинг (Phishing), потім спам (Spam), і, нарешті, чисті листи.

Розроблений запит, наведений у лістингу 3.1 виконує фільтрацію вхідного трафіку (EmailDirection == "Inbound"), нормалізацію типів загроз та агрегацію даних за діями системи доставки (DeliveryAction). Це дозволяє отримати базову

статистику ефективності первинної фільтрації та побудувати візуальну модель розподілу загроз.

Лістинг 3.1 – KQL запит для нормалізації та аналізу структури вхідного потоку загроз

```
EmailEvents
| where Timestamp > ago(30d)
| where EmailDirection == "Inbound"
| extend ThreatCategory = case(
    ThreatTypes has "Malware", "Malware",
    ThreatTypes has "Phish", "Phishing",
    ThreatTypes has "Spam", "Spam",
    "Clean")
| summarize Count = count() by ThreatCategory, DeliveryAction
| order by Count desc
| render piechart
```

Результат виконання наведеного запиту дозволяє візуалізувати співвідношення заблокованих та доставлених листів для кожної категорії загроз. Наприклад, наявність подій зі статусом ThreatCategory == "Phishing" та DeliveryAction == "Delivered" є первинним індикатором потенційних інцидентів, які потребують додаткового аналізу через механізми пост-доставки.

Результати виконання KQL запиту, наведеного у лістингу 3.1, на основі реальних даних представлені на рисунку 3.1. З нього видно, що найбільшою категорією по кількості листів є не зловмисні листи, які були доставлені отримувачу, а другою категорією є спам, який в більшості випадків був віднесений до категорії небажаних листів.

Отримані дані дають загальну картину того, які листи система блокує, а які все ж доходять до користувача. Якщо серед доставлених є повідомлення з категорією Phishing, це означає, що такі випадки потрібно перевіряти окремо, щоб зрозуміти причину пропуску та оцінити ризики.

Run query | Set in query | Save | Share link

Query

```
1  EmailEvents
2  | where Timestamp > ago(30d)
3  | where EmailDirection == "Inbound"
4  | extend ThreatCategory = case(
5      ThreatTypes has "Malware", "Malware",
6      ThreatTypes has "Phish", "Phishing",
7      ThreatTypes has "Spam", "Spam",
8      "Clean")
9  | summarize Count = count() by ThreatCategory, DeliveryAction
10 | order by Count desc
11 | render piechart
```

Getting started | Results | Query history

Export | Show empty columns

Filters: Add filter

ThreatCategory	DeliveryAction	Count
Clean	Delivered	42821
Spam	Junked	4079
Clean	Junked	1381
Phishing	Blocked	922
Spam	Blocked	295
Phishing	Junked	196
Clean	Blocked	194
Spam	Delivered	151
Phishing	Delivered	118
Malware	Blocked	3

Рисунок 3.1 – Структура вхідного потоку листів

На рисунку 3.2 представлена графічна візуалізація даних, отриманих у результаті виконання KQL запиту, наведеного у лістингу 3.1.

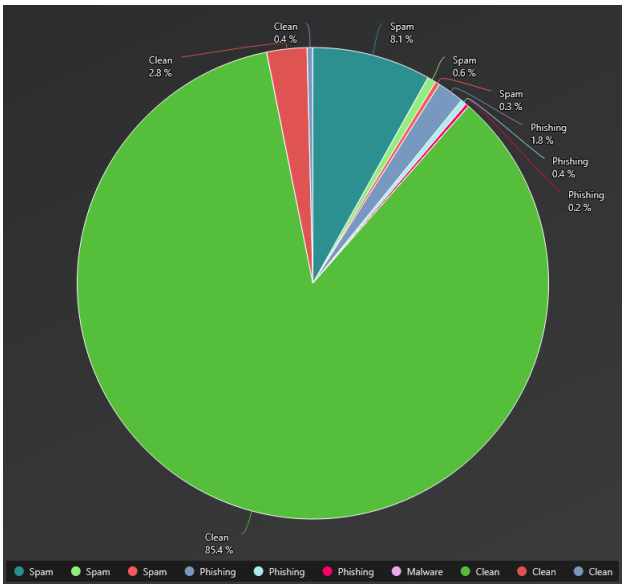


Рисунок 3.2 – Графічна візуалізація структури вхідного потоку листів

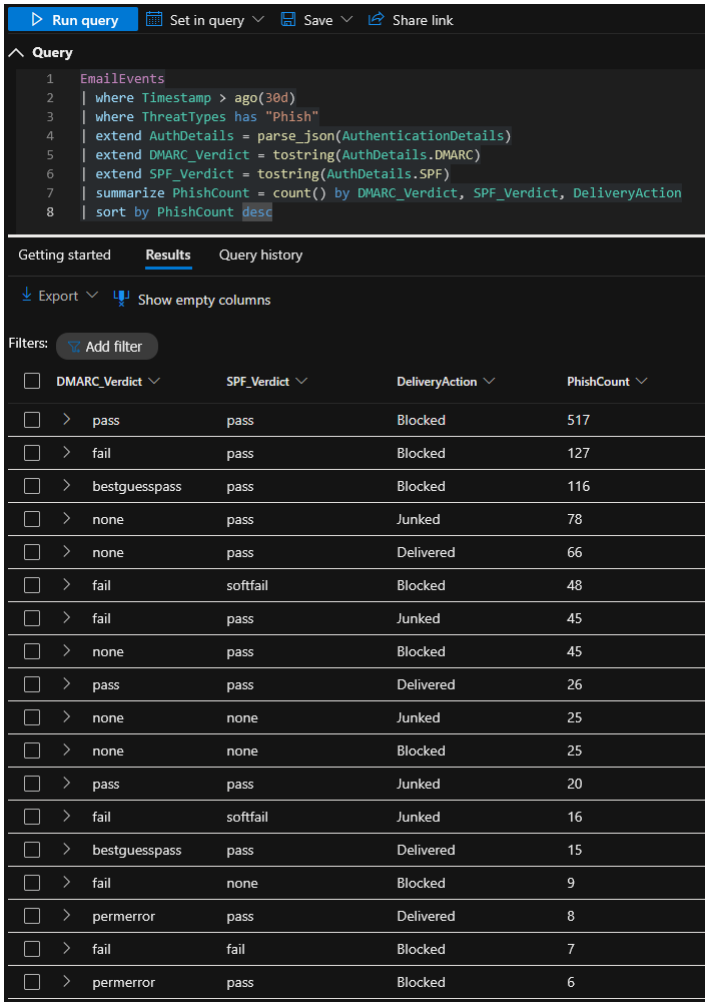
Наступним кроком є збагачення даних про фішингові листи інформацією про вектори атаки. Для цього необхідно використати дані з таблиць EmailUrlInfo та EmailAttachmentInfo. Оскільки один лист може містити безліч посилань або вкладень, проста операція з'єднання (join) може призвести до дублювання записів та викривлення статистики. Для уникнення цього у розробленому алгоритмі використовується агрегація вкладених сутностей у списки перед з'єднанням з основною таблицею подій. Це дозволяє сформувати єдиний аналітичний запис для кожного інциденту, що містить метадані листа, перелік підозрілих URL-адрес та хеш-суми вкладень.

Особливу увагу при нормалізації слід приділити полю AuthenticationDetails, яке містить результати перевірки SPF, DKIM та DMARC. У сирому вигляді це поле є текстовим рядком формату JSON або XML. Для автоматизованої оцінки ефективності політик аутентифікації необхідно розпарсити цей рядок та виділити вердикти для кожного протоколу в окремі колонки. Реалізація цього підходу представлена у Лістингу 3.2, де використовується функція parse_json для структурування даних аутентифікації, що дозволяє виявити кореляцію між налаштуваннями DMARC та ймовірністю пропуску спуфінг-атак [17].

Лістинг 3.2 – KQL запит для аналізу кореляції між результатами аутентифікації та вердиктами фільтрації

```
EmailEvents
| where Timestamp > ago(30d)
| where ThreatTypes has "Phish"
| extend AuthDetails = parse_json(AuthenticationDetails)
| extend DMARC_Verdict = tostring(AuthDetails.DMARC)
| extend SPF_Verdict = tostring(AuthDetails.SPF)
| summarize PhishCount = count() by DMARC_Verdict, SPF_Verdict,
DeliveryAction
| sort by PhishCount desc
```

Результати виконання KQL запиту, наведеного у лістингу 3.2, зображені на рисунку 3.3.



Run query | Set in query | Save | Share link

Query

```
EmailEvents
| where Timestamp > ago(30d)
| where ThreatTypes has "Phish"
| extend AuthDetails = parse_json(AuthenticationDetails)
| extend DMARC_Verdict = tostring(AuthDetails.DMARC)
| extend SPF_Verdict = tostring(AuthDetails.SPF)
| summarize PhishCount = count() by DMARC_Verdict, SPF_Verdict, DeliveryAction
| sort by PhishCount desc
```

Getting started | Results | Query history

Export | Show empty columns

Filters: Add filter

DMARC_Verdict	SPF_Verdict	DeliveryAction	PhishCount
pass	pass	Blocked	517
fail	pass	Blocked	127
bestguesspass	pass	Blocked	116
none	pass	Junked	78
none	pass	Delivered	66
fail	softfail	Blocked	48
fail	pass	Junked	45
none	pass	Blocked	45
pass	pass	Delivered	26
none	none	Junked	25
none	none	Blocked	25
pass	pass	Junked	20
fail	softfail	Junked	16
bestguesspass	pass	Delivered	15
fail	none	Blocked	9
permerror	pass	Delivered	8
fail	fail	Blocked	7
permerror	pass	Blocked	6

Рисунок 3.3 – Структура фішингових листів

Отримані в результаті виконання запитів нормалізовані набори даних формують фундамент для розрахунку складних метрик, таких як коефіцієнт хибно-негативних спрацювань (FNR) та час реакції (MTTD). Застосування мови KQL на етапі підготовки даних дозволяє автоматизувати рутинні операції обробки логів та забезпечити високу точність вимірювань, виключаючи людський фактор при підрахунку статистики інцидентів.

Загалом, підготовка й нормалізація даних через KQL потрібні для того, щоб далі коректно рахувати метрики та порівнювати результати. Коли дані приведені до одного формату, з ними простіше працювати, робити підрахунки та будувати графіки без помилок і зайвих дублювань.

3.2 Реалізація алгоритму виявлення хибно-негативних спрацювань та розрахунку часових затримок

На основі підготовлених у попередньому підрозділі нормалізованих даних телеметрії середовища Microsoft Defender XDR було реалізовано алгоритми для автоматизованого розрахунку ключових показників ефективності фільтрації. Зокрема, мова KQL дозволяє безпосередньо обчислити метрики точності (Precision), повноти (Recall), коефіцієнта хибнонегативних спрацювань (False Negative Rate, FNR) та інтегральної F_1-міри на основі статистики класифікації кожного електронного листа. Крім того, розроблено запити для оцінки затримки виявлення фішингових повідомлень (часу реакції, MTTD) та для підрахунку кількості інцидентів, пропущених первинним захистом і виявлених лише ретроспективно або вручну.

Для визначення кількісних показників правильних та помилкових рішень здійснюється підрахунок подій кожної категорії результату класифікації за вибраний період. Використовуючи розроблену раніше категоризацію ThreatCategory, KQL-запит агрегує кількість листів у кожному з чотирьох можливих станів: TP (True Positive, заблокована загроза), FP (False Positive, помилково заблокований легітимний лист), FN (False Negative, пропущена загроза) та TN (True Negative, легітимний лист, що успішно пройшов фільтри). На основі цих значень обчислюються Precision, Recall, F_1 та FNR згідно з визначеннями, наведеними в другому розділі.

У лістингу 3.3 показано відповідний KQL-запит, який реалізує описаний алгоритм розрахунку. В результаті виконання цього KQL запиту, буде формуватися таблиця з одним рядком, що міститиме значення всіх розрахованих показників.

Отриманий результат зручний для подальшого аналізу, оскільки дозволяє одразу оцінити якість фільтрації за вибраний період. Також цей підхід спрощує порівняння показників між різними проміжками часу або після змін у політиках захисту, бо структура розрахунку залишається однаковою.

Лістинг 3.3 – KQL-запит для розрахунку матриці класифікації та похідних метрик Precision, Recall, F_1, FNR

```
EmailEvents
| where Timestamp > ago(30d)
| where EmailDirection == "Inbound"
| extend ThreatCategory = case(
    ThreatTypes has "Malware", "Malware",
    ThreatTypes has "Phish", "Phishing",
    ThreatTypes has "Spam", "Spam",
    "Clean")
| summarize
    TP = countif(ThreatCategory in ("Phishing", "Malware") and
DeliveryAction in ("Blocked", "Junked")),
    FN_hidden = countif(ThreatCategory in ("Phishing", "Malware")
and DeliveryAction == "Delivered"),
    FP = countif(ThreatCategory == "Clean" and DeliveryAction in
("Blocked", "Junked")),
    TN = countif(ThreatCategory == "Clean" and DeliveryAction ==
"Delivered")
| extend Precision = round(todouble(TP) / todouble(TP + FP), 4),
         Recall = round(todouble(TP) / todouble(TP + FN_hidden), 4),
         F1_score = round(2 * todouble(TP) / todouble(2 * TP + FP +
FN_hidden), 4),
         FNR = round(todouble(FN_hidden) / todouble(TP + FN_hidden),
4)
```

Результат виконання KQL запиту, наведеного у лістингу 3.3, зображено на рисунку 3.4. Так, на реальних даних отримано Precision ≈ 0.42 (42%), Recall ≈ 0.90 (90%), F1 ≈ 0.57 (57%) і FNR ≈ 0.1 (10%).

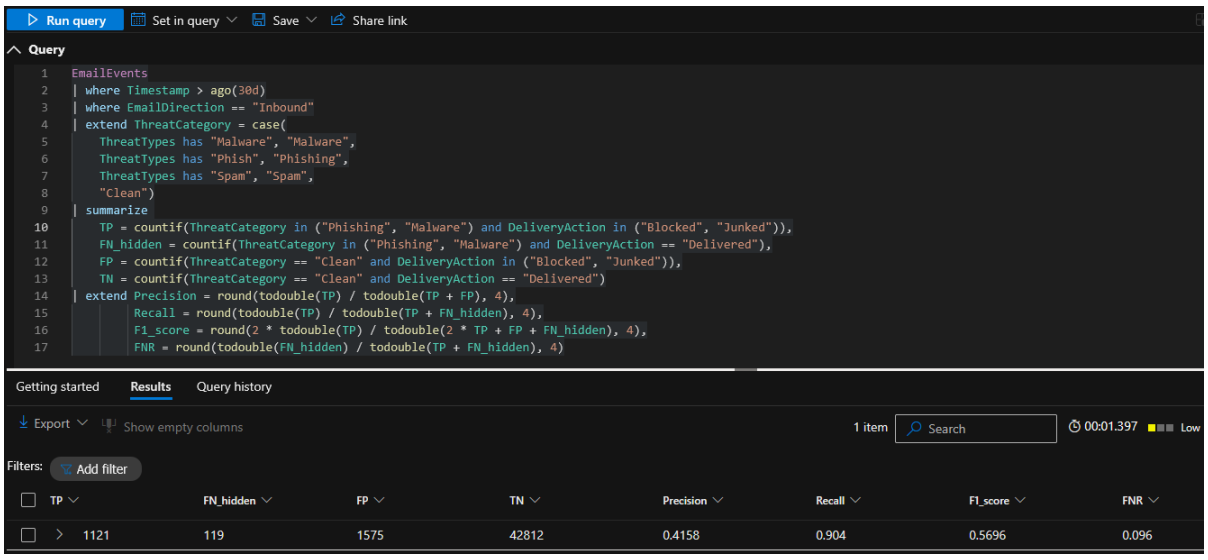


Рисунок 3.4 – Матриця класифікації та похідних метрик

Наступним етапом є аналіз часових характеристик реагування системи на загрози. Для оцінки затримки виявлення фішингових повідомлень було використано об’єднання таблиці початкових подій (EmailEvents) з таблицею пост-доставочних дій (EmailPostDeliveryEvents). Це дозволяє зіставити час доставки кожного листа з часом його подальшого видалення із поштової скриньки засобами ZAP або вручну адміністратором. У лістингу 3.4 наведено KQL-запит, що обчислює середній та медіанний час між доставкою фішингового листа і його нейтралізацією.

Лістинг 3.4 – KQL-запит для оцінки затримки виявлення фішингових листів (MTTD)

```
EmailEvents
| where Timestamp > ago(30d)
| where EmailDirection == "Inbound" and DeliveryAction == "Delivered"
| project NetworkMessageId, DeliveryTime = Timestamp, RecipientEmailAddress, Subject, ThreatTypes
| join kind=inner (
    EmailPostDeliveryEvents
    | where Timestamp > ago(30d)
    | extend RemovalTime = Timestamp
    | project NetworkMessageId, RemovalTime, ActionType
) on NetworkMessageId
| where ThreatTypes has "Phish" or ActionType in ("Phish ZAP", "Manual remediation")
| extend DetectionDelay = datetime_diff("hour", RemovalTime, DeliveryTime)
| summarize AvgDelayHours = avg(DetectionDelay), MedianDelayHours = percentile(DetectionDelay, 50)
```

У цьому запиті за допомогою операції join по полю NetworkMessageId встановлюється зв’язок між фактом доставки листа (з EmailEvents) та фактом його видалення після доставки (з EmailPostDeliveryEvents). Фільтр ActionType in ("Phish ZAP", "Manual Remediation") гарантує, що аналізуються лише ті випадки, коли лист був класифікований як фішинговий постфактум – автоматично видалений системою або вручну переміщений адміністратором до карантину. Обчислення різниці часу між доставкою та видаленням (datetime_diff) дозволяє отримати величину затримки в годинах для кожного інциденту. Далі агрегуються середнє значення цього інтервалу (AvgDelayHours) та медіана

(MedianDelayHours) для всіх загроз, які були виявлені із запізненням. Наприклад, у разі ефективної роботи механізму ZAP середня затримка виявлення може становити ~1.5 години, а медіанний час – близько 1 години, що вказує на те, що зазвичай пропущені фішингові атаки нейтралізуються протягом перших годин після доставки. Результати виконання запиту, наведеного у лістингу 3.4, зображено на рисунку 3.5.

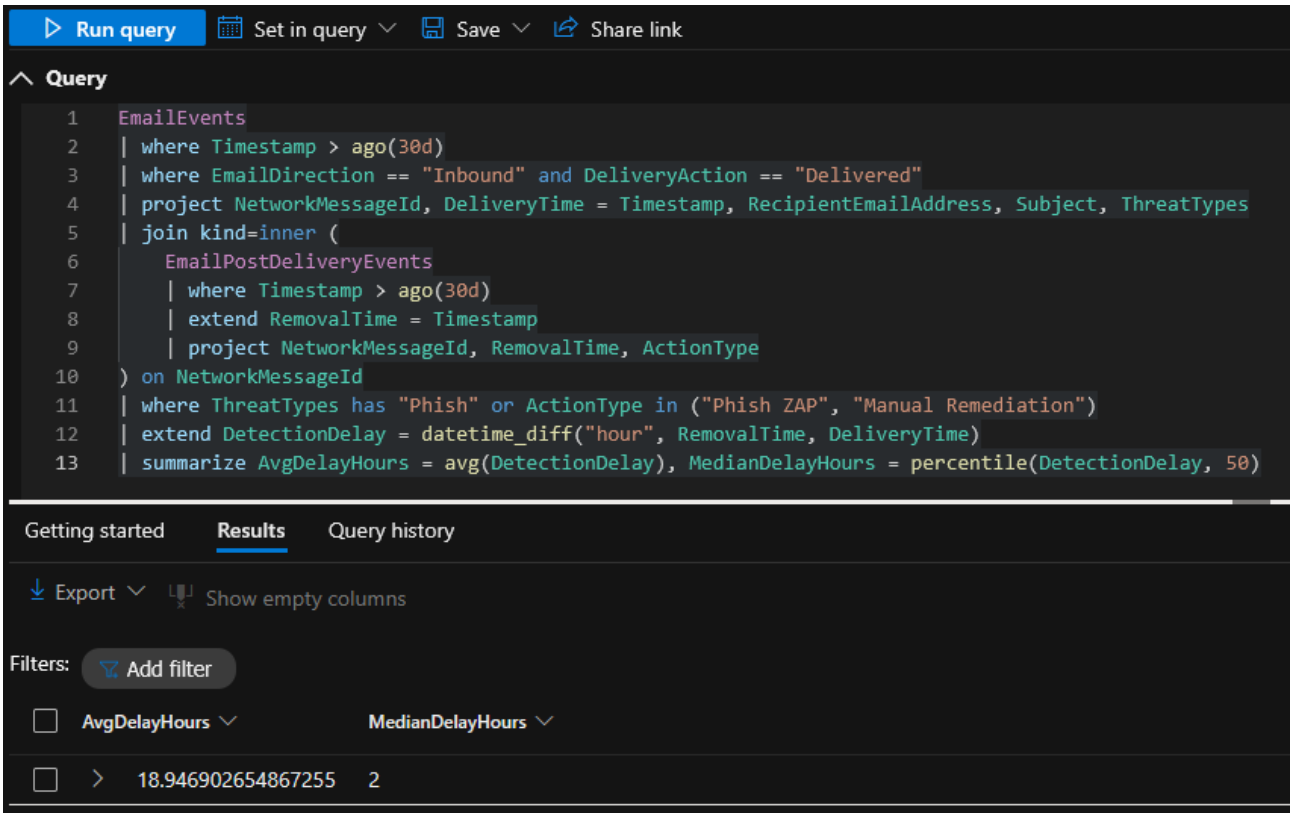


Рисунок 3.5 – Середня та медіанна затримки між доставкою та видаленням

Окрім кількісних метричних показників, доцільно оцінити спосіб виявлення пропущених атак, оскільки це характеризує роль автоматизованих засобів та людини в системі захисту. Для цього окремим запитом підраховано кількість інцидентів типу FN, які були ліквідовані механізмом Zero-hour Auto Purge (Phish ZAP), та кількість тих, що залишилися непоміченими системою і потребували ручного реагування з боку адміністратора. Лістинг 3.5 містить відповідний KQL-запит.

Лістинг 3.5 – KQL-запит для підрахунку ретроспективно автоматично виявлених і вручну виявлених інцидентів

```
EmailPostDeliveryEvents
| where Timestamp > ago(30d)
| where ActionType in ("Phish ZAP", "Manual Remediation")
| summarize
    FN_ZAP = countif(ActionType contains "ZAP"),
    FN_User = countif(ActionType == "Manual Remediation")
```

Цей запит звертається до таблиці подій після доставки та рахує кількість записів за кожним типом дії. Поле FN_ZAP відображає число фішингових листів, автоматично видалених системою після початкової доставки, а FN_User – кількість інцидентів, що були пропущені фільтрацією і виявлені лише після повідомлення користувачів (через ручне втручання адміністратора). Результати запиту дають змогу оцінити, яка частка пропущених атак була нейтралізована засобами захисту без участі людини, а які інциденти залишалися невиявленими до звернень від користувачів.

Результати виконання KQL запиту, наведеного у лістингу 3.5, зображені на рисунку 3.6.

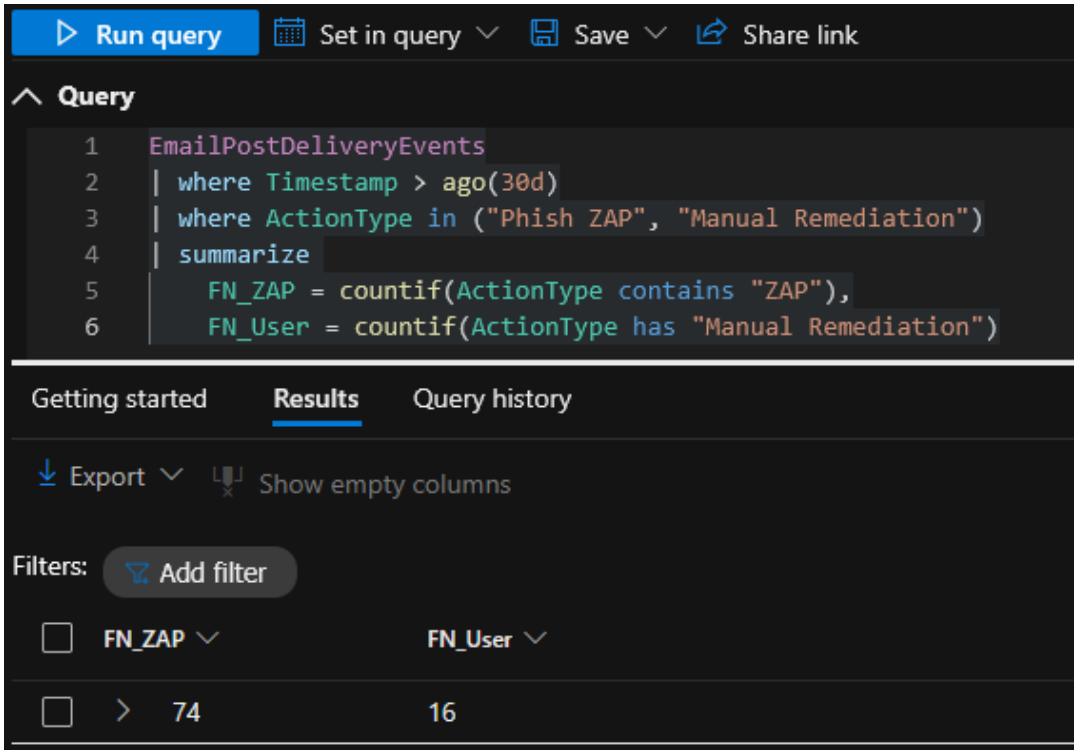


Рисунок 3.6 – Підрахунок автоматичних і ручних виявлень фішингових листів

3.3 Інтерпретація результатів експериментального оцінювання та рекомендації щодо оптимізації політик безпеки

Отримані в реальній інфраструктурі компанії результати розрахунку метрик ефективності протидії фішингу в Microsoft Defender XDR засобами Advanced Hunting дозволяють перейти від загального уявлення про «наявність захисту» до предметної оцінки того, як саме працюють антифішингові механізми в повсякденних умовах, у яких присутні і масові розсилки, і цільові атаки, і легітимні бізнес-листи, що можуть бути схожими на фішингові за окремими ознаками. Важливо підкреслити, що наведені показники не є «оцінкою продукту Microsoft» як такого, а відображають ефективність конкретної реалізації захисту в конкретному середовищі, тобто результат поєднання технічних можливостей платформи, обраних політик, інтеграцій, а також реального профілю загроз і поведінки користувачів. Саме тому інтерпретація метрик у цьому підрозділі розглядається як основа для прийняття практичних рішень щодо оптимізації політик безпеки та процесів реагування.

За результатами виконання розроблених KQL-запитів отримано значення Precision ≈ 0.42 (42%), Recall ≈ 0.90 (90%), F1 ≈ 0.57 (57%) і FNR ≈ 0.10 (10%). Таке співвідношення показників є характерним для середовищ, у яких організація робить акцент на максимальному перехопленні підозрілих повідомлень, але стикається з наслідками «агресивної» фільтрації у вигляді значної частки хибнопозитивних рішень. Високе значення Recall означає, що більшість фішингових листів так чи інакше ідентифікуються, а отже первинний контур захисту загалом виконує свою базову функцію й не дозволяє значній кількості атак безперешкодно досягати користувачів. Водночас низький Precision демонструє, що суттєва частка заблокованих або поміщених у карантин повідомлень фактично є легітимною кореспонденцією. У практичній площині це проявляється не лише додатковими запитами користувачів щодо «зниклих листів», а й системним операційним навантаженням на SOC або адміністраторів, які змушені постійно виконувати перевірку, відновлення або внесення винятків. За таких умов зростає ризик, що важливі сигнали загроз будуть «загублені» на

фоні великої кількості подій із низькою реальною небезпекою, а також з’являється спокуса послабити політики безпеки ситуативно, щоб відновити нормальний робочий процес. Таким чином, низький Precision є не лише технічною характеристикою, а й фактором, що прямо впливає на стійкість організації до атак через вплив на дисципліну реагування та якість операційних процесів.

Окремої уваги заслуговує FNR ≈ 0.10, що вказує на наявність пропусків первинного детектування. Навіть якщо частина таких випадків згодом нейтралізується, сам факт потрапляння фішингового повідомлення до поштової скриньки формує «вікно ризику», протягом якого користувач може виконати небезпечну дію. У сучасних сценаріях атак це особливо актуально, оскільки зловмисники прагнуть отримати миттєву взаємодію, використовуючи терміновість, авторитет або підробку знайомих сервісів. У цьому контексті FNR не слід трактувати як суто статистичний показник, оскільки його практичний зміст полягає в імовірності того, що окремі атаки стартуватимуть із фази взаємодії користувача, а отже подальше реагування може бути вже «після факту». Висновок із цього полягає в тому, що оптимізація політик має переслідувати одразу дві цілі – зменшення кількості помилкових блокувань та скорочення частки пропусків, але робити це потрібно системно, опираючись на спостережувані дані та повторювані вимірювання.

Часові характеристики реагування, отримані за даними розрахунку затримки виявлення, становлять AvgDelayHours = 19 та MedianDelayHours = 2. Таке поєднання медіани й середнього значення зазвичай означає нерівномірний розподіл затримок: у типових ситуаціях пропущені листи видаляються або нейтралізуються відносно швидко, але існує помітна частина інцидентів, які залишаються невиявленими значно довше та формують «довгий хвіст». З практичної точки зору саме ці випадки є найбільш критичними, оскільки збільшують час, протягом якого фішинговий лист перебуває в середовищі, а отже підвищують шанси успішної компрометації. Такі затримки можуть мати різну природу: індикатори загрози могли з’явитися лише через певний час, механізми ретроспективної переоцінки спрацювали із запізненням, лист містив

нестандартний вміст, або ж організаційні процеси реагування були ініційовані пізно. У будь-якому разі результат демонструє важливість метрик часу як окремого виміру ефективності, адже навіть «правильне» виявлення, яке відбулося із запізненням, не завжди знижує ризик до прийнятного рівня.

Додатково встановлено, що серед пропущених інцидентів після доставки суттєва частина нейтралізувалася автоматично завдяки механізмам ретроспективного реагування: FN_ZAP = 74 та FN_User = 16. Домінування автоматичних дій над ручним втручанням підтверджує практичну значущість механізмів на кшталт Zero-hour Auto Purge, які компенсують частину помилок або пропусків на етапі доставки та дозволяють зменшувати загальний ризик без прямої участі людини. Водночас наявність випадків, що потребували ручного реагування, свідчить, що не всі сценарії фішингу однаково добре «підхоплюються» автоматикою, а також що роль користувацького репортингу та аналітичних процедур SOC залишається суттєвою. На практиці це означає, що технічні політики захисту мають бути доповнені процесами виявлення та ескалації інцидентів, а також автоматизацією рутинних кроків аналізу й реагування, щоб мінімізувати час від першого сигналу до нейтралізації.

З огляду на виявлені характеристики, оптимізацію політик безпеки доцільно спрямувати на зменшення частки хибнопозитивних рішень без втрати досягнутої повноти виявлення, а також на скорочення часу нейтралізації пропущених інцидентів. У частині Precision ключовим є перехід від «широких» узагальнених правил до більш контекстно-орієнтованих рішень, які враховують довірені бізнес-процеси та профіль комунікацій організації. Практично це передбачає системний перегляд причин помилкових блокувань, виявлення повторюваних шаблонів легітимної кореспонденції, що потрапляє під карантин, та кероване застосування винятків із контролем ризику. Важливо, щоб винятки не були «глобальними» і неконтрольованими, а базувалися на перевірених ознаках, наприклад на автентифікації доменів і стабільності комунікації з конкретними контрагентами. Доцільно також посилити дисципліну керування дозволеними відправниками та доменами, підтримувати актуальність списків довірених

партнерів і використовувати принцип мінімально необхідних винятків, щоб не створювати «дірки» у захисті.

Додатковим напрямом підвищення точності є перегляд політик антиспаму й антифішингу з фокусом на узгодження рівнів агресивності для різних категорій користувачів і підрозділів. У реальних організаціях існують ролі з різним ризик-профілем: бухгалтерія, фінансові служби, керівництво, підрозділи закупівель зазвичай є пріоритетними цілями BEC-атак і потребують більш суворого контролю, тоді як для інших груп надмірна агресивність може створювати непропорційні бізнес-втрати. Тому сегментація політик за групами користувачів із різними сценаріями комунікації може дати кращий загальний баланс між безпекою та безперервністю бізнес-процесів, одночасно зменшуючи рівень «зайвих» карантинів.

У частині зниження FNR та зменшення часових затримок доцільно посилити механізми превентивної перевірки вмісту, зокрема URL-адрес і вкладень, а також забезпечити достатню швидкість ретроспективного переоцінювання повідомлень. З практичного погляду це означає коректну конфігурацію і повноцінне використання функцій аналізу посилань та вкладень у середовищі Microsoft Defender, а також контроль того, як швидко спрацьовують механізми пост-доставочного реагування для різних типів загроз. Важливим елементом є скорочення «довгого хвоста» затримок, який формує високе середнє значення, що, як правило, досягається комбінацією технічних і процесних змін. До технічних змін належить уточнення правил детектування, посилення валідації відправника, фокус на аномаліях заголовків, підвищення пріоритету реагування на окремі класи індикаторів. До процесних змін належить налаштування чітких SLA на обробку репортів користувачів, впровадження автоматизованих playbook’ів для типових сценаріїв та забезпечення швидкого поширення індикаторів на всю організацію після підтвердження інциденту.

Окремим напрямом оптимізації є підвищення якості каналів зворотного зв’язку від користувачів. З огляду на наявність випадків ручного реагування, варто забезпечити максимально простий механізм повідомлення про підозрілі листи, стандартизувати процедуру опрацювання таких повідомлень і

використовувати їх як джерело даних для подальшого налаштування політик. У практичному сенсі це сприяє скороченню часу виявлення в тих сценаріях, які автоматичні механізми не класифікують достатньо швидко, а також формує культуру кібергігієни, що є критичною складовою протидії фішингу як соціально-технічній загрозі.

Сформовані результати мають як діагностичний, так і прикладний характер. Вони демонструють, що запропонований підхід до оцінювання ефективності засобів протидії фішингу на основі Advanced Hunting та бібліотеки KQL-запитів є працездатним у реальному середовищі та забезпечує вимірюваність ключових характеристик захисту, включно з тими аспектами, які складно оцінити на підставі лише стандартних зведень. Власне, цінність підходу полягає в можливості організувати керований цикл вдосконалення, коли зміни конфігурацій, правил детектування або операційних процедур оцінюються на підставі повторюваних метрик до та після впровадження, а не на основі інтуїції чи одиничних кейсів. У результаті організація отримує практичний інструмент управління якістю антифішингового захисту, який дозволяє збалансувати безпеку й безперервність комунікацій, скоротити витрати на обробку «шуму», зменшити критичні затримки реагування та підвищити загальну кіберстійкість без необхідності розгортання додаткових систем збору телеметрії.

РОЗДІЛ 4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1 Охорона праці

Виконання кваліфікаційної роботи магістра, присвяченої оцінці ефективності заходів протидії фішингу в середовищі Microsoft Defender XDR, передбачає проведення комплексу науково-дослідних робіт, які за своїм характером належать до сфери інформаційних технологій та аналітичної діяльності. Специфіка роботи дослідника в даному контексті полягає у необхідності тривалої взаємодії з візуальними дисплейними терміналами електронно-обчислювальних машин, що використовується для моніторингу інцидентів безпеки, написання скриптів для Advanced Hunting та аналізу великих масивів даних. Цей трудовий процес характеризується значним інтелектуальним навантаженням, необхідністю концентрації уваги, напруженням зорового аналізатора, а також гіпокінезією, викликаною тривалим перебуванням у вимушеній робочій позі сидячи. З огляду на це, забезпечення безпечних умов праці є невід’ємною складовою ефективного виконання дослідження, оскільки комфортне робоче середовище безпосередньо впливає на продуктивність праці та якість отриманих результатів.

Проведений аналіз умов праці дозволив ідентифікувати потенційно шкідливі та небезпечні виробничі фактори, що можуть впливати на дослідника під час роботи з програмно-апаратним комплексом. До основних фізичних факторів належать підвищений рівень електромагнітних випромінювань від системного блоку та периферійного обладнання, можлива недостатня освітленість робочої зони, наявність відблисків на екрані монітора, підвищений рівень статичної електрики, а також небезпека ураження електричним струмом у разі пошкодження ізоляції струмопровідних частин. Психофізіологічні фактори ризику обумовлені специфікою кібербезпекової діяльності та включають нервово-емоційне напруження при обробці інформації, перевтому зорового апарату внаслідок тривалого спостереження за об’єктами розрізнення на екрані,

а також монотонність праці. Для мінімізації впливу зазначених факторів розроблено комплекс інженерно-технічних та організаційних заходів, що відповідають сучасним нормативним вимогам України.

Організація робочого місця дослідника здійснюється відповідно до чинних «Вимог до безпеки та захисту здоров'я працівників під час роботи з екранними пристроями», затверджених наказом Мінсоцполітики України № 207 від 14.02.2018 [18]. Робоче місце спроектовано таким чином, щоб забезпечити достатній простір для розміщення засобів праці, зокрема моніторів, клавіатури та документації, дозволяючи змінювати робочу позу та здійснювати необхідні рухи. Робочий стіл має низьку відбивну здатність, а його розміри дозволяють гнучко розташовувати екран, клавіатуру та документи. Клавіатура розміщується окремо від екрана, що дозволяє обрати зручну робочу позу та уникнути втоми рук. Простір перед клавіатурою достатній для опори рук та зап'ясть, а поверхня клавіатури матова, щоб уникнути відблисків. Робоче крісло є стійким, дозволяє користувачеві легко рухатися та займати зручне положення, а сидіння регулюється по висоті. Екран монітора встановлено на такій відстані від очей, що відповідає вимогам ергономіки (600–700 мм), а зображення на ньому є стабільним, без мерехтінь, що критично важливо при тривалому аналізі логів.

Суттєвий вплив на працездатність та самопочуття дослідника мають параметри мікроклімату в приміщенні, до яких відносяться температура, відносна вологість та швидкість руху повітря. Робота з налаштування політик безпеки та аналізу логів відноситься до категорії 1а згідно з ДСанПіН 3.3-6.042-99 «Санітарні норми мікроклімату виробничих приміщень», як така, що пов'язана з незначним фізичним напруженням. Для забезпечення теплового комфорту в холодний період року температура повітря підтримується в межах 22-24 градусів Цельсія, а в теплий період – 23-25 градусів Цельсія. Відносна вологість повітря контролюється на рівні 40-60 відсотків, що запобігає пересиханню слизових оболонок дихальних шляхів та очей, а швидкість руху повітря обмежується величиною 0,1 метра на секунду для уникнення протягів. Забезпечення цих параметрів досягається за допомогою системи припливно-

витяжної вентиляції та побутових кондиціонерів, а також шляхом регулярного провітрювання приміщення під час регламентованих перерв.

Для створення комфортних умов зорової роботи спроектовано систему освітлення, яка відповідає вимогам ДБН В.2.5-28:2018 «Природне і штучне освітлення» [19]. У приміщенні застосовується система суміщеного освітлення, де природне світло через віконні прорізи доповнюється штучним. Штучне освітлення реалізоване як загальне рівномірне, що забезпечується стельовими світильниками з розсіювачами світла. Нормативний рівень освітленості на поверхні робочого столу в зоні роботи з документами становить від 300 до 500 люкс. Особлива увага приділена запобіганню прямого блиску від джерел світла та відбитого блиску від екранів моніторів, для чого вікна обладнані регульованими жалюзі, а розташування світильників виключає їх попадання в поле зору працівника. Використання сучасних джерел світла дозволяє уникнути пульсації світлового потоку, яка може викликати стробоскопічний ефект та підвищену втому очей.

Зважаючи на те, що виконання роботи пов’язане з експлуатацією електроустановок напругою 220 В, важливим етапом є забезпечення електробезпеки. Приміщення лабораторії або офісу класифікується як приміщення без підвищеної небезпеки ураження електричним струмом. Відповідно до «Правил безпечної експлуатації електроустановок споживачів» [20], усі металеві частини обладнання, які нормально не перебувають під напругою, але можуть опинитися під нею внаслідок пошкодження ізоляції, підлягають захисному заземленню або зануленню. Це реалізується через використання трипровідної мережі живлення та розеток із заземлюючим контактом. Для захисту від струмів витоку в електричному щиті встановлено пристрої захисного вимкнення (ПЗВ), які забезпечують миттєве відключення живлення у аварійних ситуаціях. Дослідник зобов’язаний перед початком роботи візуально перевіряти цілісність корпусів, вилок, з’єднувальних шнурів та розеток, а також не допускати потрапляння вологи на електрообладнання.

Окремим важливим питанням є пожежна безпека, оскільки наявність великої кількості електронної техніки, кабельних комунікацій та паперових

носіїв інформації створює пожежне навантаження. Організація пожежної безпеки здійснюється згідно з Кодексом цивільного захисту України та НАПБ А.01.001-2014 «Правила пожежної безпеки в Україні» [21]. Для запобігання виникненню пожежі суворо забороняється перевантажувати електромережу підключенням великої кількості споживачів до однієї розетки, залишати обладнання увімкненим без нагляду після завершення робочого дня та використовувати пошкоджені електроприлади. Приміщення обладнане автоматичними пожежними сповіщувачами, що реагують на появу диму. В якості первинних засобів пожежогасіння обрано вуглекислотні вогнегасники типу ВВК, оскільки вони ефективно гасять загоряння електроустановок під напругою і не завдають шкоди комп'ютерній техніці, випаровуючись без сліду. Розміщення вогнегасників виконано у помітних та легкодоступних місцях поблизу виходів з приміщення.

Таким чином, дотримання обґрунтованих вимог охорони праці є критично важливим фактором для успішного виконання кваліфікаційної роботи. Забезпечення ергономічних умов та належного освітлення безпосередньо впливає на здатність дослідника ефективно працювати з інтерфейсом Microsoft Defender XDR, мінімізуючи зорову втому та підвищуючи концентрацію уваги при аналізі складних інцидентів безпеки. Відсутність відволікаючих факторів та фізичного дискомфорту дозволяє уникнути помилок при написанні пошукових запитів KQL під час Advanced Hunting, що гарантує точність виявлення фішингових атак та достовірність отриманих результатів дослідження їх ефективності.

4.2 Безпека в надзвичайних ситуаціях

Забезпечення безпеки життєдіяльності під час виконання науково-дослідної роботи передбачає не лише створення комфортних умов праці, але й готовність до ефективних дій у разі виникнення надзвичайних ситуацій техногенного, природного або воєнного характеру. Правовою основою для розробки заходів безпеки є Кодекс цивільного захисту України, який визначає механізми захисту

населення, територій та майна від надзвичайних ситуацій. Оскільки робота над темою оцінки ефективності заходів протидії фішингу виконується в умовах офісного приміщення або спеціалізованої лабораторії, найбільш ймовірними загрозами є пожежа, аварії в системах життєзабезпечення (електропостачання), а також загрози, пов’язані з воєнним станом, зокрема повітряні атаки [22]. Метою цього підрозділу є розробка чіткого алгоритму дій дослідника, спрямованого на збереження життя та мінімізацію матеріальних втрат у критичних умовах [23].

Планування заходів цивільного захисту на об’єкті, де проводиться дослідження, базується на принципах превентивності та оперативності реагування. Ключовим елементом системи безпеки є наявність та знання плану евакуації, який графічно відображає шляхи виходу з приміщення, місця розташування первинних засобів пожежогасіння, ручних пожежних сповіщувачів та аптечок домедичної допомоги. Дослідник, перебуваючи на робочому місці, повинен заздалегідь ознайомитися з маршрутами евакуації, щоб уникнути паніки та дезорієнтації в умовах задимлення або відключення освітлення. Важливим аспектом є також знання розташування найближчих захисних споруд цивільного захисту (укриттів або сховищ), оскільки в сучасних умовах загроза ракетних ударів вимагає миттєвої реакції на сигнали оповіщення.

Однією з найбільш поширених небезпек в адміністративних будівлях є виникнення пожежі. У разі виявлення ознак горіння, таких як полум’я, дим або запах горілої ізоляції, першочерговою дією є негайне повідомлення про подію пожежно-рятувальної служби за телефоном «101». При цьому необхідно чітко назвати адресу об’єкта, місце виникнення пожежі та надати інформацію про наявність людей у будівлі. Паралельно з викликом рятувальників слід активувати систему ручного пожежного оповіщення, щоб попередити інших співробітників про небезпеку. Евакуація повинна відбуватися організовано, без паніки, використовуючи визначені евакуаційні виходи та сходові клітки. Використання ліфтів під час пожежі суворо заборонено, оскільки існує високий ризик відключення електроенергії та блокування кабіни у шахті, яка може наповнитися токсичними продуктами горіння.

Якщо осередок пожежі незначний і його гасіння не загрожує життю дослідника, допускається використання первинних засобів пожежогасіння. Враховуючи наявність дороговартісного комп'ютерного обладнання та серверів Microsoft Defender XDR, пріоритетним є використання вуглекислотних вогнегасників. Вони ефективно збивають полум'я та охолоджують поверхню, не пошкоджуючи електронні компоненти, на відміну від порошкових або пінних засобів. Під час гасіння електроустановок під напругою необхідно дотримуватися безпечної відстані не менше одного метра від сопла вогнегасника до струмопровідних частин. У випадку неможливості самостійного гасіння слід негайно залишити приміщення, щільно зачинивши за собою двері, щоб обмежити доступ кисню до осередку горіння та сповільнити розповсюдження вогню.

В умовах воєнного стану критично важливим є алгоритм дій під час оголошення сигналу «Увага всім!» (повітряна тривога). Почувши звук сирени або отримавши сповіщення через офіційні канали комунікації, дослідник зобов'язаний негайно припинити роботу. Специфіка роботи з даними кіберрозвідки та Advanced Hunting вимагає, за можливості, виконання екстреного збереження даних та блокування робочої станції для запобігання несанкціонованому доступу до конфіденційної інформації під час відсутності персоналу. Після цього необхідно вимкнути основне електроживлення комп'ютера та периферійних пристроїв, взяти особисті речі, документи та мобільний телефон і найкоротшим шляхом дістатися до визначеного укриття. Перебування в захисній споруді триває до моменту отримання офіційного повідомлення про відбій тривоги. Ігнорування сигналів оповіщення є неприпустимим порушенням правил безпеки.

У разі виникнення надзвичайних ситуацій, що супроводжуються травмуванням людей, дослідник повинен володіти навичками надання домедичної допомоги. При ураженні електричним струмом, що є специфічним ризиком при роботі з ЕОМ, першим кроком є звільнення потерпілого від дії струму шляхом вимкнення рубильника або відкидання проводу сухим діелектричним предметом. Після цього необхідно оцінити стан потерпілого,

перевірити наявність свідомості та дихання. У разі відсутності ознак життя негайно розпочинаються серцево-легенева реанімація, яка проводиться до відновлення дихання або прибуття бригади екстреної медичної допомоги. При термічних опіках уражену ділянку слід охолодити холодною водою, накласти стерильну пов’язку та не допускати пошкодження пухирів. У випадку отруєння чадним газом потерпілого необхідно винести на свіже повітря, розстібнути одяг, що утруднює дихання, та забезпечити спокій.

Окрему увагу в системі безпеки слід приділити інформаційній складовій. Оскільки тема кваліфікаційної роботи стосується кібербезпеки, надзвичайна ситуація може бути використана зловмисниками як фактор дестабілізації для проведення кібератак. Тому план дій у надзвичайних ситуаціях включає процедури резервного копіювання критично важливих даних на віддалені хмарні сховища або захищені носії. Це забезпечує безперервність дослідницького процесу та збереження результатів аналізу навіть у випадку фізичного знищення обладнання внаслідок пожежі чи інших руйнівних факторів.

Отже, організація безпеки в надзвичайних ситуаціях є комплексним завданням, що вимагає від дослідника знання нормативних документів, чіткого розуміння алгоритмів евакуації та вміння надавати домедичну допомогу. Дотримання розроблених інструкцій та планів реагування дозволяє мінімізувати ризики для життя та здоров’я, а також забезпечити збереження наукових напрацювань та матеріальних цінностей в умовах непередбачуваних обставин. Свідоме ставлення до питань цивільного захисту є показником професійної компетентності та відповідальності фахівця.

ВИСНОВКИ

У кваліфікаційній роботі досліджено підхід до оцінювання ефективності протидії фішингу в Microsoft Defender XDR із використанням Advanced Hunting та запитів KQL. Запропоновано і апробовано метод, що дозволяє на основі телеметрії поштових подій формувати кількісні показники якості виявлення, а також оцінювати часові характеристики реагування та роль ретроспективних механізмів нейтралізації загроз.

Практична апробація на даних реальної інфраструктури підтвердила, що розроблений набір запитів і підхід до інтерпретації метрик є придатними для використання в операційній роботі, оскільки надають вимірювану основу для аналізу результативності налаштувань захисту, контролю змін політик і визначення пріоритетів їх оптимізації. Отримані результати демонструють характерні для реальних середовищ компроміси між рівнем перехоплення загроз і кількістю помилкових рішень, а також підкреслюють важливість скорочення часу між доставкою підозрілих повідомлень та їх нейтралізацією.

Загалом, застосування Advanced Hunting як інструмента метричного контролю забезпечує можливість системно підвищувати ефективність антифішингового захисту, підтримувати безперервне вдосконалення політик і підсилювати керованість процесів реагування без залучення додаткових засобів збору даних.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Proofpoint. (2024). 2024 State of the Phish – Today's cyber threats and phishing protection [Threat report]. https://www.proofpoint.com/us/resources/threat-reports/state-of-phish. Дата доступу: 03.12.2025
2. Verizon. (2024). 2024 Data Breach Investigations Report (DBIR). Verizon Business. https://www.verizon.com/business/resources/reports/dbir.html. Дата доступу: 03.12.2025
3. Hoxhunt. Phishing trends report (Updated for 2025). https://hoxhunt.com/guide/phishing-trends-report. Дата доступу: 03.12.2025
4. Federal Bureau of Investigation. (2025, April 23). FBI releases annual Internet crime report [Press release]. https://www.fbi.gov/news/press-releases/fbi-releases-annual-internet-crime-report. Дата доступу: 03.12.2025
5. Zscaler ThreatLabz. (2024, April 23). Phishing attacks rise 58% year-over-year as AI takes root in the threat landscape: ThreatLabz 2024 phishing report. https://www.zscaler.com/blogs/security-research/phishing-attacks-rise-58-year-ai-threatlabz-2024-phishing-report. Дата доступу: 03.12.2025
6. Panjawani, A., & Brandt, A. (2024, October 16). Quishing: The new trend in QR code attacks. Sophos News. https://news.sophos.com/en-us/2024/10/16/quishing/. Дата доступу: 03.12.2025
7. Microsoft. (2025, August 7). Phishing triage agent in Microsoft Defender XDR (Preview). Microsoft Learn. https://learn.microsoft.com/en-us/defender-xdr/phishing-triage-agent. Дата доступу: 03.12.2025
8. Microsoft. (2025, July 28). Anti-phishing protection in Microsoft Defender for Office 365. Microsoft Learn. https://learn.microsoft.com/en-us/defender-office-365/anti-phishing-protection-about. Дата доступу: 03.12.2025
9. Microsoft. (2025, June 3). Zero-hour auto purge (ZAP) in Microsoft Defender for Office 365. Microsoft Learn. https://learn.microsoft.com/en-us/defender-office-365/zero-hour-auto-purge. Дата доступу: 03.12.2025

10. Microsoft. (2025, July 23). URLClickEvents table in advanced hunting. Microsoft Learn. https://learn.microsoft.com/en-us/defender-xdr/advanced-hunting-urlclickevents-table. Дата доступу: 03.12.2025
11. Microsoft Threat Intelligence. (2025, September 24). AI vs AI: Detecting an AI-obfuscated phishing campaign. Microsoft Security Blog. https://www.microsoft.com/en-us/security/blog/2025/09/24/ai-vs-ai-detecting-an-ai-obfuscated-phishing-campaign/. Дата доступу: 03.12.2025
12. Cysec4psych. (2023, July 30). Signal detection theory. CySec4Psych. https://cysec4psych.eu/psych-cyber-concept/signal-detection-theory/. Дата доступу: 03.12.2025
13. McClelland, G. H. (2011). Use of signal detection theory as a tool for enhancing performance and evaluating tradecraft in intelligence analysis. In National Research Council, Intelligence analysis: Behavioral and social scientific foundations. National Academies Press. https://www.nationalacademies.org/read/13062/chapter/7. Дата доступу: 03.12.2025
14. Unchit, P., Das, S., Kim, A., & Camp, L. J. (2020). Quantifying susceptibility to spear phishing in a high school environment using signal detection theory (arXiv:2006.16380v2) [Preprint]. arXiv. https://arxiv.org/abs/2006.16380. Дата доступу: 03.12.2025
15. Microsoft. Advanced hunting query language. Microsoft Learn. https://learn.microsoft.com/en-us/defender-xdr/advanced-hunting-query-language. Дата доступу: 03.12.2025
16. Microsoft. Advanced hunting best practices. Microsoft Learn. https://learn.microsoft.com/en-us/defender-xdr/advanced-hunting-best-practices. Дата доступу: 03.12.2025
17. Microsoft. (2025, September 15). KQL quick reference. Microsoft Learn. https://learn.microsoft.com/en-us/kusto/query/kql-quick-reference?view=microsoft-fabric. Дата доступу: 03.12.2025
18. ZAGORODNA, N., STADNYK, M., LYPA, B., GAVRYLOV, M., & KOZAK, R. (2022). Network Attack Detection Using Machine Learning

Methods. Challenges to national defence in contemporary geopolitical situation, 2022(1), 55-61.

19. Boltov, Y., Skarga-Bandurova, I., & Derkach, M. (2023, September). A Comparative Analysis of Deep Learning-Based Object Detectors for Embedded Systems. In 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) (Vol. 1, pp. 1156-1160). IEEE.
20. Y. Skorenkyy, R. Zolotyy, S. Fedak, O. Kramar, R. Kozak. Digital Twin Implementation in Transition of Smart Manufacturing to Industry 5.0 Practices. CEUR Workshop Proceedings, 2023, 3468, pp. 12–23
21. Kozak R., Skorenkyy Yu., Kramar O., Lechachenko T. and Brevus H. Cybersecurity provisioning for Industry 4.0 digital twin with AR components // CEUR Workshop Proceedings, 3rd International Workshop on Computer Information Technologies in Industry 4.0 (CITI 2025), Ternopil, 2025.- vol. 4057, pp. 166–178.
22. Микитишин, А. Г., Митник, М. М., Голотенко, О. С., & Карташов, В. В. (2023). Комплексна безпека інформаційних мережевих систем. Навчальний посібник для студентів спеціальності 174 «Автоматизація, комп'ютерно-інтегровані технології та робототехніка».
23. Tymoshchuk D., Yasniy O., Mytnyk M., Zagorodna N., Tymoshchuk V. Detection and classification of DDoS flooding attacks by machine learning method. CEUR Workshop Proceedings. 2024. Vol. 3842. P. 184-195.
24. Стручок В.С. Методичний посібник для здобувачів освітнього ступеня «магістр» всіх спеціальностей денної та заочної (дистанційної) форм навчання «БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ» / Тернопіль: ФОП Паляниця В. А., –156 с.
25. Стручок В.С. Навчальний посібник «ТЕХНОЕКОЛОГІЯ ТА ЦИВІЛЬНА БЕЗПЕКА. ЧАСТИНА «ЦИВІЛЬНА БЕЗПЕКА»» / Тернопіль: ФОП Паляниця В. А., – 156 с.

ДОДАТОК А ПУБЛІКАЦІЯ

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ

М А Т Е Р І А Л И

XIII НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

«ІНФОРМАЦІЙНІ МОДЕЛІ, СИСТЕМИ ТА ТЕХНОЛОГІЇ»

17-18 грудня 2025 року

ТЕРНОПІЛЬ
2025

УДК 004.056.5
Т. Маслянка; Ю. Олійник; Н. Загородна, к.т.н., доцент
(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

ОБМЕЖЕННЯ СТАНДАРТНИХ МЕТРИК ЕФЕКТИВНОСТІ XDR-СИСТЕМ ПРИ АНАЛІЗІ ФІШИНГОВИХ КАМПАНІЙ

UDC 004.056.5
T. Maslianka; Yu. Oliinyk; N. Zagorodna, Ph.D, Assoc. Prof.

LIMITATIONS OF STANDARD EFFICIENCY METRICS OF XDR SYSTEM IN ANALYZING PHISHING CAMPAIGNS

У сучасному ландшафті кіберзагроз фішинг залишається домінуючим вектором первинного доступу зловмисників до корпоративних мереж. Згідно зі звітом Microsoft Digital Defense Report, лише за останній рік обсяг атак на основі паролів зріс у десять разів, а складність методів обходу фільтрації, таких як Adversary-in-the-Middle (AiTM), значно підвищилася [1]. У відповідь на це організації впроваджують платформи розширеного виявлення та реагування, зокрема Microsoft Defender XDR, які консолідують захист пошти, кінцевих точок та ідентичності. Однак, покладання виключно на автоматизовані засоби блокування та стандартні звіти створює хибне відчуття безпеки.

Проблематика стандартних інструментів звітності полягає у використанні обмеженого набору кількісних показників, які можуть не враховувати критичних факторів контексту та хронології інцидентів. Типові дашборди демонструють кількість заблокованих листів, але часто не дають відповіді на питання про час перебування фішингового листа у поштовій скриньці до моменту його видалення або про дії користувача, які не призвели до негайного спрацювання сигнатур. Як зазначають дослідники SANS Institute, перехід до проактивного захисту вимагає зміни парадигми від пасивного моніторингу до активного пошуку загроз [2].

Для об’єктивної оцінки ефективності контрзаходів необхідний перехід до аналізу «сирих» даних. У екосистемі Microsoft Defender XDR ключовим інструментом для цього є Advanced Hunting, що базується на мові запитів Kusto Query Language (KQL). На відміну від статичних звітів, KQL дозволяє корелювати події з різних доменів безпеки. Наприклад, поєднання даних з таблиць EmailEvents, EmailUrlInfo та DeviceNetworkEvents дозволяє виявити ланцюжки атак, де користувач перейшов за посиланням, яке на момент доставки вважалося безпечним [3].

Аналіз функціональних можливостей XDR-систем свідчить про необхідність застосування спеціалізованих метрик аудиту, які виходять за межі стандартних звітів. Використання певних запитів, таких як Advanced Hunting в Microsoft Defender XDR, дозволяє вирахувати такі критичні показники, як середній час між доставкою листа та кліком користувача та ефективність пост-детекції. Такий підхід дозволяє трансформувати механізми пошуку загроз у засіб верифікації надійності системи захисту та виявляти слабкі місця в політиках безпеки, які залишаються непоміченими при використанні стандартних інструментів моніторингу.

Література

1. Microsoft Digital Defense Report 2024 [Електронний ресурс]. – Режим доступу: https://www.microsoft.com/en-us/security/security-insider/microsoft-digital-defense-report-2024.
2. Lee R. M. The Who, What, Where, When, Why and How of Effective Threat Hunting / R. M. Lee, D. Wolpoff. – SANS Institute, 2016. – 18 p.
3. Advanced hunting schema [Електронний ресурс] // Microsoft Learn. – 2024. – Режим доступу: https://learn.microsoft.com/en-us/defender-xdr/advanced-hunting-schema-tables.