МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ КІБЕРБЕЗПЕКИ ТА ЗАХИСТУ ІНФОРМАЦІЇ
КАФЕДРА УПРАВЛІННЯ КІБЕРБЕЗПЕКОЮ ТА ЗАХИСТОМ ІНФОРМАЦІЇ

КВАЛІФІКАЦІЙНА РОБОТА

на тему: “МЕХАНІЗМИ УПРАВЛІННЯ КІБЕРБЕЗПЕКОЮ З УРАХУВАННЯМ ПСИХОЛОГІЧНИХ АСПЕКТІВ ПОВЕДІНКИ КОРИСТУВАЧІВ”

на здобуття освітнього ступеня магістра
зі спеціальності 125 Кібербезпека та захист інформації
освітньо-професійної програми Управління інформаційною та кібернетичною безпекою

Кваліфікаційна робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ Дмитро КОНДРАТЮК
(підпис) *Ім’я, ПРІЗВИЩЕ здобувача*

Виконав: Здобувач вищої освіти гр. УБДМ-61

Керівник:

Рецензент:

Київ 2025

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

Навчально-науковий інститут кібербезпеки та захисту інформації

Кафедра Управління кібербезпекою та захистом інформації
Ступінь вищої освіти магістр
Спеціальність 125 Кібербезпека та захист інформації
Освітньо-професійна програма Управління інформаційною та кібернетичною безпекою

ЗАТВЕРДЖУЮ
Завідувач кафедрою УКБЗІ
_____ Світлана ЛЕГОМІНОВА
"____" _____ 2025 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Студенту КОНДРАТЮКУ Дмитру Олеговичу
(прізвище, ім'я, по батькові здобувача)

1. Тема кваліфікаційної роботи: "Механізми управління кібербезпекою з урахуванням психологічних аспектів поведінки користувачів"

керівник кваліфікаційної роботи Іван ОПІРСЬКИЙ, д-р техн. наук, професор
(Ім'я, ПРІЗВИЩЕ, науковий ступінь, вчене звання)
затверджені наказом Державного університету інформаційно-комунікаційних технологій від "30" жовтня 2025 р. № 467.

2. Строк подання кваліфікаційної роботи "11" грудня 2025 р.
3. Вихідні дані до кваліфікаційної роботи:.
4. Перелік питань, які потрібно розробити:
   1. Аналіз теоретико-психологічних моделей поведінки користувачів та сучасного стану систем управління кібербезпекою.
   2. Розробка концептуальної моделі адаптивного управління на основі поведінкової аналітики та ключових індикаторів психологічного ризику.
   3. Обґрунтування та деталізація механізмів адаптивного реагування: динамічного управління привілеями та методів психологічного впливу.

5. Перелік ілюстративного матеріалу: *презентація*
6. Дата видачі завдання "02" жовтня 2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назви етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1.	Визначення об’єкту, предмету, мети та завдань дослідження.	10.10.2025	
2.	Збір та аналіз літератури.	23.10.2025	
3.	Аналіз основних характеристик інформаційного протиборства.	27.10.2025	
4.	Дослідження особливостей поведінкової аналітики користувачів та розробка системи ключових індикаторів психологічного ризику.	10.11.2025	
5.	Визначення напрямів та методів впровадження адаптивних механізмів управління у загальну систему кібербезпеки підприємства.	15.11.2025	
6.	Формулювання висновків за результатами дослідження.	22.11.2025	
7.	Оформлення роботи.	04.12.2025	
8.	Оформлення презентації.	14.12.2025	
9.	Отримання рецензії на роботу.	18.12.2025	
10.	Захист в ЕК.	20.01.2026	

Здобувач вищої освіти _____ Дмитро КОНДРАТЮК
(підпис) *(Ім’я, ПРІЗВИЩЕ)*

Керівник
кваліфікаційної роботи _____ Іван ОПІРСЬКИЙ
(підпис) *(Ім’я, ПРІЗВИЩЕ)*

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ КІБЕРБЕЗПЕКИ ТА ЗАХИСТУ ІНФОРМАЦІЇ

ПОДАННЯ
ГОЛОВІ ЕКЗАМЕНАЦІЙНОЇ КОМІСІЇ
ЩОДО ЗАХИСТУ КВАЛІФІКАЦІЙНОЇ РОБОТИ
на здобуття освітнього ступеня магістра

Направляється здобувач Кондратюк Д.О. до захисту кваліфікаційної роботи
(прізвище та ініціали)
за спеціальністю 125 Кібербезпека та захист інформації
(код, найменування спеціальності)
Освітньо-професійної програми Управління інформаційною та кібернетичною безпекою
(назва)
на тему: "Механізми управління кібербезпекою з урахуванням психологічних аспектів поведінки користувачів"
Кваліфікаційна робота і рецензія додаються.

Директор ННІКБЗІ _____ Євгенія ІВАНЧЕНКО
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

Висновок керівника кваліфікаційної роботи

Здобувач КОНДРАТЮК Дмитро у кваліфікаційній роботі проаналізував теоретичні засади механізмів управління кібербезпекою, вивчив психологічні аспекти поведінки користувачів та їхній вплив на захищеність системи, а також дослідив практичне застосування методів контролю для зміцнення кібербезпеки підприємства».

КОНДРАТЮК Дмитро показав високу теоретичну і практичну підготовку, володіння науково-дослідницькими методами, вміння самостійно знаходити шляхи вирішення проблеми дослідження. Результати дослідження апробовані на конференції "Стратегії кіберстійкості: управління ризиками та безперервність бізнесу" 28 лютого 2024 року.

Все це дозволяє оцінити кваліфікаційну роботу здобувача **КОНДРАТЮКА Дмитра** на оцінку "добре" та присвоїти йому кваліфікацію "Магістр з кібербезпеки та захисту інформації за освітньо-професійною програмою Управління інформаційною та кібернетичною безпекою".

Керівник кваліфікаційної роботи _____ Іван ОПІРСЬКИЙ
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

"___"_____2025 року

Висновок кафедри про кваліфікаційну роботу

Кваліфікаційна робота розглянута. Здобувач Кондратюк Д.О. допускається до захисту даної роботи в Екзаменаційній комісії.

Завідувач кафедрою
Управління кібербезпекою та захистом
інформації _____ Світлана ЛЕГОМІНОВА
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

ВІДГУК РЕЦЕНЗЕНТА
на кваліфікаційну магістерську роботу

здобувача вищої освіти Кондратюка Дмитра Олеговича
на тему "Механізми управління кібербезпекою з урахуванням психологічних аспектів поведінки користувачів"

Актуальність. Як свідчать реалії, людський фактор залишається однією з найуразливіших ланок у системі захисту будь-якої організації. Традиційні статичні засоби безпеки часто не враховують психоемоційний стан користувача, що призводить до помилок або навмисних порушень. Тому розробка механізмів управління кібербезпекою, які враховують психологічні аспекти поведінки, є актуальним науковим завданням для забезпечення стійкості сучасного підприємства до внутрішніх та зовнішніх загроз.

Позитивні сторони

1. У роботі досліджено теоретичні засади механізмів управління кібербезпекою та проведено аналіз психологічних факторів, що впливають на поведінку користувачів. Визначено особливості ідентифікації психологічних аномалій за допомогою систем поведінкової аналітики (UBA) та запропоновано динамічну модель оцінювання персонального кіберризику.

2. Кваліфікаційна робота оформлена відповідно до вимог. Виклад матеріалу здійснено логічно та послідовно згідно з планом. Ключові положення, зокрема архітектура моделі ризику та матриця адаптивних привілеїв, представлені у вигляді наочних рисунків та таблиць. Автор опрацював значну джерельну базу, включаючи сучасні англомовні дослідження у сферах HCI (Human-Computer Interaction) та поведінкової психології.

3. За результатами дослідження розроблено практичні рекомендації щодо впровадження адаптивного контролю доступу та системи проактивних поведінкових втручань, що дозволяє суттєво підвищити рівень захищеності підприємства.

Недоліки

1. Доцільно було б приділити більше уваги порівняльному аналізу конкретних програмних засобів класу UBA, представлених на вітчизняному ринку, та особливостям їхнього технічного налаштування для збору психологічних метрик.

Однак, вищезгадані зауваження не впливають на загальну позитивну оцінку кваліфікаційної роботи.

Висновок: Кваліфікаційна робота виконана на належному науково-методичному рівні і заслуговує позитивної оцінки, а здобувач Кондратюк Дмитро Олегович заслуговує присвоєння кваліфікації "Магістр кібербезпеки за освітньо-професійною програмою "Управління інформаційною та кібернетичною безпекою".

Рецензент:

_____
підпис *(Ім'я, ПРІЗВИЩЕ)*

РЕФЕРАТ

Текстова частина кваліфікаційної роботи на здобуття освітнього ступеня магістра: 87 с., 6 рис., 2 ф., 20 джерел.

Метою роботи є розроблення та обґрунтування концептуальної моделі управління кібербезпекою, яка динамічно адаптує захисні механізми, беручи до уваги психологічну дестабілізацію користувачів.

Об'єктом дослідження є процес управління кібербезпекою в корпоративних інформаційних системах з фокусом на моніторингу та корекції поведінки користувачів.

Предметом дослідження є модель адаптивного управління привілеями та механізми поведінкового втручання, що ґрунтуються на показниках психологічного ризику із застосуванням UBA-систем.

Методи дослідження. Емпіричні, теоретичні методи дослідження, прикладне використання знань, математичне та комп'ютерне моделювання динамічної моделі ризику.

Короткий зміст роботи. Робота містить – огляд психологічних моделей, що пояснюють порушення правил безпеки; обґрунтування застосування поведінкової аналітики для формування Ключових Індикаторів Психологічного Ризику; розробку архітектури модуля прийняття рішень та деталізацію механізмів адаптивного реагування ALP та Nudges.

У дипломній роботі досліджено механізми перетворення сирих даних моніторингу на показники психологічного ризику та їх інтеграції в систему контролю доступу. Також досліджено організацію проактивного управління ризиком, спричиненим людським фактором, на основі адаптивного підходу.

Для цього була розроблена концептуальна схема архітектури UBA-системи та обрано алгоритм динамічної зміни привілеїв. Проведені обґрунтування з метою оптимізації рівня безпеки та поліпшення процесів ідентифікації внутрішніх загроз,

підвищення надійності системи за рахунок врахування когнітивного навантаження користувачів.

Галузь застосування – підприємства IT індустрії.

ABSTRACT

The Master's qualification work includes: 87 pp., 6 figs., 2 formulas, 20 sources.

The objective of the work is to develop and justify a conceptual model of cybersecurity management that dynamically adapts defense mechanisms by considering the psychological destabilization of users.

The object of research is the process of cybersecurity management in corporate information systems, focusing on the monitoring and correction of user behavior.

The subject of research is the adaptive privilege management model and behavioral intervention mechanisms based on psychological risk indicators using UBA (User Behavior Analytics) systems.

Research methods. Empirical and theoretical research methods, applied knowledge, mathematical and computer modeling of the dynamic risk model.

Summary. The work includes: an overview of psychological models explaining security policy violations; justification for using behavioral analytics to form Key Psychological Risk Indicators (KIR); development of the decision-making module architecture and detailing of the adaptive response mechanisms, specifically ALP (Adaptive Least Privilege) and Nudges.

The thesis investigates the mechanisms for transforming raw monitoring data into psychological risk indicators and their integration into the access control system. It also explores the organization of proactive human-factor risk management based on an adaptive approach.

To achieve this, a conceptual architecture scheme for a UBA system was developed, and an algorithm for dynamic privilege modification was selected. Substantiations were provided to optimize the security level, improve internal threat identification processes, and increase system reliability by accounting for user cognitive load.

Field of application – IT industry enterprises.

ЗМІСТ

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

СУКБ	—	Система Управління Кібербезпекою
CIA	—	Конфіденційність, Цілісність, Доступність
PoLP	—	Принцип Найменших Привілеїв
PDCA	—	Плануй-Виконуй-Перевіряй-Дій
CMMI	—	Інтеграція Моделі Зрілості Кібербезпеки
UBA	—	Аналітика Поведінки Користувачів
TPB	—	Теорія Планованої Поведінки
TAM	—	Модель Прийняття Технологій
PU	—	Сприйнята Корисність
PEOU	—	Сприйнята Легкість Використання
PMT	—	Теорія мотивації захисту
HCI	—	Взаємодія Людини та Комп'ютера
KIR	—	Ключові Індикатори Ризику
ALP	—	Жорстке Технічне Блокування

ВСТУП

Сьогодні інформаційні системи та цифрові інфраструктури розвиваються з безпрецедентною швидкістю, перетворюючись на життєво важливу основу для функціонування не лише економіки, але й держави загалом. Розрив між фізичним світом та кіберпростором постійно скорочується, оскільки майже всі критичні бізнес-процеси та особисті дані інтегровані у глобальні та локальні обчислювальні мережі. У цьому контексті захист інформації перетворюється з технічного завдання на ключовий елемент стратегічної стійкості та національної безпеки.

Основною технологією захисту є забезпечення цілісності, конфіденційності та доступності даних. Проте, якщо раніше зусилля були зосереджені на побудові периметральної оборони (міжмережеві екрани, системи виявлення вторгнень), то останні роки показали, що ця стратегія є недостатньою. Непереборна кількість інцидентів, за різними оцінками, до 90% від загальної кількості, є результатом експлуатації людського фактора. Під цим терміном ховається цілий спектр вразливостей, які не можуть бути усунені лише технічними засобами. Ці вразливості включають не лише недостатню обізнаність, але й втому, стрес, когнітивне перевантаження та неуважність- чинники, що мають глибоку психологічну природу.

Технології кібербезпеки розроблялися як технології захисту мережі та об’єктів, і донедавна мережі цього класу й були єдиною областю її застосування. Однак безперечний успіх соціальної інженерії та внутрішніх загроз, які витиснули всі інші вектори атак за ефективністю, привів до ідеї про необхідність використання проактивних, адаптивних механізмів, які спрямовані не на об'єкт, а на суб'єкт- самого користувача. Комп’ютерна мережа, як основа будь-якого виробничого чи управлінського процесу, несе в собі як багато потенційних можливостей, так і зростаючу кількість внутрішніх ризиків. Це дозволяє нам приймати розробку адаптивних механізмів управління ризиками як ключове питання сучасної безпеки. Рішення питань з побудови системи управління кібербезпекою на базі гнучкого й технологічного комплексу, який відповідає сучасним технічним вимогам, з

визначенням зростаючих психологічних потреб і подальшим розвитком механізмів контролю у зв’язку з появою новіших методів аналізу поведінки дозволить нам позитивно взяти участь у сучасному ринку кіберстійкості.

В даній магістерській роботі досліджені можливі варіанти рішень для побудови адаптивних механізмів управління кібербезпекою з урахуванням психологічних аспектів поведінки користувачів. Серед них:

1. Розробка теоретичної основи для поведінкової аналітики користувачів (UBA) та формалізація набору ключових індикаторів ризику (KIR), які є об’єктивними показниками психологічного стану.

2. Створення та валідація динамічної моделі оцінки персонального кіберризику, що забезпечує кількісне перетворення індикаторів KIR у єдиний інтегральний показник ризику.

3. Розробка архітектури та алгоритмів адаптивного механізму контролю доступу (ALP) для динамічного коригування привілеїв.

4. Створення механізмів проактивних поведінкових втручань «Nudges» для ненав’язливої корекції ризикової поведінки.

5. Обґрунтування архітектурної та адміністративної інтеграції розроблених механізмів у загальну Систему управління кібербезпекою (СУКБ) та оцінка економічної ефективності.

РОЗДІЛ 1 ТЕОРЕТИЧНІ ОСНОВИ УПРАВЛІННЯ КІБЕРБЕЗПЕКОЮ ТА РОЛЬ ПОВЕДІНКОВИХ ЧИННИКІВ

Управління кібербезпекою в сучасних умовах є завданням, що виходить далеко за межі суто технологічного захисту. Хоча інвестиції в антивірусні системи, міжмережеві екрани та складні архітектури захисту постійно зростають, статистика незмінно демонструє: людський фактор залишається найбільш уразливою ланкою. За даними більшості досліджень, значна частина успішних кібератак зумовлена помилками, недбалістю або маніпуляціями користувачів, а не прорахунками техніки.

Таким чином, для побудови справді стійкої системи кібербезпеки необхідний перехід від виключно технологічно-орієнтованого підходу до людино-орієнтованого управління ризиками. Це вимагає глибокого розуміння не лише технічних загроз, але й психологічних механізмів, які лежать в основі поведінки користувачів: їхніх когнітивних упереджень, схильності до ризику, реакції на стрес та мотивації до дотримання правил.

1.1 Концептуальні засади та моделі управління кібербезпекою

1.1.1. Еволюція та сучасні стандарти систем управління кібербезпекою (СУКБ)

Система управління кібербезпекою (СУКБ) є організаційним та методологічним фундаментом для забезпечення конфіденційності, цілісності та доступності інформаційних активів (тріада CIA). Її еволюція відображає перехід від локалізованих заходів захисту окремих пристроїв до цілісного, ризик-орієнтованого підходу, що охоплює всю архітектуру підприємства.

Тріада CIA (Confidentiality, Integrity, Availability) є наріжним каменем інформаційної безпеки та визначає базові цілі будь-якої системи захисту. Конфіденційність (Confidentiality) гарантує, що доступ до інформації мають лише

авторизовані суб'єкти, запобігаючи несанкціонованому ознайомленню та витоку даних. Це забезпечується механізмами шифрування та контролю доступу. Цілісність (Integrity) забезпечує точність, повноту та надійність інформації та методів її обробки, гарантуючи, що дані не були змінені несанкціонованим або випадковим чином. Це підтримується хешуванням, електронними підписами та механізмами контролю змін. Доступність (Availability) забезпечує, що авторизовані користувачі, системи та процеси можуть отримати доступ до інформаційних активів та пов'язаних з ними ресурсів (мережевих сервісів, обчислювальної потужності), коли це необхідно. Всі заходи безпеки, включаючи розробку СУКБ та механізмів контролю, спрямовані на підтримку рівноваги між цими трьома елементами, оскільки порушення будь-якого з них вважається інцидентом безпеки.

Рис. 1.1 Тріада CIA

Початкові етапи розвитку СУКБ були зосереджені переважно на технологічних та фізичних бар'єрах. З появою глобальних мереж (Internet) та зростанням обсягів даних, фокус змістився на розробку універсальних практик та стандартів. Ключові концепції, такі як Defense-in-Depth (багатоешелонний захист) та Least Privilege (принцип

найменших привілеїв), стали основою для розробки формалізованих систем управління, які інтегрують процеси, персонал та технології.

Концепція Defense-in-Depth передбачає побудову багатошарової, ешелонованої системи захисту, де відсутність одного захисного бар'єра компенсується наявністю наступних. Замість того, щоб покладатися на один сильний механізм (наприклад, зовнішній фаєрвол), ця стратегія вимагає розміщення контрольних точок на різних рівнях архітектури- від периметра мережі, ідентифікації та аутентифікації користувачів, рівня додатків, операційних систем, аж до рівня даних. Мета полягає в тому, щоб навіть у разі подолання зловмисником одного захисного шару, він зіткнувся б з наступним. Цей принцип критично важливий, оскільки він визнає, що жоден окремий захід безпеки не є досконалим, і спрямований на уповільнення або повне блокування кібератаки на різних етапах її реалізації.

Принцип найменших привілеїв (PoLP) є наріжним каменем управління доступом. Його основна ідея полягає у тому, що будь-який суб'єкт- користувач, процес або система - повинен мати лише той мінімальний набір прав доступу, який абсолютно необхідний для виконання його поточних, визначених функцій. Наприклад, звичайному користувачу не потрібні адміністративні права для встановлення програмного забезпечення. Застосування цього принципу має подвійну мету: по-перше, воно мінімізує потенційну зону збитку у разі компрометації облікового запису або успішного використання вразливості, оскільки зловмисник, що працює під обмеженим обліковим записом, не зможе завдати значної шкоди. По-друге, він знижує ризик ненавмисних помилок або внутрішніх порушень, оскільки фізично унеможливлює виконання неавторизованих або деструктивних дій.

Сучасна СУКБ розглядається не як статичний набір інструментів, а як безперервний цикл, що функціонує за принципом PDCA (Plan-Do-Check-Act) – плануй, виконуй, перевіряй, дій. Цей циклічний підхід забезпечує постійне вдосконалення механізмів захисту відповідно до динамічних змін кіберзагроз та бізнес-середовища.

Рис. 1.2 Plan-Do-Check-Act

Впровадження СУКБ в організації зазвичай ґрунтується на визнаних міжнародних стандартах, які забезпечують структуровану основу для управління ризиками.

1. Стандарт ISO/IEC 27001 (та сімейство 27000):

Стандарт ISO/IEC 27001:2013 є найбільш визнаним міжнародним стандартом для створення, впровадження, підтримки та постійного вдосконалення СУКБ. Він не містить конкретних технічних рішень, а надає вимоги до системи управління.

Фокус: Встановлення методології управління ризиками, яка вимагає від організації ідентифікувати активи, оцінити загрози та вразливості, а також обрати відповідні заходи контролю (викладені у додатку А – ISO/IEC 27002).

Значення для теми: ISO 27001 неявно визнає людський фактор через вимоги до навчання, підвищення обізнаності (А.7.2.2) та дисциплінарних процедур (А.7.2.3), проте він фокусується на організаційних вимогах, а не на психологічному моделюванні поведінки.

2. Фреймворк кібербезпеки NIST (NIST Cybersecurity Framework – CSF):

Розроблений Національним інститутом стандартів і технологій США, фреймворк NIST CSF пропонує гнучку, добровільну основу, призначену для зменшення кіберризиків у критичній інфраструктурі. Він є менш нормативним, ніж ISO, і більш орієнтований на інтеграцію з бізнес-процесами.

Фокус: Структура побудована навколо п'яти ключових функцій: Ідентифікація (Identify), Захист (Protect), Виявлення (Detect), Реагування (Respond) та Відновлення (Recover).

Значення для теми: Функція "Захист" включає категорію "Навчання та обізнаність" (Awareness and Training), що безпосередньо стосується людського фактора. CSF допомагає визначити, на якому етапі циклу управління необхідне втручання з урахуванням психології.

3. Модель зрілості кібербезпеки CMMI (Cybersecurity Maturity Model Integration):

Ця модель використовується для оцінки зрілості процесів кібербезпеки в організації. Вона допомагає визначити, наскільки формалізовані та постійно вдосконалюються практики безпеки, що корелює з культурою безпеки та поведінкою співробітників.

Взаємозв'язок стандартів та актуальність інтеграції психології

Основна цінність зазначених стандартів полягає у їхній універсальності та структурованості. Однак, вони мають спільне обмеження:

Вони встановлюють що потрібно робити (наприклад, "проводити навчання"), але не вказують, як це робити ефективно, враховуючи когнітивні упередження, мотивацію та емоційні стани користувачів.

Таким чином, для досягнення максимальної ефективності СУКБ, імплементованої на базі ISO чи NIST, необхідна її доробка в частині Human Factor Engineering. Це формує передумови для подальшого дослідження у роботі, яке буде зосереджене на розробці механізмів управління, що враховують психологічні аспекти.

1.1.2. Класифікація та механізми управління основними видами кіберризиків

Управління кіберризиками є центральним елементом будь-якої Системи управління кібербезпекою (СУКБ), яка була розглянута у попередньому підрозділі. Під кіберризиком розуміють потенційну можливість втрати, пошкодження чи несанкціонованого доступу до інформаційних активів, що є наслідком реалізації загроз, використовуючи існуючі вразливості. Ефективне управління ризиками – це безперервний і структурований процес, що включає ідентифікацію, оцінювання, обробку (або мітигацію) та моніторинг ризиків.

Для забезпечення комплексного підходу до захисту критично важливою є класифікація ризиків, що дозволяє застосовувати відповідні механізми управління. Традиційно кіберризики поділяють на три основні категорії: технічні, організаційні та персонал-орієнтовані ризики. Технічні ризики пов’язані з апаратним та програмним забезпеченням, включно з вразливостями операційних систем, мережевих протоколів, або помилками конфігурації. Організаційні ризики стосуються недоліків у політиках, процесах, процедурах та юридичних аспектах, зокрема, відсутності плану безперервності бізнесу чи нечіткого розподілу відповідальності.

Найбільш складними для кількісної оцінки та управління є персонал-орієнтовані ризики, які прямо виникають із дій або бездіяльності користувачів та співробітників. До них відносять ризики, пов’язані із соціальною інженерією, інсайдерськими загрозами, а також випадковими помилками, спричиненими неуважністю, втомою або недостатньою обізнаністю. У контексті теми даної роботи, ця категорія є ключовою, оскільки вона вимагає застосування психологічного моделювання, а не лише технічних обмежень.

Механізми управління ризиками, які використовуються для їх обробки, також класифікуються відповідно до їхньої природи. Технічні механізми включають використання криптографічних засобів, міжмережевих екранів, систем виявлення вторгнень IDS-IPS, багатофакторної аутентифікації та автоматизованих систем резервного копіювання. Вони спрямовані на створення фізичних та логічних бар'єрів

для захисту інформації. Організаційні механізми охоплюють розробку та впровадження політик інформаційної безпеки, процедур доступу та інструкцій реагування на інциденти.

Традиційні підходи до управління кіберризиками, що ґрунтуються переважно на технічних та організаційних контролях, часто недооцінюють динамічну та ірраціональну складову персонального ризику. Наприклад, політика використання складних паролів є організаційним контролем, але психологічна схильність користувачів до спрощення або записування паролів зводить його ефективність до нуля. Це підкреслює фундаментальну прогалину: відсутність інтеграції інструментів поведінкової психології у класичні методи оцінки ризиків. Таким чином, для створення ефективного механізму управління кібербезпекою необхідно перейти до детального аналізу та моделювання саме психологічних факторів, що і стане предметом розгляду наступного підрозділу.

1.1.3. Принципи побудови багатоешелонного захисту (Defense-in-Depth) в контексті впливу користувачів

Концепція багатоешелонного захисту- Defense-in-Depth- є фундаментальним архітектурним підходом в інформаційній безпеці. Її основна ідея полягає у запобіганні повного компрометування системи внаслідок прориву одного захисного механізму. Замість покладання на один «непробивний» бар’єр, створюється багаторівнева структура контрольних механізмів, які послідовно уповільнюють або зупиняють зловмисника, незалежно від початкової точки проникнення.

Традиційно, модель Defense-in-Depth включає сім основних шарів захисту, починаючи від політик і закінчуючи фізичним захистом. Усі ці шари можна умовно поділити на технологічні, фізичні та адміністративні. Однак, у контексті людського фактору, особливої уваги вимагають шари, які безпосередньо взаємодіють з користувачем або залежать від його поведінки та рішень.

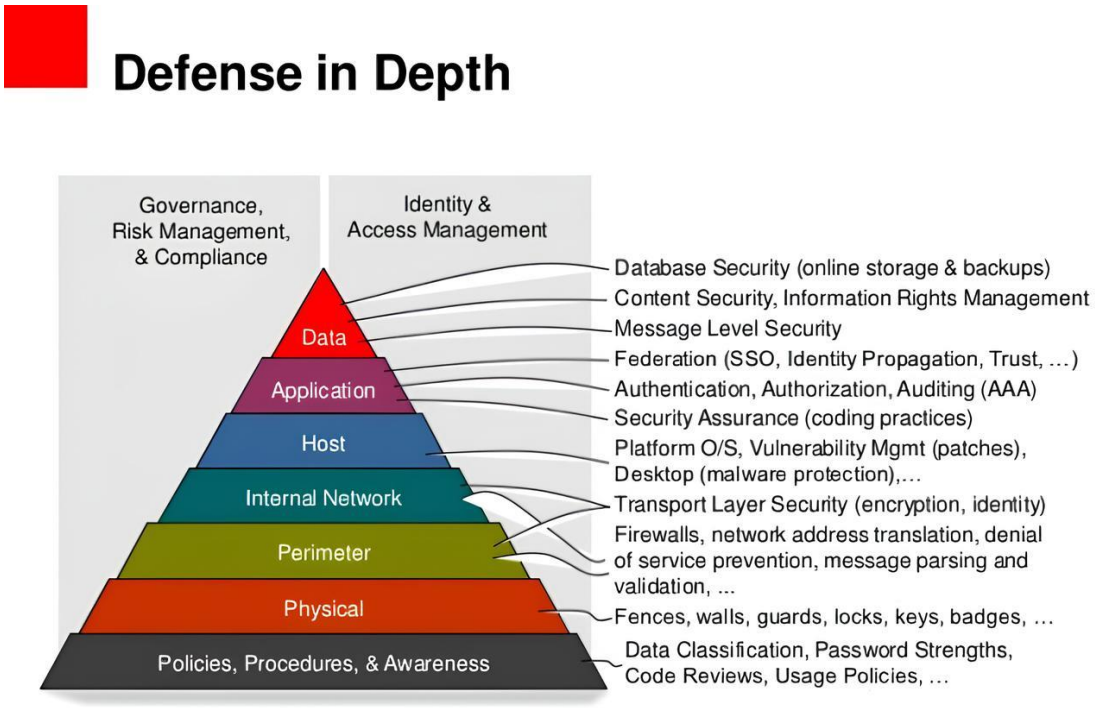

Рис. 1.3 Сім основних шарів захисту Defense-in-Depth

З погляду управління кібербезпекою з урахуванням психологічних аспектів, модель багатоешелонного захисту трансформується у послідовність, що має компенсувати ймовірність людської помилки на кожному етапі. Ефективний ешелонний захист має працювати не лише проти зовнішнього ворога, а й проти внутрішніх загроз, які часто є результатом ненавмисних дій співробітників.

Ключові шари захисту, де людський фактор відіграє вирішальну роль:

1. Політики, процедури та обізнаність (Адміністративний шар). Це перший, найширший шар. Його мета- встановити правила і забезпечити, щоб користувачі їх розуміли та дотримувалися. Психологічна проблема тут полягає у «відчуженні»- користувачі схильні ігнорувати складні або незручні політики. Ефективний багатоешелонний захист повинен адаптувати політики, роблячи їх інтуїтивно зрозумілими та мінімізуючи когнітивне навантаження.

2. Периметр (Мережевий шар). Хоча це технологічний шар- міжмережеві екрани, системи IDS/IPS - саме тут реалізуються атаки соціальної інженерії, такі як фішинг. Якщо електронний лист, що містить шкідливе посилання, проходить через технічні фільтри, подальший успіх атаки повністю залежить від психологічної стійкості та

пильності користувача. Отже, навчання користувачів виступає як додатковий, «поведінковий» фільтр.

3. Хост (Кінцева точка). Сюди входять антивірусне програмне забезпечення, персональні фаєрволи та контроль доступу. На цьому рівні користувач може вчинити фатальну помилку- наприклад, вимкнути захист для «зручності» або надати надмірні права неперевіреній програмі. Тут необхідні механізми, які не лише блокують, але й аргументують користувачеві ризик його дії, використовуючи психологічні прийоми переконання.

4. Дані (Шар даних). Це фінальний об'єкт захисту. Контроль доступу, шифрування та DLP-системи є обов'язковими. У цьому контексті, інсайдерська загроза, мотивована образою чи фінансовою вигодою, є найскладнішою. Технології User Behavior Analytics (UBA) мають на меті не просто фіксувати факт доступу, а оцінювати аномальність поведінки, що безпосередньо пов’язано з психологічним станом інсайдера.

Таким чином, для управління кібербезпекою з урахуванням психологічних аспектів, Defense-in-Depth повинна трансформуватися у Behavioral-Aware Defense-in-Depth. Кожен ешелон має не лише технологічний, але й поведінковий контроль, який активно нівелює вразливості, породжені людським фактором. Це підводить до необхідності створення інструментів для профілювання та моделювання поведінки, що і буде детально розглянуто в наступних підрозділах.

1.2 Психологічні аспекти поведінки користувачів як критичний фактор кібербезпеки

Після встановлення організаційно-методологічних засад Системи управління кібербезпекою, необхідно здійснити глибокий аналіз найменш передбачуваного елементу системи - людини. Психологічні аспекти поведінки користувачів відіграють критично важливу роль у формуванні кіберризиків. Це вимагає розгляду когнітивних, мотиваційних та емоційних чинників, які або сприяють дотриманню політик безпеки, або, навпаки, призводять до виникнення інцидентів.

У цьому підрозділі буде детально розглянуто, як саме психологія впливає на рівень захищеності інформаційних систем та закладено теоретичне підґрунтя для розробки поведінкової моделі.

1.2.1. Психологічна природа людських помилок Human Error у кіберпросторі

Людські помилки, або Human Error, є однією з найпоширеніших причин інцидентів кібербезпеки. Вони не є проявом злого умислу, а виникають внаслідок збігу обставин, пов’язаних із внутрішніми психологічними процесами та зовнішнім середовищем. Класифікація помилок, розроблена Джеймсом Різоном, поділяє їх на три основні категорії:

1. Прослизання (slips) та Промахи (lapses). Це помилки, що виникають на етапі виконання вже спланованих дій. Прослизання- це невірна фізична дія (наприклад, натискання не на ту кнопку), а промах- це помилка пам’яті (наприклад, забути виконати аутентифікацію). У кібербезпеці це може бути неправильно введена адреса електронної пошти або випадкове видалення файлу.

2. Помилки на основі правил (rule-based mistakes). Виникають, коли користувач неправильно застосовує відоме правило або слідує невірному правилу. Наприклад, використання складного пароля, але зберігання його у загальнодоступному місці, оскільки це «технічно» відповідає політиці, але порушує її дух.

3. Помилки на основі знань (knowledge-based mistakes). Відбуваються у нових або складних ситуаціях, коли людина не має готового рішення і вимушена покладатися на загальні знання, що часто призводить до некоректних висновків. Це може бути невірне оцінювання загрози під час нової, раніше не баченої фішингової атаки.

Ключові причини цих помилок у кіберпросторі часто кореняться у когнітивних навантаженнях, стресі, втомі та ефекті рутини. Щоденна багаторазова взаємодія з попередженнями безпеки призводить до звикання або втоми від безпеки (security fatigue), коли користувач починає ігнорувати або автоматично закривати попереджувальні повідомлення, не вчитуючись у їхній зміст.

1.2.2. Психологія обману та соціальна інженерія

Соціальна інженерія є найбільш прямим застосуванням психології для обходу технологічного захисту. Цей метод спрямований на маніпулювання людськими емоціями та когнітивними упередженнями для отримання конфіденційної інформації або виконання небажаних дій.

Основними психологічними тригерами, які експлуатуються зловмисниками, є:

– Авторитет. Цей тригер базується на глибоко вкоріненій у суспільстві схильності автоматично довіряти особам, що представляються керівництвом, технічною підтримкою, аудиторами або представниками державних органів, незалежно від логічності та обґрунтованості їхніх вимог. Психологічно, це є наслідком прагнення уникнути конфлікту та страху перед можливими негативними наслідками непокори вищій інстанції. Зловмисники використовують це, імітуючи офіційну комунікацію, включаючи логотипи та підписи високих посадових осіб, щоб паралізувати критичне мислення жертви. Наприклад, користувач отримує електронний лист, нібито від «Фінансового директора», з вимогою «негайно і строго конфіденційно» надати доступ до певного файлу або здійснити транзакцію, оскільки «це критичний аудит, який має бути завершений протягом години, а мій доступ заблоковано». Під тиском авторитету та страхом підвести керівництво, жертва не дотримується стандартних процедур перевірки.

– Дефіцит та терміновість. Цей механізм спрямований на маніпуляцію часом, що залишається у жертви на прийняття рішення, різко обмежуючи її здатність до раціонального аналізу. Створення штучного відчуття невідкладності або обмеженості пропозиції (дефіцит) активує ірраціональну реакцію: люди схильні переоцінювати те, що є рідкісним або скоро може бути втрачене. Це є основою успіху багатьох фішингових атак, які вимагають «негайного» підтвердження, зміни пароля або оплати рахунку. Такі повідомлення часто містять погрози, наприклад: «Ваш обліковий запис буде заблоковано через 30 хвилин», або «Залишилося лише 5 ліцензій за спеціальною ціною». Терміновість змушує користувача діяти імпульсивно, не перевіряючи

посилання чи джерело повідомлення, оскільки первинним імпульсом стає бажання уникнути втрати або штрафу.

– Взаємність. Цей тригер експлуатує соціальний принцип, згідно з яким людина відчуває підсвідомий обов'язок відповісти на отриману послугу або «подарунок». Зловмисник може спочатку надати невелику, але відчутну послугу, або навіть просто вислухати проблему жертви, встановлюючи таким чином психологічний «борг». Згодом, коли жертва відчуває себе зобов'язаною, зловмисник вимагає поступок, які можуть стосуватися безпеки. Наприклад, шахрай може, представляючись IT-фахівцем, допомогти користувачеві відновити доступ до його особистої пошти, а потім, користуючись моментом подяки, попросити «у відповідь» надіслати тимчасовий токен доступу до корпоративного сервісу, пояснюючи це технічною необхідністю та неможливістю доступу через свій обліковий запис. Відчуття боргу значно знижує опір жертви вимогам.

– Симпатія та довіра. Схильність погоджуватися на вимоги людей, які нам подобаються, знайомі або яким ми довіряємо, є потужним важелем у соціальній інженерії. Зловмисники використовують це, витрачаючи час на побудову стосунків (претекстинг), імітуючи знайомих осіб або апелюючи до спільних інтересів чи схожого життєвого досвіду. Чим більше жертва симпатизує зловмиснику або вважає його «своїм», тим нижчою стає її пильність. Наприклад, зловмисник, використовуючи скомпрометований обліковий запис колеги або друга в соціальній мережі, звертається з особистим проханням про допомогу, наголошуючи на важливості та конфіденційності, і просить перейти за посиланням або завантажити файл. Оскільки комунікація імітує довірене джерело, жертва, керуючись емоційним зв’язком, ігнорує зовнішні ознаки ризику.

Окрім експлуатації тригерів, соціальна інженерія активно використовує когнітивні упередження- систематичні помилки у мисленні, що впливають на прийняття рішень. Наприклад, упередження нормальності (normality bias) змушує людей недооцінювати ймовірність катастрофи чи загрози, оскільки «раніше цього не було». Ілюзія контролю (illusion of control) призводить до переконання, що ризик стосується інших, але не мене, зменшуючи мотивацію до виконання рутинних процедур безпеки. Усі ці фактори

у сукупності перетворюють користувача на вразливу ціль, що вимагає розробки контрзаходів, заснованих на поведінкових науках.

1.2.3. Мотиваційні та поведінкові моделі дотримання політик безпеки

Для розробки ефективних механізмів управління кібербезпекою недостатньо лише класифікувати помилки- важливо зрозуміти мотивацію користувачів до виконання або ігнорування правил безпеки. Користувач- це не лише джерело помилок, але і потенційний суб’єкт захисту. Саме тому, застосування психологічних та соціологічних моделей допомагає прогнозувати та коригувати поведінку, перетворюючи її на активний елемент системи захисту.

Основним завданням тут є пояснення поведінки людини у контексті використання технологій та дотримання норм. З цією метою можуть бути застосовані декілька фундаментальних поведінкових теорій:

1. Теорія планованої поведінки (TPB), розроблена І. Ажзеном (Icek Ajzen), є одним з найбільш впливових фреймворків у соціальній психології для прогнозування людської поведінки. Вона є розширенням раніше сформульованої Теорії обґрунтованої дії. Ключове твердження цієї моделі полягає в тому, що реальна поведінка людини визначається не лише її бажанням, а й наміром її виконати. Намір, у свою чергу, є інтегративною функцією трьох основних, незалежних, але взаємопов'язаних компонентів:

– Ставлення до поведінки (Attitude toward the behavior): Цей компонент відображає особистісну оцінку людиною певної дії. Чи вважає індивід виконання дії (наприклад, регулярне оновлення паролів або використання двофакторної аутентифікації) позитивною, корисною та сприятливою для себе, чи, навпаки, негативною, обтяжливою або зайвою. Позитивне ставлення дотримання правил безпеки часто корелює зі сприйняттям користі та ефективності цих правил.

– Суб'єктивні норми (Subjective norms): Цей елемент відображає сприйняття людиною соціального тиску. Це переконання щодо того, як інші важливі для неї люди (керівництво, колеги, друзі) ставляться до цієї поведінки і чи очікують вони від неї

виконання цієї дії. Якщо в організації панує культура, де колеги вважають за норму обходити правила безпеки заради швидкості, це негативно впливає на намір окремого співробітника дотримуватися політик.

– Сприйняття поведінкового контролю (Perceived Behavioral Control): Цей компонент є найважливішим доповненням до оригінальної теорії. Він відображає суб'єктивну оцінку людиною того, наскільки легко чи важко виконати цю дію. Фактори контролю можуть бути внутрішніми (навички, знання, впевненість) або зовнішніми (наявність часу, зручність системи, складність інтерфейсу).

У контексті кібербезпеки, TPB є надзвичайно цінною, оскільки вона допомагає пояснити розрив між знанням та дією — чому користувачі, які знають правила, все одно їх ігнорують. Часто це відбувається не через брак знань, а через низьке сприйняття контролю (наприклад, система 2FA занадто незручна і вимагає багато часу) або негативне ставлення, викликане сприйняттям незручності (правила безпеки уповільнюють роботу). Таким чином, для зміни поведінки недостатньо підвищити знання; необхідно поліпшити сприйняття контролю та сформувати позитивні суб'єктивні норми.

2. Модель прийняття технологій (TAM), розроблена Ф. Девісом (Fred Davis), є потужним інструментарієм для прогнозування того, наскільки охоче користувачі прийматимуть і регулярно використовуватимуть нові інформаційні системи та технології, що особливо актуально для засобів безпеки. Ця модель постулює, що Намір використання (Intention to Use) є прямим провісником фактичного використання, і що цей намір формується двома ключовими, взаємопов'язаними когнітивними змінними:

– Сприйнята корисність (Perceived Usefulness — PU): Ця змінна відображає ступінь, до якого користувач вірить, що використання механізму безпеки (наприклад, VPN, наскрізне шифрування або менеджер паролів) підвищить його продуктивність, ефективність або рівень особистого захисту. Якщо користувач не бачить прямої вигоди від інструменту, який, на його думку, лише сповільнює його роботу, корисність сприймається як низька.

– Сприйнята легкість використання (Perceived Ease of Use — PEOU): Цей чинник визначає, наскільки користувач вважає, що використання механізму є вільним від зусиль — тобто, наскільки легко його освоїти, інтегрувати у щоденну рутину та застосовувати без значного когнітивного навантаження чи необхідності подолання складних інтерфейсів.

TAM чітко показує, що ці два фактори мають сильний вплив на поведінку. Зокрема, якщо механізм безпеки є надмірно складним (низька PEOU) або не сприймається як корисний для щоденної роботи (низька PU), користувач майже гарантовано активно шукатиме шляхи його обходу (наприклад, запише пароль на стікері або вимкне двофакторну аутентифікацію, якщо це можливо), незалежно від адміністративних вимог та накладених санкцій. Цей принцип є критичним для проєктування ефективних механізмів управління, оскільки безпека повинна бути вбудована та невидима, а не нав'язана як незручна перешкода.

3. Теорія мотивації захисту (PMT), розроблена Р. Роджерсом (R.W. Rogers), є однією з найбільш релевантних моделей для пояснення поведінки у сфері ризиків та здоров'я, що безпосередньо проєктується на кібербезпеку. Вона передбачає, що намір захистити себе (тобто мотивація захисту) залежить від інтеракції двох основних, складних когнітивних процесів:

– Оцінка загрози (Threat Appraisal): Цей процес включає сприйняття серйозності (Severity) загрози (наприклад, який фінансовий або репутаційний збиток може завдати успішний фішинг) та сприйняття власної вразливості (Vulnerability) до цієї загрози (наприклад, наскільки ймовірно, що саме я клікну на фішингове посилання). Якщо користувач вважає, що загроза несерйозна або його це не стосується, мотивація захисту буде низькою.

– Оцінка копінгу/подолання (Coping Appraisal): Цей процес фокусується на засобах реагування і також має два виміри: сприйняття ефективності захисної реакції (Response Efficacy) (наприклад, чи справді оновлення антивірусу або використання VPN допоможе уникнути загрози) та сприйняття власної здатності виконати цю реакцію (Self-Efficacy) (наскільки я впевнений, що зможу правильно налаштувати та використовувати цей засіб захисту).

Критичний висновок PMT для кібербезпеки полягає у наступному: якщо користувач вважає, що загроза є серйозною (висока оцінка загрози), але при цьому засоби її подолання складні або неефективні (низька оцінка копінгу), мотивація захисту буде низькою. У цьому випадку індивід, зіткнувшись із загрозою, обере дезадаптивний шлях - заперечення, ігнорування ризику або припинення використання захисних механізмів, щоб зменшити психологічний дискомфорт від непереборної загрози. Це безпосередньо пояснює, чому складні корпоративні політики безпеки часто саботуються.

Всі ці моделі є критично важливим інструментом для переходу до другого розділу роботи. Вони забезпечують теоретичне підґрунтя для розробки моделі оцінювання ризиків. Аналізуючи, які саме психологічні змінні є найбільш впливовими у конкретній організації (наприклад, легкість використання чи суб'єктивні норми), можна створювати цільові та ефективні механізми управління, замість універсальних, які часто не спрацьовують. Це дозволяє впровадити психологічно обґрунтовані механізми управління кібербезпекою.

1.3. Методологічні підходи до дослідження поведінки користувачів

Якісна інтеграція психологічних аспектів у механізми управління кібербезпекою вимагає застосування науково обґрунтованих методів збору та аналізу даних про поведінку користувачів. Методологічна база дослідження охоплює як традиційні соціологічні методи оцінки рівня обізнаності, так і сучасні інструменти поведінкової аналітики, що дозволяють кількісно оцінити ризик, спричинений людським фактором. Вибір методів залежить від мети - чи потрібно виявити суб'єктивне сприйняття ризику, чи об'єктивно виміряти фактичну поведінку.

1.3.1. Кількісні та якісні методи оцінки рівня кібербезпекової культури

Дослідження культури кібербезпеки в організації - це важливий етап, який виявляє, наскільки глибоко принципи безпеки інтегровані у щоденну діяльність співробітників.

Кількісні методи є основою для отримання статистично значущих даних. Вони включають масові опитування та анкетування. Розробка опитувальників повинна ґрунтуватися на теоретичних моделях (як PMT, TAM), щоб вимірювати не лише знання (що робити), а й ставлення (мотивація робити) та самоефективність (здатність це зробити). Прикладами кількісних метрик можуть бути індекс схильності до ризику, рівень сприйнятої загрози та частота порушень правил безпеки. Використання стандартизованих шкал Лайкерта дозволяє перетворити суб'єктивні оцінки на числові дані, придатні для статистичного аналізу.

Якісні методи слугують для поглибленого розуміння мотивів, бар’єрів та контексту, що стоїть за кількісними показниками. До них відносяться індивідуальні інтерв’ю, фокус-групи та аналіз історій інцидентів. Наприклад, інтерв’ю з користувачами, які стали жертвами фішингу, може виявити емоційні або організаційні фактори, що сприяли їхній помилці, які неможливо зафіксувати у закритому опитувальнику. Ці методи особливо цінні для виявлення прихованих «суб'єктивних норм» або неформальних практик, що порушують безпеку.

Комбінація кількісних та якісних методів, так званий змішаний підхід, забезпечує високу валідність дослідження. Кількісні дані показують масштаб проблеми, а якісні - її причини та механізми.

1.3.2. Використання поведінкової аналітики UBA у системах управління безпекою

Якщо соціологічні методи вимірюють, як користувачі сприймають безпеку, то поведінкова аналітика UBA, User Behavior Analytics - вимірює, як користувачі фактично поводяться. UBA є ключовим інструментом для об'єктивного вимірювання психологічних аспектів.

UBA системи збирають та аналізують широкий спектр даних про дії користувачів, включаючи, але не обмежуючись ними, час входу, географічне розташування, послідовність доступу до ресурсів, частоту друку документів, використання зовнішніх носіїв та швидкість набору тексту. Мета UBA- створити базовий, або «нормальний» поведінковий профіль для кожного співробітника.

Після створення профілю, система безперервно моніторить відхилення від цього шаблону - аномалії. Ці аномалії можуть бути індикаторами двох типів ризику:

1. Компрометація облікового запису (Compromised User) - цей тип ризику виникає, коли ідентифікаційні дані користувача були викрадені, а його обліковий запис перехоплений зовнішнім зловмисником. Аномалії у цьому випадку часто носять різкий, інтенсивний та очевидно нелогічний характер у порівнянні з історичною поведінкою. Це можуть бути різкі зміни в логіці роботи — наприклад, одночасний вхід у систему з двох географічно віддалених місць (Impossible Travel), спроба доступу до ресурсів, які ніколи раніше не використовувалися, або виконання автоматизованих, послідовних команд, характерних для шкідливих скриптів. UBA-система у цьому випадку має не просто фіксувати, та миттєво ідентифікувати цей збіг факторів, що прямо вказує на Account Takeover (захоплення облікового запису) або Credential Theft (крадіжку облікових даних), вимагаючи негайного технічного втручання.

2. Психологічна дестабілізація та Внутрішня загроза (Insider Threat) - цей тип ризику є набагато тоншим і пов'язаний з внутрішнім станом легітимного користувача. Поведінка, що відхиляється від норми, тут може бути як ненавмисною (спричиненою втомою, стресом, когнітивним перевантаженням, що призводить до помилок), так і навмисною (свідчить про незадоволеність, підготовку до крадіжки даних або саботажу). Індикаторами можуть бути: раптовий експорт великого обсягу даних перед звільненням, нетипово висока активність у неробочий час або постійні, але обережні спроби доступу до заборонених ресурсів. На відміну від компрометації, ці аномалії часто розвиваються повільно, але є стійкими. Їхня коректна ідентифікація вимагає психологічної інтерпретації, оскільки вони можуть свідчити про критичний рівень незадоволеності, вигорання (Burnout) або формування зловмисного наміру, що і робить цей ризик найскладнішим для традиційних засобів захисту.

Використання UBA переводить управління персональними ризиками на проактивний рівень. Воно дозволяє застосовувати адаптивний контроль- наприклад, якщо поведінка користувача стає підозрілою, система може автоматично вимагати повторну аутентифікацію або тимчасово обмежити доступ до критичних даних.

Разом з тим, впровадження UBA вимагає ретельного розгляду етичних та юридичних аспектів. Необхідно дотримуватися політик конфіденційності та забезпечити, щоб моніторинг був спрямований виключно на безпеку, а не на тотальний контроль персоналу. Методи, розглянуті у цьому підрозділі, закладають основу для моделі оцінювання ризиків, яка буде розроблена у наступному розділі.

1.4. Аналіз існуючих підходів до інтеграції психології в СУКБ

Усвідомлення критичної ролі людського фактора в кібербезпеці призвело до пошуку нових шляхів, які б дозволили ефективно інтегрувати знання поведінкової психології у класичні механізми управління. Сучасні дослідники та практики дедалі більше відходять від суто технологічних контролів, визнаючи, що найбільш стійкий захист забезпечується зміною поведінки користувачів. Це вимагає розробки та впровадження так званих людино-орієнтованих механізмів управління.

Інтеграція психології в СУКБ здійснюється за кількома основними напрямками, кожен з яких намагається вирішити конкретну проблему- від підвищення загальної обізнаності до тонкої настройки політик безпеки. В даному підрозділі буде систематизовано наявний світовий досвід та проаналізовані найбільш ефективні практики. Зокрема, значна увага приділяється еволюції програм підвищення обізнаності, які відходять від формального читання лекцій до використання гейміфікації та адаптивного контенту. Окрім того, важливим є огляд розробок у сфері поведінково-орієнтованих політик та технічних засобів моніторингу.

Цей аналіз дозволить чітко визначити прогалини в існуючих механізмах. Більшість наявних підходів є реактивними або фокусуються на зміні поведінки без достатнього кількісного її вимірювання. Саме тому, кінцевою метою підрозділу є обґрунтування

необхідності розробки уніфікованої моделі оцінки психологічного ризику, яка і стане центральним елементом наступного розділу дипломної роботи.

1.4.1. Світові практики підвищення обізнаності Security Awareness з урахуванням психології

Традиційна парадигма підвищення обізнаності (Security Awareness) у сфері кібербезпеки була переважно сфокусована на передачі знань - що таке фішинг, як вибрати пароль тощо. Проте, численні інциденти свідчать про так званий розрив знань і поведінки, коли співробітники знають, як діяти правильно, але в критичний момент ухвалюють помилкове рішення. Сучасні світові практики спрямовані на подолання цього розриву шляхом інтеграції поведінкових наук та психології у навчальні програми.

Еволюція програм обізнаності від пасивних до адаптивних

Першою зміною стало усвідомлення, що ефективність програм обізнаності вимірюється не кількістю прослуханих годин, а реальною зміною поведінки. Це призвело до переходу від пасивних лекцій до симуляційних навчань, зокрема імітації фішингових атак. Психологічна цінність таких симуляцій полягає у створенні ефекту емоційного навчання. Коли користувач робить помилку і стикається з її імовірними наслідками у безпечному середовищі, це формує стійкіший рефлекс уважності, ніж просте теоретичне попередження.

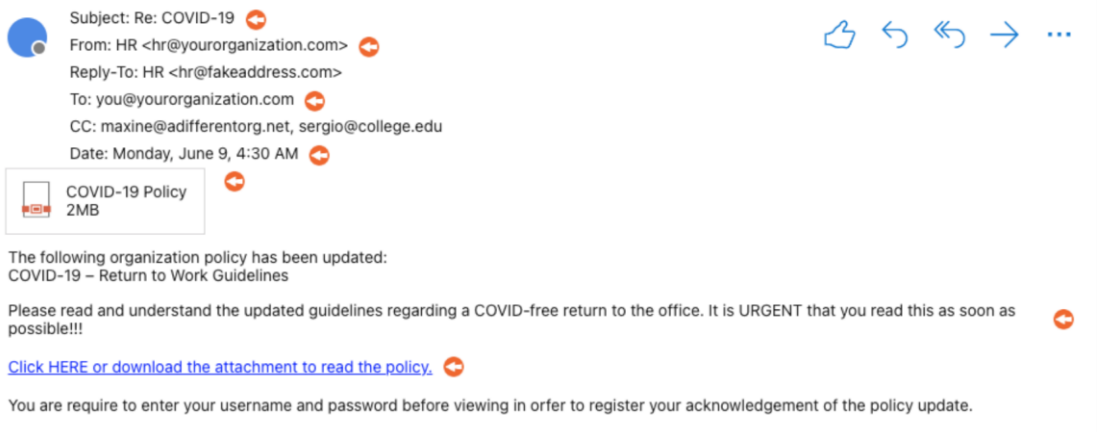

Рис. 1.4 Приклад вправи, включений до щорічного тренінгу з кібербезпеки

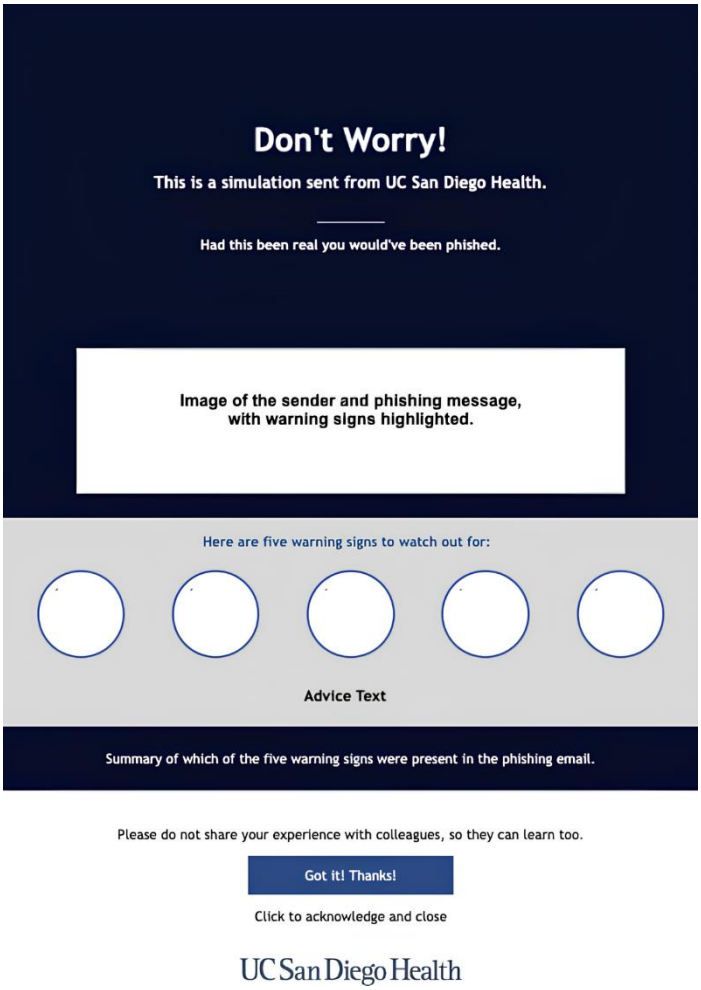

Рис. 1.5 Приклад зображення, яке відображалося працівникам, що провалили тест

Ключовим психологічним інструментом, що використовується у сучасних програмах, є гейміфікація. Вона застосовує ігрові елементи- змагання, нарахування

балів, рейтинги, віртуальні нагороди- для підвищення внутрішньої мотивації та залученості користувачів. Використання гейміфікації ґрунтується на теорії самовизначення, яка показує, що люди більш охоче виконують дії, коли відчувають автономію і компетентність. Навчання перетворюється з нудного обов'язку на виклик, що стимулює до постійного вдосконалення знань.

Персоналізація та мікронавчання як інструменти проти втоми від безпеки

Однією з головних психологічних перешкод є втома від безпеки (security fatigue), спричинена надмірною кількістю складних правил, частими попередженнями та необхідністю запам’ятовувати багато паролів. Для подолання цієї втоми використовуються такі підходи:

1. Мікронавчання - це сучасний підхід до підвищення обізнаності, який передбачає надання навчального контенту невеликими, легко засвоюваними порціями (тривалістю від 30 секунд до 5 хвилин), безпосередньо у момент, коли знання є найбільш релевантними, відповідно до принципу Just-in-Time (JIT) Learning. Замість тривалих курсів, контент подається у форматі коротких відео, інфографіки, інтерактивних підказок або міні-вікторин. Наприклад, система може автоматично видати коротке нагадування про правила безпеки даних та політику обробки персональних даних безпосередньо перед їхнім завантаженням на зовнішній хмарний сервіс або перед відкриттям критичного внутрішнього документа. Психологічна перевага цього методу полягає у зниженні когнітивного навантаження та мінімізації порушення «стану потоку» (flow state) користувача. Оскільки інформація надається саме тоді, коли вона потрібна для виконання конкретного завдання, це підвищує ймовірність негайного застосування знань та перетворення їх на стійку, безпечну поведінку.

2. Адаптивне навчання являє собою технологічний стрибок у підвищенні обізнаності, де навчальний контент персоналізовано підлаштовується під індивідуальний рівень знань, навичок та поведінковий профіль користувача, який може бути визначений за результатами попередніх тестів, аналізу інцидентів або даних UBA-системи. Цей підхід використовує дані для ідентифікації прогалин у знаннях та поведінці. Наприклад, якщо аналіз показує, що користувач регулярно провалює симуляції фішингу, але є компетентним у питаннях безпеки паролів, система

зосереджується лише на темах, пов'язаних з ідентифікацією соціальної інженерії та фішингових ознак, ігноруючи теми, в яких він уже компетентний. Такий фокусований підхід робить навчання більш цілеспрямованим, ефективним та залучаючим, оскільки користувачі не витрачають час на повторення вже засвоєного матеріалу, що, у свою чергу, підвищує їхню мотивацію до навчання.

Ефективні програми обізнаності, засновані на психології, розглядають користувачів не як пасивних отримувачів інформації, а як активних учасників процесу захисту. Вони використовують принципи позитивного підкріплення замість покарання, акцентуючи увагу на правильних діях, а не лише на помилках. Однак, незважаючи на значний прогрес, більшість цих програм залишаються відокремленими від основних механізмів управління ризиками, що створює необхідність у більш глибокій інтеграції психологічних метрик.

1.4.2. Поведінково-орієнтовані політики та процедури кібербезпеки

Ефективність Системи управління кібербезпекою (СУКБ) значною мірою залежить від якості адміністративних контролів - політик, правил та процедур. Якщо традиційні політики часто створювалися з погляду технічної зручності чи юридичної необхідності, то поведінково-орієнтовані підходи прагнуть адаптувати ці правила до реальних психологічних обмежень користувачів. Мета полягає у зменшенні когнітивного навантаження та мінімізації необхідності приймати складні рішення, які можуть призвести до помилки.

Одним із ключових напрямків є застосування принципів Human-Centric Design (Дизайн, орієнтований на людину) до розробки політик. Замість того, щоб карати за порушення, такі політики створюють так звані «захисні архітектури», які роблять правильну поведінку найлегшою. Наприклад, замість того, щоб вимагати від користувача запам'ятовувати складний і довгий пароль, політика може вимагати від системи автоматичного використання менеджера паролів або двофакторної аутентифікації. У цьому випадку, правильна поведінка, така як використання

складного пароля, досягається без зусиль з боку користувача, нівелюючи його психологічну схильність до спрощення.

Важливим елементом є архітектура вибору та застосування Теорії підштовхування (Nudge Theory). Ця теорія, запозичена з поведінкової економіки, пропонує використовувати невеликі, ненав'язливі втручання для спрямування користувачів до безпечного вибору. Наприклад, під час збереження файлу на зовнішній носій система може не просто видати технічне попередження, а надати контекстуальне нагадування про конфіденційність даних, використовуючи візуальний або текстовий «підштовхувач», що апелює до відчуття відповідальності.

Інтеграція психології у процедури реагування на інциденти

Поведінкові принципи також інтегруються у процедури реагування на інциденти. Коли інцидент, як-от фішинг або витік даних, вже стався, важливо, щоб користувачі діяли швидко і без зайвого стресу. Складні та бюрократичні процедури звітності про інциденти часто призводять до того, що користувачі приховують помилки, боячись покарання.

Політики, орієнтовані на психологію, вимагають створення культури безпеки без провини (No-Blame Culture). У таких процедурах акцент робиться на швидкому виявленні та усуненні технічної причини інциденту, а не на пошуку винного. Психологічно це стимулює співробітників швидше повідомляти про свої помилки чи підозрілі дії, що є критично важливим для мінімізації збитків.

Таким чином, поведінково-орієнтовані політики та процедури вимагають від СУКБ постійної адаптації:

- Мінімізація тертя: зменшення кроків, необхідних для виконання безпечної дії.
- Контекстуалізація: надання інформації про безпеку лише тоді, коли вона є найбільш релевантною.
- Стимулювання: використання позитивного підкріплення для заохочення безпечної поведінки.

Незважаючи на прогрес у розробці цих практик, існуючі політики часто залишаються статичними. Вони орієнтовані на ідеального користувача і не враховують

індивідуальну динаміку психологічного стану, що робить їх недостатньо ефективними проти складних загроз, які експлуатують саме ці індивідуальні особливості.

1.4.3. Визначення прогалин в управлінні кібербезпекою, що потребують психологічного моделювання

Комплексний аналіз існуючих систем управління кібербезпекою, стандартів ISO, NIST та світових практик підвищення обізнаності свідчить про значний прогрес у розумінні ролі людського фактора. Проте, незважаючи на впровадження гейміфікації, адаптивного навчання та поведінково-орієнтованих політик, більшість СУКБ все ще функціонують з критичною методологічною прогалиною. Ця прогалина полягає у нездатності ефективно перетворити якісні психологічні фактори на кількісні метрики, придатні для інтеграції у формалізовані процеси оцінки ризиків.

Ця проблема виявляється у трьох ключових аспектах:

1. Статичний характер механізмів управління. Більшість контролів безпеки - політики, навчання, навіть двофакторна аутентифікація - є статичними, тобто вони застосовуються однаково до всіх користувачів і не змінюються залежно від поточних внутрішніх умов. Однак, психологічний стан користувача (рівень втоми, стресу, емоційна стійкість) є динамічним. Статичний контроль не може компенсувати ризик, який виникає, коли висококомпетентний користувач перебуває у стані сильного стресу. Це вимагає розробки адаптивних механізмів, здатних реагувати на зміни поведінкового профілю в реальному часі.

2. Відсутність уніфікованої кількісної моделі психологічного ризику. Існуючі моделі оцінки ризиків, наприклад, на основі формули R = P x I (Ризик = Ймовірність x Наслідки) ефективно оцінюють технічну вразливість, але недостатньо точно визначають Ймовірність (P), коли вона зумовлена людською помилкою. Сучасні інструменти поведінкової аналітики UBA лише фіксують аномалії, але не надають інструментарію для інтерпретації цих аномалій з погляду психологічної причини (навмисний злий умисел, чи ненавмисна помилка через втому). Це створює нагальну

потребу у розробці математичної або структурної моделі, яка дозволить перевести психологічні змінні у кількісні показники ризику.

3. Ігнорування когнітивного навантаження як фактора ризику. Багато правил безпеки, які вважаються ефективними на папері(часта зміна паролів, складні VPN-процедури) насправді збільшують когнітивне навантаження на користувача. Психологічно, це призводить до компенсаторної поведінки - спрощення паролів, записування їх на стікерах, або пошук шляхів обходу захисту. Існуючі підходи не пропонують систематичних методів для вимірювання та оптимізації політик з погляду мінімізації цього навантаження, що є життєво необхідним для стійкої кібербезпеки.

Таким чином, критичною прогалиною є відсутність механізму, який би поєднував психологічні теорії (розглянуті у підрозділі 1.2) з практичними інструментами (UBA, обізнаність) для створення проактивного, адаптивного управління кібербезпекою. Розробка такого механізму, що перетворює психологічні аспекти поведінки користувачів на інтегровану, вимірювану частину СУКБ, є безпосередньою науковою новизною даної дипломної роботи та головною задачею, що має бути вирішена у наступних розділах.

1.5. Висновки до розділу 1

Проаналізувавши теоретичні засади управління кібербезпекою та роль людського фактора, можна стверджувати, що сучасна парадигма захисту інформації зазнає суттєвої трансформації, переходячи від суто технологічного до людино-орієнтованого підходу. Дослідження еволюції систем управління кібербезпекою (СУКБ) та міжнародних стандартів, таких як ISO/IEC 27001 та NIST CSF, продемонструвало, що хоча вони й створюють надійний організаційний фундамент, їхня ефективність часто нівелюється ігноруванням ірраціональної природи користувача та його когнітивних упереджень. Це зумовлює виникнення «прірви» між формальними вимогами безпеки та реальною поведінкою персоналу, де статичні принципи, як-от тріада CIA або концепція Defense-in-Depth, стають вразливими перед методами соціальної інженерії та психологічними тригерами.

Розуміння природи людських помилок через призму моделі Дж. Різона та аналіз психологічних механізмів обману (авторитету, терміновості, взаємності) підтверджують, що користувач є не просто «слабкою ланкою», а динамічним елементом системи, стан якого постійно змінюється під впливом стресу, втоми або когнітивного навантаження. Саме тому застосування таких поведінкових теорій, як TPB, TAM та PMT, є критично важливим, оскільки вони дозволяють виявити внутрішні чинники - ставлення до правил, суб'єктивні норми та сприйняту легкість використання засобів захисту - які безпосередньо формують намір дотримуватися або саботувати політику безпеки.

З огляду на це, методологічна інтеграція поведінкової аналітики (UBA) у загальну стратегію управління ризиками постає як необхідний інструмент для об'єктивного моніторингу цих процесів. Перехід від суб'єктивної оцінки обізнаності до кількісного вимірювання фактичних поведінкових аномалій дозволяє ідентифікувати реальний рівень загрози ще до моменту вчинення інциденту. Таким чином, результати теоретичного аналізу підтверджують гіпотезу про необхідність розробки динамічних механізмів реагування, які б адаптували рівень доступу та методи взаємодії з користувачем залежно від його поточного психологічного профілю, що і стане об'єктом проектування у наступному розділі роботи.

РОЗДІЛ 2 АНАЛІЗ ПСИХОЛОГІЧНИХ ФАКТОРІВ КОРИСТУВАЧІВ ТА РОЗРОБКА МОДЕЛІ ОЦІНЮВАННЯ РИЗИКІВ

2.1. Аналіз архітектур User Behavior Analytics (UBA) та методи ідентифікації психологічних аномалій

Ефективна інтеграція психологічних аспектів у механізми управління кібербезпекою вимагає об'єктивного вимірювання поведінки користувачів, що є можливим завдяки використанню систем поведінкової аналітики. User Behavior Analytics. UBA - це клас рішень, що застосовують машинне навчання, статистику та алгоритми для виявлення аномалій у поведінкових патернах співробітників, які можуть свідчити про ризик, або компрометацію, або наявність інсайдерської загрози, або, що найбільш важливо для нашої теми, про схильність до помилок через психологічні фактори.

2.1.1. Концептуальні засади та архітектура систем User Behavior Analytics (UBA) у кібербезпеці

Системи UBA є еволюційним розвитком традиційних систем управління інформаційною безпекою та подіями - SIEM, Security Information and Event Management. Якщо SIEM-системи зосереджені на кореляції логів та подій з різних джерел для виявлення відомих атак, то UBA-системи фокусуються на користувачеві як точці ризику. UBA має на меті не просто фіксувати подію, а створити та підтримувати базовий профіль нормальної поведінки кожного окремого користувача, незалежно від його посади.

Архітектурно UBA-системи складаються з чотирьох основних функціональних блоків:

1. Блок збору та агрегації даних. На цьому етапі відбувається збір сирих даних з великої кількості джерел- систем аутентифікації, мережевих логів, журналів доступу

до файлових систем, електронної пошти та кінцевих точок. Ключовою вимогою є уніфікація різнорідних даних для подальшої обробки.

2. Блок нормалізації та профілювання. Зібрані дані очищаються та використовуються для побудови базового профілю. Профіль включає такі параметри, як-то типовий час роботи, географічне розташування, часто використовувані ресурси, обсяги даних, що завантажуються. Цей етап критично важливий, оскільки саме він визначає, що є «нормальною» поведінкою для конкретного співробітника.

3. Блок аналізу та виявлення аномалій. Це ядро системи, де застосовуються алгоритми машинного навчання(як керованого, так і некерованого) для постійного порівняння поточної поведінки користувача з його базовим профілем. Відхилення від норми класифікуються як аномалії.

4. Блок оцінки ризику та візуалізації. Виявлені аномалії перетворюються на показник ризику, який інтегрується в загальний рейтинг. Системи UBA надають оператору не просто список подій, а ранжований перелік користувачів з найвищим інтегральним ризиком.

На відміну від традиційних систем безпеки, які працюють за принципом «якщо X, то Y», UBA використовує ймовірнісний підхід - «якщо поведінка користувача відхиляється від його власної норми на Z відсотків, то ймовірність ризику становить P». Це дозволяє виявляти загрози, які є унікальними для конкретного користувача, що ідеально підходить для оцінки психологічно-зумовлених ризиків, які неможливо уніфікувати.

2.1.2. Вибір та обґрунтування технічних показників, що відображають психологічні стани користувачів

Для розробки ефективної моделі оцінювання психологічного ризику необхідно встановити прямий зв’язок між сирими технічними даними, які збирає UBA-система, та внутрішніми психологічними станами користувача- втомою, стресом, когнітивним навантаженням чи навіть прихованою ворожістю. Ці технічні показники мають бути обґрунтовані з погляду поведінкової психології та human-computer interaction(HCI).

Взаємодія Людини та Комп'ютера (Human-Computer Interaction- HCI) - це міждисциплінарна галузь, що вивчає проєктування, оцінку та впровадження інтерактивних обчислювальних систем для використання людиною, а також вивчає самі феномени, що їх оточують. Головною метою HCI є підвищення Us (зручності та ефективності використання) та доступності систем, забезпечуючи, щоб взаємодія користувача з технологією була максимально інтуїтивно зрозумілою, безпечною та продуктивною. Дисципліна поєднує знання з комп'ютерних наук (проєктування інтерфейсів), когнітивної психології (розуміння того, як люди сприймають і обробляють інформацію) та ергономіки. У контексті кібербезпеки, HCI стає критично важливим, оскільки погано спроєктовані інтерфейси можуть збільшувати когнітивне навантаження та призводити до помилок користувача, які, у свою чергу, відкривають шляхи для інцидентів безпеки.

Індикатори когнітивного навантаження та втоми

Психологічні дослідження показують, що підвищене когнітивне навантаження та втома безпосередньо призводять до зниження пильності та збільшення ймовірності помилок (slips and lapses). Ці стани можна відстежити за такими технічними метриками:

- Частота та динаміка помилок автентифікації. Незначне, але стійке зростання кількості помилок при введенні пароля або PIN-коду може свідчити про втому, відволікання або розсіяну увагу, а не про спробу несанкціонованого доступу.
- Швидкість та патерни набору тексту (Keystroke Dynamics). Зменшення середньої швидкості набору, збільшення кількості виправлень або нерівномірний ритм введення можуть бути індикаторами фізичної та розумової втоми. З іншого боку, різка зміна патерну може свідчити про використання стороннього програмного забезпечення або навіть про компрометацію облікового запису.
- Тривалість робочої сесії та час відпочинку. Надмірно довга робоча сесія без перерв або нетипово пізня активність може бути прямим індикатором втоми і, відповідно, підвищеного ризику прийняття хибних рішень (наприклад, відкриття фішингового листа).

– Час, витрачений на типові операції. Збільшення часу, необхідного користувачеві для виконання рутинних завдань (наприклад, сортування пошти, заповнення форми), порівняно з його базовим профілем, також може вказувати на зниження когнітивної ефективності.

Індикатори емоційної дестабілізації та умисного ризику

Окрім неуважності, психологічний ризик включає й умисні дії, часто зумовлені емоційною дестабілізацією, наприклад, незадоволеністю, образою, або фінансовою необхідністю. Ці індикатори є ключовими для виявлення інсайдерської загрози, що є однією з найбільш руйнівних.

– Аномальний доступ до конфіденційних або заборонених ресурсів. Різка зміна шаблонів доступу- спроба отримати доступ до даних, які не потрібні для виконання посадових обов'язків, або до інформації, яку користувач ніколи раніше не використовував.

– Обсяги та типи завантажених даних. Нетипово велике завантаження даних на зовнішні носії, хмарні сховища або використання несанкціонованих сервісів передачі файлів може свідчити про намір витоку даних, часто спричинений особистою мотивацією.

– Ігнорування попереджень системи безпеки. Багаторазове та послідовне ігнорування або відключення попереджувальних повідомлень (наприклад, про необхідність оновлення ПЗ або про ризикований вебсайт), порівняно з базовою нормою, може бути проявом умисного ігнорування політик.

Ці технічні метрики, зібрані UBA-системою, є сирими даними. Наступний етап дослідження полягатиме у застосуванні методів машинного навчання, щоб навчитися інтерпретувати ці показники як істинні індикатори психологічного ризику.

2.1.3. Методи машинного навчання для виявлення аномалій, зумовлених психологічними факторами

Виявлення аномалій у поведінці користувачів є фундаментальною задачею UBA-систем. У контексті оцінки психологічного ризику, методи машинного навчання

застосовуються для двох ключових цілей: по-перше, для точного визначення базового (нормального) профілю, і по-друге, для класифікації відхилень від цього профілю як потенційно психологічно-зумовлених ризиків.

Для побудови базового профілю нормальної поведінки найчастіше використовуються некеровані методи навчання, оскільки вони не вимагають попереднього маркування даних про інциденти.

– Кластеризація (Clustering): Алгоритми, такі як K-Means або DBSCAN, використовуються для групування користувачів зі схожими поведінковими шаблонами. Це допомагає визначити "нормальність" не лише для індивідуума, але й для його функціональної групи. Наприклад, поведінка розробника уночі може бути нормальною, тоді як така ж поведінка бухгалтера є аномалією.

– Статистичне моделювання часових рядів: Для динамічних метрик, наприклад, обсяг завантажень або час роботи, застосовуються моделі, такі як ARIMA, для прогнозування очікуваної поведінки у певний момент часу. Відхилення фактичної поведінки від прогнозованого діапазону (викиди) розглядається як аномалія.

Ідентифікація та інтерпретація аномалій

Після визначення нормального профілю, наступний крок - це виявлення та інтерпретація відхилень. Аномалія може бути спричинена технічним збоєм, або ж вона може бути результатом психологічного фактора- втоми, стресу чи злого наміру.

– Методи виявлення відхилень (Outlier Detection), також відомі як виявлення аномалій, є фундаментальним і найбільш поширеним підходом у поведінковій аналітиці користувачів (UBA). Суть цих методів полягає у безперервному порівнянні поточної, миттєвої поведінки суб'єкта (користувача або системи) з його історичним базовим профілем (Baseline Profile). Цей профіль являє собою статистично обґрунтований "портрет норми", що включає такі параметри, як типовий час входу та виходу з системи, звичайний обсяг завантажених даних, частота використання певних додатків або типовий географічний регіон активності. Кожна дія, здійснена користувачем (наприклад, вхід у нетиповий час або спроба доступу до нехарактерного ресурсу), аналізується та присвоюється їй оцінка аномальності (Anomaly Score). Ця оцінка є кількісним вираженням ступеня відхилення від статистичної норми,

встановленої у профілі. Якщо ця оцінка перевищує попередньо встановлений динамічний поріг (Threshold), дія класифікується як аномальна. Ключова перевага цих методів полягає в їхній здатності виявляти як раптові, різкі відхилення (наприклад, аномальна кількість невдалих спроб входу), що можуть свідчити про компрометацію облікового запису, так і повільні, ледь помітні зміни в поведінці, що часто вказують на зростаючу втому, стрес або підготовку до внутрішнього інциденту. Саме на основі агрегації цих аномальних оцінок і формується інтегральний показник ризику.

– Байєсівські мережі (Bayesian Networks) є особливо цінним та інтелектуально просунутим інструментом у поведінковій аналітиці, оскільки вони виходять за рамки простого виявлення аномалій і переходять до їхньої контекстуальної інтерпретації та прогнозування. Ці моделі використовують теорію ймовірностей, графічно представляючи причинно-наслідкові зв'язки між різними подіями та змінними (вузлами). Вони можуть ефективно інтегрувати кілька незалежних, але взаємопов'язаних аномалій, наприклад, стійке зростання кількості помилок аутентифікації, нетипово збільшений час реакції користувача на запити системи та одночасний несанкціонований доступ до папок, які раніше не використовувалися. Байєсівська мережа обчислює умовну ймовірність того, що ця комбінація подій є результатом певної психологічної причини (наприклад, втома чи стрес), або ж результатом зовнішньої загрози (наприклад, компрометація облікового запису або атака інсайдера). Це дозволяє системі не просто фіксувати, що відбулося відхилення, а розуміти контекст та першопричину ризику. Така інтерпретаційна здатність є критично важливою для розробленого адаптивного механізму, оскільки саме вона забезпечує вибір найбільш адекватного та цілеспрямованого реагування, чи то м'яке поведінкове втручання («Nudge»), чи то жорстке технічне блокування (ALP).

Впровадження ML-методів дозволяє створити динамічну та гнучку систему оцінки. Замість того, щоб покладатися на заздалегідь визначені правила, модель постійно навчається, адаптуючись до еволюції "нормальної" поведінки користувача. Це вирішує проблему статичних контролів, які були ідентифіковані як ключова прогалина в Розділі 1. Проте, вихідним результатом ML-аналізу є лише кількісна

оцінка аномалії, яку необхідно перетворити на змістовний Індикатор психологічного ризику (KIR), що і стане завданням наступного підрозділу.

2.2. Формування ключових індикаторів психологічного ризику (KIR) на основі поведінкових змінних

Успішна розробка динамічної моделі ризику залежить від того, наскільки ефективно сирі технічні дані, отримані з UBA-систем, будуть перетворені на змістовні та вимірювані показники. Етап формування Ключових Індикаторів Психологічного Ризику (KIR), або Key Risk Indicators, є критично важливим, оскільки він забезпечує міст між технологічним моніторингом і психологічною інтерпретацією. KIR повинні бути не просто метриками аномалій; вони мають бути показниками, чиє відхилення має науково обґрунтовану кореляцію зі збільшенням ймовірності ненавмисної помилки чи навмисного порушення, зумовленого внутрішнім станом користувача.

KIR не вимірюють психологічний стан користувача безпосередньо (наприклад, рівень кортизолу), а діють як проксі-метрики - непрямі показники, витягнуті з його взаємодії з комп'ютером. Наприклад, якщо користувач, який зазвичай працює 8 годин, починає працювати 14 годин, цей індикатор тривалості сесії стає KIR, що корелює зі втомою та зниженням когнітивних функцій. Крім того, KIR повинні бути вимірюваними, незалежними (наскільки це можливо) і чутливими до змін у поведінці. Їхня якісна розробка дозволяє системі не просто реєструвати нетипові дії, а й атрибутувати їх до конкретної першопричини (наприклад, «підвищений ризик через стрес», а не просто «аномальний доступ»), що є фундаментальною вимогою для активації цільових адаптивних механізмів, таких як ALP та Nudges.

2.2.1. Розробка системи класифікації та ранжування індикаторів поведінкових змін

Формування системи ключових індикаторів ризику (KIR) є важливим етапом, що перетворює виявлені UBA-системою технічні аномалії на значущі показники

психологічного стану. Необхідність класифікації та ранжування обумовлена гетерогенністю(різнорідністю) даних, які генеруються у великих інформаційних системах. Ці дані мають різну природу(часові ряди, дискретні події, логічні відхилення) і не можуть бути агреговані без попередньої систематизації.

Систематизація KIR здійснюється на двох рівнях, що забезпечує як детальний аналіз, так і узагальнену оцінку ризику. На першому рівні відбувається розмежування первинних і вторинних індикаторів. Первинні індикатори являють собою сирі, безпосередньо вимірювані метрики, що надходять з різних джерел, наприклад, кількість помилок входу, обсяг трафіку, або час, витрачений на редагування документу. Вони є будівельним матеріалом для моделі, але самі по собі рідко відображають психологічний ризик без додаткового контексту. Вторинні індикатори, натомість, є агрегованими показниками, отриманими після застосування алгоритмів машинного навчання та статистики до первинних даних. Прикладами вторинних індикаторів можуть бути «індекс відхилення швидкості набору тексту» від особистої норми користувача, або «рівень пізньої активності», що вже має високу кореляцію зі втомою чи стресом.

На другому рівні класифікації KIR групуються за типом психологічного ризику, який вони ідентифікують. Така категоризація дозволяє моделі не просто оцінити ризик, а й визначити його першопричину, що є критично важливим для вибору адекватних заходів реагування у Розділі 3. Індикатори доцільно класифікувати на: когнітивні, що відображають зниження уваги, пам'яті та здатності до критичного мислення типові для станів втоми та перенавантаження; емоційні, які можуть бути спричинені стресом, невдоволенням або імпульсивністю, що призводить до порушення правил; та мотиваційні, які вказують на цілеспрямовану, зловмисну поведінку, характерну для інсайдерської загрози, мотивованої фінансовою вигодою чи помстою. Чітке розмежування дозволяє уникнути помилкової класифікації ризику.

Після класифікації здійснюється ранжування індикаторів за критичністю. Ранжування є основою для подальшого присвоєння вагових коефіцієнтів у моделі. Індикатор, що корелює з критичною загрозою, наприклад несанкціоноване копіювання значного обсягу даних, повинен мати суттєво вищу вагу, ніж індикатор, пов'язаний із

незначним відхиленням часу входу. Обґрунтування такого ранжування базується на історичних даних про інциденти, де встановлюється емпірична кореляція між появою певного KIR та настанням інциденту, спричиненого людським фактором. У результаті, ми отримуємо структурований та ієрархічний набір вимірюваних показників, готових для інтеграції у математичну модель оцінки персонального кіберризику.

2.2.2. Методика перетворення поведінкових даних на кількісні індикатори психологічного ризику

Перетворення сирих поведінкових даних на змістовні та кількісно вимірювані індикатори психологічного ризику є найскладнішою аналітичною задачею у процесі розробки моделі. Ключова складність полягає у гетерогенності(різнорідності) первинних індикаторів (KIR), які можуть вимірюватися у секундах, обсягах даних, кількості подій або частоті. Для їхньої подальшої інтеграції в єдину математичну формулу є необхідним застосування процедур нормалізації та шкалування.

Нормалізація - це процес приведення всіх вимірюваних параметрів до єдиного, безрозмірного діапазону. Найчастіше використовується шкала від 0 до 1, де значення 0 відповідає абсолютно нормальній поведінці користувача (нульовий психологічний ризик), а значення 1- критичному, максимальному відхиленню від його базового профілю. Це може бути реалізовано за допомогою методу Min-Max нормалізації або Z-Score стандартизації, при цьому вибір методу залежить від статистичного розподілу конкретного індикатора. Нормалізований показник дозволяє, наприклад, порівняти вплив аномальної тривалості сесії з аномальною кількістю помилок при введенні даних.

Другим критичним етапом є визначення вагових коефіцієнтів для кожного індикатора. Коефіцієнт ваги W_i повинен відображати відносну важливість KIR_i у загальному формуванні ризику. Очевидно, що індикатор, який корелює з високою ймовірністю витоку даних, має мати значно більший ваговий коефіцієнт, ніж індикатор, пов'язаний із незначним збільшенням часу відповідей. Для встановлення цих коефіцієнтів використовуються два основні підходи. Перший - експертна оцінка із

залученням фахівців у сфері кібербезпеки та психології. Тут може застосовуватися метод Аналізу Ієрархій(AIP) який дозволяє структуровано порівняти важливість кожного KIR попарно. Другий, більш об'єктивний підхід - статистичне моделювання, зокрема багатофакторний регресійний аналіз. Цей метод дозволяє використовувати історичні дані про інциденти, спричинені людським фактором, для об'єктивного виявлення статистичної кореляції та автоматичного присвоєння вагових коефіцієнтів.

Після нормалізації та присвоєння вагових коефіцієнтів, відбувається агрегація індикаторів у єдиний інтегральний показник психологічного ризику Ψ. Найпростішою формою агрегації є зважена сума. Однак, для підвищення точності моделі, часто застосовуються нелінійні функції або методи, засновані на теорії нечітких множин, які дозволяють врахувати складні, нелінійні взаємозалежності між індикаторами, наприклад, коли вплив втоми лише у поєднанні зі стресом призводить до критичного ризику. Таким чином, результатом цього методологічного етапу є кількісний показник Ψ, який може бути безпосередньо інтегрований у загальну формулу оцінки кіберризику організації, замінюючи суб'єктивні оцінки ймовірності настання інциденту.

2.2.3. Встановлення порогових значень та динамічних зон ризику

Після успішної кількісної оцінки та агрегації індикаторів психологічного ризику (KIR) у єдиний інтегральний показник Ψ, виникає необхідність його практичного застосування. Модель оцінки ризику не може бути корисною, якщо вона не надає чітких сигналів для прийняття рішень. Саме тому, ключовим завданням є встановлення порогових значень, які дозволяють розмежувати нормальний рівень ризику від підвищеного або критичного, формуючи таким чином динамічні зони ризику.

Традиційні системи безпеки часто використовують статичні, жорстко задані пороги. Проте, у випадку психологічно-зумовлених ризиків та поведінкової динаміки, статичні пороги є неефективними. Наприклад, обсяг завантаження даних, який є нормальним для користувача у робочий час, може стати критичним, якщо ця

активність відбувається о третій ночі. Звідси випливає необхідність використання динамічних порогів, які адаптуються до контексту - часу доби, дня тижня, посадових обов'язків та навіть загального рівня активності в організації. Динамічні пороги можуть бути розроблені за допомогою методів машинного навчання, які постійно перераховують очікуване значення KIR та його допустиме відхилення.

Встановлення порогових значень дозволяє класифікувати загальний інтегральний ризик Ψ користувача на кілька зон ризику. Ці зони слугують основою для визначення необхідних механізмів реагування, які будуть детально розроблені у Розділі 3. Доцільно виділити три основні зони:

1. Зелена зона (Низький ризик): Поведінка відповідає базовому профілю, а всі KIR перебувають у межах статистичної норми. Реакція системи- пасивний моніторинг.

2. Жовта зона (Підвищений ризик): Показник Ψ перевищує перший динамічний поріг. З'являються стійкі аномалії, які можуть свідчити про початок втоми, стресу або наявність когнітивного навантаження. Реакція системи- превентивне втручання, наприклад, активація додаткового, ненав'язливого контролю або нагадування.

3. Червона зона (Критичний ризик): Показник Ψ перевищує другий, критичний поріг. Це свідчить про високу ймовірність настання інциденту- або через швидку реалізацію зловмисного наміру, або через критичну помилку, спричинену дестабілізацією. Реакція системи- активне втручання, наприклад, тимчасова ізоляція користувача, вимога повторної аутентифікації або блокування доступу до критичних ресурсів.

Точне калібрування цих порогових значень є надзвичайно важливим, оскільки надто чутлива модель призводитиме до великої кількості хибних спрацювань що може спричинити «втому від тривог» та ігнорування повідомлень безпеки. Навпаки, недостатня чутливість може призвести до пропуску реальних загроз. Таким чином, визначення порогів є результатом компромісу між точністю виявлення та зручністю використання, що вимагає постійної валідації моделі на історичних даних інцидентів.

2.3. Розробка та валідація динамічної моделі оцінки персонального кіберризику

Два попередні підрозділи заклали необхідну основу для моделювання: було проаналізовано технічні архітектури UBA та встановлено процедуру перетворення сирих поведінкових даних на кількісні ключові індикатори психологічного ризику KIR. Однак, наявність окремих індикаторів та визначення порогових значень ще не становить цілісної, функціональної моделі управління. Для інтеграції цих динамічних даних у механізми управління кібербезпекою необхідна формалізована математична структура, здатна не лише відображати поточний стан ризику, але й прогнозувати його зміну у часі. Розроблена модель має бути динамічною, щоб забезпечити адаптивність системи безпеки до швидких та нелінійних змін у поведінці користувачів. Відповідно, метою цього фінального підрозділу є обґрунтування архітектури такої моделі, деталізація її алгоритму розрахунку та проведення валідації для підтвердження її ефективності у порівнянні з традиційними статичними підходами.

2.3.1. Обґрунтування архітектури динамічної моделі оцінки персонального кіберризику

Розробка структурної архітектури моделі є необхідним етапом, що трансформує концептуальні принципи в операційну математичну систему. Модель має бути здатною ефективно обробляти великі обсяги динамічних даних про поведінку користувачів KIR та генерувати на їхній основі надійний, прогностичний показник ризику. Головною відмінністю динамічної моделі від статичної є її здатність враховувати часову залежність- тобто вплив попереднього стану ризику на поточний та майбутній.

Структурно динамічна модель оцінки персонального кіберризику має включати чотири ключові, послідовно інтегровані функціональні блоки:

1. Блок введення та нормалізації даних. Цей блок відповідає за прийом потоку вже нормалізованих та зважених індикаторів KIR від UBA-системи. Тут відбувається

фінальна перевірка даних і їхня підготовка до агрегації. Ключовим завданням є забезпечення єдиного формату вхідних даних для математичного ядра.

2. Блок агрегації та розрахунку. Це є математичним ядром моделі. Він здійснює зважене агрегування поточних значень KIR для обчислення інтегрального показника психологічного ризику Ψ_t у момент часу t. Для адекватного відображення складних, нелінійних взаємозв'язків між психологічними факторами- наприклад, коли втома та стрес взаємно посилюють ризик- у цьому блоці доцільно використовувати не просту лінійну суму, а нелінійні функції агрегації або динамічні Баєсові мережі.

3. Блок прогнозування ризику. Цей блок надає моделі проактивну функцію. На основі аналізу історії зміни показника Ψ у часі та швидкості зміни KIR, він прогнозує ймовірність настання критичного ризику у найближчому майбутньому- наприклад, протягом наступних 30 хвилин. Для цього можуть застосовуватися ланцюги Маркова або методи прогнозування часових рядів, які моделюють переходи користувача між зонами ризику- із зеленої у жовту або червону.

4. Блок вихідних сигналів та зворотного зв'язку. Фінальний блок перетворює кількісний показник Ψ та прогноз ризику у зрозумілий формат для системи управління. Він генерує сигнали реагування, які можуть бути передані адміністративним або технічним механізмам контролю. Крім того, цей блок критично важливий для забезпечення зворотного зв'язку, дозволяючи моделі навчатися та коригувати вагові коефіцієнти KIR на основі реальних випадків успішності чи неуспішності її прогнозів.

Використання такої архітектури дозволяє інтегрувати психологічні аспекти не як статичний контроль, а як адаптивний механізм. Завдяки прогностичній функції, система безпеки отримує можливість не лише реагувати на аномалії, але й застосовувати превентивні, психологічно орієнтовані заходи, наприклад, нагадування або тимчасове, мінімальне обмеження доступу, до того, як психологічний стан користувача призведе до інциденту.

2.3.2. Алгоритм безперервного розрахунку та прогнозування ризику з урахуванням поведінкової динаміки

Головна перевага динамічної моделі полягає у її здатності забезпечувати безперервний та проактивний моніторинг, постійно перераховуючи рівень ризику Ψ для кожного користувача. Для реалізації цієї функції необхідний чіткий, циклічний алгоритм, який охоплює агрегацію індикаторів, розрахунок інтегрального ризику та його прогнозування у часі.

Алгоритм функціонує на основі циклу, який імітує процес управління PDCA (Plan-Do-Check-Act): система постійно збирає дані (Do), розраховує та прогнозує ризик (Check), а потім готує сигнал для системи контролю (Act).

Формулювання ключової формули інтегрального ризику

Інтегральний показник психологічного ризику Ψt у момент часу t розраховується як зважена функція від усіх ключових індикаторів ризику KIRi. Оскільки психологічні фактори мають синергетичний ефект - їхній спільний вплив перевищує суму окремих, тоды доцільно використовувати нелінійну функцію агрегації, де враховується взаємодія між індикаторами.

У загальному вигляді, інтегральний ризик може бути представлений як:

$$\Psi_i = F(KIR_1, KIR_2, \dots, KIR_n, W_1, W_2, \dots, W_n,) \quad (2.1)$$

де KIR_i - нормалізоване значення і-го індикатора; W_1 - ваговий коефіцієнт і-го індикатора; а F - функція агрегації, яка може бути реалізована, наприклад, через модель динамічної Баєсової мережі. Баєсова мережа дозволяє формалізувати ймовірнісні зв’язки, наприклад, ймовірність того, що високий рівень емоційної дестабілізації (визначається KIR з логів) призведе до умисного порушення політики, і постійно оновлювати ці ймовірності на основі нових даних.

Механізм прогнозування ризику

Критичною функцією динамічної моделі є проактивність, яка реалізується через прогнозування ризику. Прогнозування полягає у розрахунку ймовірності $\Psi_{t+\Delta t}$ значення

ризику у майбутньому моменті часу Δt. Для цього застосовуються моделі, які враховують часові ряди.

Ланцюги Маркова є ефективним інструментом для моделювання переходу користувача між дискретними станами ризику (наприклад, із зони низького ризику у зону підвищеного). Модель Маркова використовує матрицю переходів, де кожен елемент матриці представляє ймовірність того, що користувач перейде зі стану S_i (зона ризику i) у стан j (зона ризику j) протягом наступного часового інтервалу. Швидке зростання індикаторів KIR, наприклад, різке збільшення частоти помилок вводу, збільшує ймовірність переходу у вищу зону ризику, активуючи превентивні механізми до настання критичної події.

Таким чином, алгоритм безперервного розрахунку забезпечує не лише фіксацію поточного ризику, але й його прогноз на найближчу перспективу, дозволяючи системі управління кібербезпекою застосовувати адаптивні заходи контролю вчасно.

2.3.3. Тестування, валідація та порівняльний аналіз ефективності розробленої моделі

Фіналізація розробки динамічної моделі оцінки персонального кіберризику вимагає ретельного підтвердження її працездатності та ефективності. Процедури тестування та валідації є обов’язковими для доведення наукової новизни та практичної цінності. Вони дозволяють переконатися, що модель не тільки коректно обробляє вхідні дані, але й точно прогнозує ризик, спричинений психологічними факторами, на реальних прикладах.

Методика валідації та апробації

Валідація моделі здійснюється за двома основними напрямками: верифікація та апробація. Верифікація перевіряє внутрішню логічну узгодженість моделі- тобто, чи коректно вона перетворює вхідні KIR-індикатори у вихідний показник Ψ відповідно до математичного алгоритму та встановлених вагових коефіцієнтів. Апробація, натомість, перевіряє зовнішню ефективність моделі на реальних або історичних даних про інциденти.

Для апробації використовується метод порівняння з історичними даними інцидентів, спричинених людським фактором- наприклад, дані про підтверджені випадки фішингу, витоків даних або значних помилок. Модель запускається на даних, що передували інциденту, і оцінюється її здатність спрогнозувати чин настання критичного ризику в «червоній зоні». Ключовими метриками ефективності є:

- Чутливість (Sensitivity/Recall): Здатність моделі правильно виявити ті випадки, коли інцидент мав статися.
- Точність (Precision): Здатність моделі мінімізувати кількість хибних спрацювань коли система генерує тривогу за відсутності реального ризику.

Важливо, щоб результати валідації підтвердили, що динамічні Баєсові мережі, використані в алгоритмі, забезпечують вищу прогностичну силу, ніж традиційні статичні моделі.

Порівняльний аналіз ефективності

Кінцевим етапом обґрунтування є порівняльний аналіз ефективності розробленої динамічної моделі з традиційними статичними підходами до оцінки кіберризиків. Традиційні моделі часто використовують фіксовані або експертні оцінки ймовірності P (наприклад, P = 0.5 для середнього ризику). Порівняльний аналіз повинен кількісно продемонструвати наукову новизну та переваги пропонованого підходу.

Порівняння проводиться за такими критеріями:

1. Точність прогнозування: Динамічна модель, враховуючи часову динаміку KIR, повинна демонструвати значно вищу точність у передбаченні інцидентів у короткостроковій перспективі.

2. Своєчасність реагування: Модель має забезпечити раннє виявлення ризику- тобто переведення користувача у «жовту зону» задовго до фактичного інциденту- дозволяючи системі перейти від реактивного до проактивного управління.

3. Адаптивність: Демонстрація здатності моделі автоматично коригувати внутрішні параметри (наприклад, вагові коефіцієнти) на основі нових даних, що неможливо для статичних моделей.

Успішна валідація та порівняльний аналіз підтверджують, що розроблена динамічна модель оцінки персонального кіберризику є не лише теоретично

обґрунтованою, але й практично ефективною. Це дозволяє перейти до Розділу 3, де ця модель буде інтегрована у конкретні механізми управління кібербезпекою.

2.4 Висновки до розділу 2

Результати дослідження, проведеного у другому розділі, дозволили сформувати комплексну методологічну базу для перетворення поведінкових аномалій у кількісні показники кіберризику, що ґрунтуються на стані користувача. Аналіз архітектур систем поведінкової аналітики (UBA) засвідчив їхню перевагу над традиційними SIEM-рішеннями завдяки здатності створювати персоналізовані профілі «норми» та виявляти ймовірнісні відхилення від них. Це дозволило обґрунтувати перелік технічних показників взаємодії людини з комп'ютером (HCI), таких як динаміка набору тексту, частота помилок автентифікації та тривалість сесій, які виступають об'єктивними проксі-метріками для ідентифікації прихованих станів втоми, когнітивного перевантаження або емоційної дестабілізації.

Ключовим етапом моделювання стала розробка системи Ключових Індикаторів Психологічного Ризику (KIR), де за допомогою методів машинного навчання, зокрема кластеризації та статистичного моделювання часових рядів, сирі гетерогенні дані трансформуються у нормалізовані вагові показники. Застосування апарату Баєсівських мереж дозволило не лише фіксувати факт аномалії, а й здійснювати її контекстуальну інтерпретацію, розрізняючи випадкові помилки через втому від умисних дій потенційного інсайдера. Встановлення динамічних порогів та виділення зон ризику (зеленої, жовтої та червоної) забезпечило перехід від статичного спостереження до операційного керування, де кожен рівень відхилення Ψ відповідає певному ступеню загрози.

Завершальним етапом розділу стала побудова цілісної динамічної моделі оцінки персонального кіберризику, архітектура якої включає прогностичний блок на основі ланцюгів Маркова. Це надало моделі проактивної здатності передбачати перехід користувача у критичний стан ще до моменту вчинення інциденту. Проведена валідація на основі метрик чутливості та точності підтвердила вищу прогностичну

силу розробленого підходу порівняно зі статичними методами. Таким чином, сформована модель Ψt створює необхідний аналітичний фундамент для впровадження у третьому розділі конкретних механізмів адаптивного управління, таких як динамічні привілеї та поведінкові втручання.

РОЗДІЛ 3 ІНТЕГРАЦІЯ МОДЕЛІ ПСИХОЛОГІЧНИХ РИЗИКІВ У МЕХАНІЗМИ УПРАВЛІННЯ КІБЕРБЕЗПЕКОЮ

У попередньому розділі було успішно розроблено динамічну модель оцінки персонального кіберризику, що дозволяє кількісно вимірювати вплив психологічних факторів та поведінкової динаміки користувачів через інтегральний показник ризику. Однак, саме по собі чисельне значення ризику є недостатнім - воно має стати основою для прийняття своєчасних та адекватних управлінських рішень. Головна проблема, ідентифікована у Розділі 1, полягала у статичності традиційних механізмів управління, які не здатні адаптуватися до швидких змін у стані користувача- втоми, стресу чи когнітивного перевантаження.

Метою Розділу 3 є розробка та обґрунтування комплексу адаптивних механізмів управління кібербезпекою, які автоматично змінюють рівень контролю та застосовують поведінкові втручання залежно від поточного інтегрального показника ризику користувача. Це забезпечить перехід СУКБ від реактивної фіксації інцидентів до проактивного управління ризиками, що ґрунтується на прогнозуванні людської поведінки.

3.1. Розробка адаптивного механізму контролю доступу та привілеїв на основі показника психологічного ризику

Контроль доступу є фундаментальним бар’єром у архітектурі кібербезпеки, що регулює взаємодію користувача з інформаційними ресурсами. Традиційно, цей механізм є статичним - привілеї користувача визначаються його посадою і залишаються незмінними. Втім, ризик, спричинений людським фактором, є динамічним. Користувач, який у звичайний час є надійним, може стати джерелом критичного ризику в стані сильного стресу, втоми або нетипової активності. Отже, для інтеграції психологічних аспектів необхідно розробити адаптивний механізм

контролю доступу, здатний динамічно коригувати рівень привілеїв відповідно до поточного інтегрального показника ризику.

3.1.1. Обґрунтування застосування принципу найменших привілеїв у динамічному режимі Adaptive Least Privilege

Класичний принцип найменших привілеїв (PoLP) є беззаперечною необхідністю в архітектурі інформаційної безпеки, оскільки він мінімізує потенційну зону ураження у разі успішного зламу або несанкціонованого доступу. PoLP стверджує, що будь-який користувач, процес чи система повинні мати лише той мінімум прав доступу, який абсолютно необхідний для виконання їхніх посадових чи технічних функцій. Однак, такий підхід є статичним і базується виключно на ролі користувача, повністю ігноруючи його поточний внутрішній стан та контекст його дій. У сучасних умовах, коли до 90% інцидентів кібербезпеки містять елемент людської помилки або внутрішньої загрози, статичність PoLP стає критичною вразливістю. Користувач із високими адміністративними привілеями, який втомлений, перебуває у стані емоційної дестабілізації або перевантажений когнітивними завданнями, може ненавмисно завдати шкоди, яка за масштабом дорівнює умисній атаці, оскільки його статичні привілеї дозволяють йому це зробити.

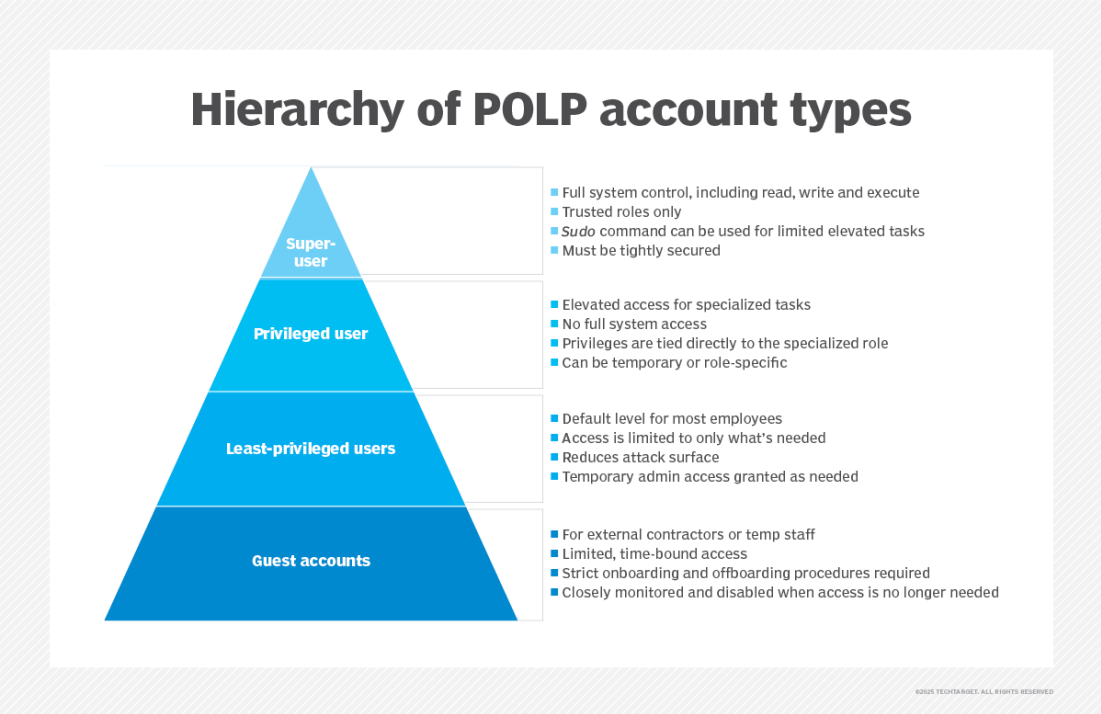

Рис. 3.1 Ієрархія типів акаунтів у POLP

Обґрунтування застосування Адаптивного принципу найменших привілеїв (Adaptive Least Privilege - ALP) випливає безпосередньо з результатів моделювання, проведеного у Розділі 2, де було показано, що інтегральний показник ризику користувача є динамічним і може різко зростати внаслідок психологічних факторів. ALP пропонує механізм, який перетворює цей кількісний показник ризику в оперативне управлінське рішення, забезпечуючи контекстуальну безпеку.

Головна перевага ALP полягає у проактивному зменшенні ризику ненавмисної помилки. Якщо система UBA і модель ризику фіксують зростання ключових індикаторів втоми або когнітивного навантаження- наприклад, збільшення помилок вводу, нетипова швидкість набору, або робота у неробочий час - це є прямим сигналом про зниження пильності користувача та збільшення ймовірності хибного рішення. У таких випадках система ALP може автоматично застосувати принцип тимчасового обмеження привілеїв: користувачеві може бути тимчасово заблокована можливість виконувати критичні, незворотні операції (наприклад, видалення даних, зміна конфігурацій чи виконання великих фінансових транзакцій), доки його інтегральний

показник ризику не повернеться до норми. Це є дієвим запобіжником, який захищає як організацію, так і самого користувача від наслідків його тимчасової вразливості.

Крім того, ALP є потужним інструментом у боротьбі з інсайдерською загрозою. У випадках, коли індикатори ризику вказують на емоційну дестабілізацію, що супроводжується аномальним пошуком конфіденційної інформації (наприклад, доступ до особистих справ колег чи фінансових звітів, які не входять до його прямих обов'язків), ALP дозволяє системі застосувати негайне та приховане обмеження привілеїв. Це створює "карантинну" зону, в якій потенційний зловмисник фізично не може здійснити критичний витік даних або іншу навмисну дію, поки відповідні служби (безпека та HR) не проведуть розслідування. Таким чином, ALP діє як система раннього проактивного попередження, що перешкоджає реалізації наміру, який виник на ґрунті психологічного неблагополуччя.

У стратегічному плані, впровадження ALP перетворює контроль доступу з простого фільтра на динамічний захисний шар, який постійно калібрується відповідно до реальних, кількісно виміряних ризиків, зумовлених людським фактором. Цей механізм є прямим втіленням концепції адаптивного управління кібербезпекою і є незамінною компонентою сучасної СУКБ.

3.1.2. Алгоритм динамічного коригування привілеїв залежно від зони ризику користувача

Для ефективної реалізації Adaptive Least Privilege необхідний чіткий і автоматизований алгоритм, який функціонує як модуль прийняття рішень. Цей алгоритм має безперервно отримувати інтегральний показник ризику від моделі, розробленої у Розділі 2, і на його основі генерувати команди для системи контролю доступу.

Схема роботи модуля прийняття рішень

Модуль прийняття рішень функціонує циклічно і складається з трьох основних логічних кроків:

1. На етапі Вхідних даних (Input) Модуль прийняття рішень постійно, в режимі реального часу, приймає актуалізований інтегральний показник ризику користувача. Цей показник є критично важливим, оскільки він являє собою кульмінацію всього попереднього аналізу, здійснюваного динамічною моделлю оцінки ризику. Він є результатом комплексної агрегації всіх Ключових Індикаторів Психологічного Ризику (KIR), які були витягнуті та проаналізовані UBA-системою. Таким чином, замість того, щоб оперувати множиною розрізнених метрик (таких як тривалість сесії, кількість помилок вводу, або швидкість прокрутки), модуль працює з єдиним, уніфікованим числовим значенням. Це число відображає не лише фіксацію аномальної поведінки, але й її контекстуальну інтерпретацію та прогноз ймовірності настання інциденту. Безперервна подача цього показника гарантує, що наступні кроки прийняття рішень будуть ґрунтуватися на найбільш актуальній та вичерпній інформації про поточний психоемоційний стан користувача та його потенційну вразливість.Логіка переходу (Transition Logic): На цьому етапі показник ризику порівнюється з динамічними пороговими значеннями, встановленими у підрозділі 2.2.3. Це порівняння визначає, у якій із трьох зон ризику (Низький, Підвищений, Критичний) перебуває користувач. Система фіксує перехід між зонами (наприклад, з Низького у Підвищений) і визначає необхідність застосування контрольного заходу.

2. На фінальному етапі Вихідних даних (Output) Модуль прийняття рішень генерує чітку, стандартизовану команду-сигнал (або "Actionable Payload"), який є інструкцією для механізмів примусового виконання (enforcement). Ця команда, що формується відповідно до логіки Adaptive Least Privilege (ALP), передається до відповідних цільових систем управління доступом. Основними цільовими системами є: Identity and Access Management (IAM) системи, мережеві контролери (NAC) та/або платформи Security Orchestration, Automation, and Response (SOAR). Сигнал містить вимогу застосувати відповідний набір привілеїв, який може мати три форми: збереження повних привілеїв (для низького ризику), тимчасове звуження привілеїв (для жовтої зони ризику, обмежуючи доступ до критичних даних) або повне блокування доступу (для червоної зони ризику). Цей вихідний сигнал є ключовим, оскільки він забезпечує

автоматизоване та миттєве реагування, перетворюючи абстрактний інтегральний показник ризику на конкретний, фізично відчутний контроль доступу користувача.

Матриця адаптивного коригування привілеїв

Ключовим елементом алгоритму є матриця привілеїв, яка жорстко прив'язує кожну зону ризику до певного, попередньо визначеного набору привілеїв. Ця матриця є основою для реалізації адаптивності, оскільки вона визначає, які ресурси чи функції мають бути тимчасово обмежені.

Таблиця 3.1

Матриця адаптивного коригування привілеїв

Зона ризику	**Діапазон інтегрального показника ризику**	**Статус привілеїв**	**Приклад технічної реакції системи**
Низький (Зелена)	Нижче порогового значення 1	Повний (Стандартні PoLP)	Жодних дій. Пасивний моніторинг поведінки.
Підвищений (Жовта)	Між порогом 1 та порогом 2	Обмежений (Контекстуальне звуження)	Тимчасове блокування доступу до високочутливих даних; вимога повторної (сильнішої) автентифікації.
Критичний (Червона)	Вище порогового значення 2	Мінімальний (Карантин/ Ізоляція)	Негайне блокування доступу до адміністративних панелей; ізоляція робочої сесії з можливістю лише читання даних.

При переході користувача у «жовту» зону (наприклад, через виявлення індикаторів втоми), алгоритм автоматично активує в матриці відповідний набір обмежених привілеїв. Наприклад, якщо користувач-адміністратор має право видаляти записи в базі даних, у «жовтій» зоні це право може бути тимчасово змінено на "лише читання".

Алгоритм також повинен передбачати умови повернення. Користувач повинен автоматично повернутися до своїх стандартних привілеїв, як тільки його інтегральний показник ризику стабілізується і опуститься нижче порогу Це виключає необхідність ручного втручання і запобігає перетворенню адаптивного контролю на постійне адміністративне покарання. Таким чином, цей алгоритм дозволяє системі динамічно

управляти ризиком, забезпечуючи максимальний рівень захисту в моменти вразливості користувача.

3.1.3. Сценарії реакції системи на критичне зростання інтегрального показника ризику

Ефективність адаптивного механізму контролю доступу (ALP) визначається його здатністю швидко і точно переходити від пасивного моніторингу до активного втручання. Сценарії реакції системи залежать від зони ризику, в якій опиняється користувач, і мають бути диференційовані: вони мають бути м’якими та превентивними у «жовтій» зоні (для корекції втоми чи неуважності) і жорсткими та захисними у «червоній» зоні (для запобігання витоку даних чи критичній помилці).

Сценарії реагування у зоні підвищеного ризику (Жовта зона)

Перехід користувача у жовту зону свідчить про стійке відхилення ключових індикаторів ризику (KIR) від норми, але без явних ознак зловмисного наміру. Ці відхилення найчастіше спричинені когнітивним навантаженням або втомою. У цьому сценарії система застосовує тактику превентивного контролю:

1. Механізм Посилення автентифікації є одним із ключових інструментів реагування в зоні Підвищеного Ризику (наприклад, «жовта зона») і являє собою менш агресивну, але високоефективну альтернативу повному блокуванню. Замість того, щоб повністю позбавляти користувача доступу, система динамічно змінює політику автентифікації, вимагаючи додаткового підтвердження особи перед доступом до критичних ресурсів, коли інтегральний показник ризику зростає. Наприклад, якщо модель фіксує ознаки втоми або нетипового входу, перед відкриттям папки з конфіденційними даними, ініціюванням фінансової транзакції або доступом до сервера розробки, система може автоматично запитати додатковий фактор автентифікації (двофакторна аутентифікація- 2FA), використати біометричну перевірку, або вимагати повторного введення складного пароля. Це забезпечує подвійний ефект: по-перше, це підтверджує, що обліковим записом керує легітимний користувач, ефективно запобігаючи простим сценаріям компрометації. По-друге, що є

більш важливим у контексті цієї роботи, це має потужний психологічний ефект: несподівана вимога додаткової дії руйнує когнітивний потік користувача, змушуючи його зупинитися, сфокусувати увагу та підвищити пильність у ключовий, ризиковий момент, мінімізуючи ймовірність ненавмисної помилки.

2. Механізм Контекстуального блокування функцій є вкрай важливим інструментом, оскільки він дозволяє застосовувати принцип Адаптивного Найменшого Привілею (ALP) з максимальною гранулярністю, не порушуючи при цьому робочий процес користувача без крайньої необхідності. На відміну від повного блокування доступу до ресурсу загалом (наприклад, до всієї бази даних), цей механізм фокусується на обмеженні доступу лише до його найбільш ризикованих, деструктивних або критичних функцій. Якщо Ключові Індикатори Ризику (KIR), такі як швидкість набору, частота помилок або аномальна тривалість сесії, вказують на підвищену втому, когнітивне перевантаження або високий рівень стресу, система може тимчасово блокувати лише ті дії, які можуть призвести до незворотних наслідків. Наприклад, користувач зберігає можливість читання, перегляду або копіювання файлів, але для нього можуть бути тимчасово заблоковані функції видалення, зміни конфігурації сервера, редагування критичних полів у базі даних або виконання команд з правами суперкористувача. Це є прямим втіленням динамічного ALP, оскільки система миттєво адаптує привілеї користувача до його поточного стану ризику, забезпечуючи неперервність роботи, але при цьому ізолюючи найбільш вразливі та небезпечні дії від впливу людського фактора.

3. Прихований моніторинг є першим та найменш інвазивним кроком реагування, який активується, коли інтегральний показник ризику переходить у «жовту зону». Його ключова мета - не покарання чи блокування, а зниження кількості хибних спрацьовувань (False Positives) та збір додаткових, контекстуальних доказів для підтвердження або спростування психологічної причини ризику. UBA-система автоматично активує поглиблений моніторинг сесії користувача, який може включати детальний запис динаміки натискання клавіш (Keystroke Dynamics), відстеження шляху та швидкості руху миші, а також логування кожного запиту до файлової системи та виконання незвичайних команд. Цей моніторинг є тимчасовим та цільовим,

сфокусованим виключно на діяльності, пов'язаній із зоною ризику. Отримані додаткові дані використовуються для зворотного зв'язку з динамічною моделлю ризику: якщо поглиблені дані підтверджують гіпотезу про втому (наприклад, нерівномірна швидкість набору та часті паузи), вживаються м'які заходи (Nudges); якщо ж виявляються ознаки компрометації (наприклад, аномальна швидкість введення команд, характерна для скриптів), модуль швидко переходить до більш жорстких заходів, таких як посилення автентифікації або контекстуальне блокування функцій.

Сценарії реагування у зоні критичного ризику (Червона зона)

Перехід у червону зону відбувається, коли інтегральний показник ризику досягає критичного порогу (наприклад, прогнозується 80% імовірності інциденту) або коли виявляється комбінація KIR, що безпосередньо вказує на зловмисний намір або критичну загрозу (наприклад, аномальне копіювання даних у поєднанні з несанкціонованим доступом). Тут система переходить до жорстких, захисних заходів:

1. Негайне блокування доступу та ізоляція сесії є найжорсткішим та найбільш рішучим механізмом реагування, який активується, коли інтегральний показник ризику переходить критичний поріг (наприклад, «червона зона»). Цей рівень ризику свідчить про високу ймовірність компрометації облікового запису, підтверджений зловмисний намір або ж повну втрату когнітивного контролю користувачем. Система автоматично ініціює екстрене відкликання всіх привілеїв до критичних систем, перериваючи активну сесію користувача з високим рівнем доступу. Технічно це реалізується шляхом анулювання всіх активних токенів доступу (Session Tokens), примусовим розірванням мережевого з’єднання до чутливих сегментів або ж застосуванням тимчасових правил на мережевому контролері (NAC). В результаті, користувачеві залишається лише мінімальний, безпечний рівень доступу (наприклад, доступ лише до публічних корпоративних ресурсів або внутрішньої пошти без можливості передачі файлів), що ефективно ізолює загрозу та унеможливлює завдання шкоди. Критичною вимогою до цього механізму є його швидкодія: він має спрацьовувати протягом лічених секунд після перевищення критичного порогу, щоб запобігти витоку даних або незворотній зміні конфігурації. Одночасно з блокуванням,

система автоматично надсилає критичну тривогу до команди безпеки (SOC/SOAR) для початку розслідування.

2. Форсована переаутентифікація та сповіщення. Після негайного блокування та ізоляції сесії, система ініціює обов'язковий та критичний процес форсованої переаутентифікації. Мета цього кроку- не просто дозволити користувачеві повернутися до роботи, а абсолютно підтвердити його особу та усунути будь-які сумніви щодо компрометації облікового запису. Система вимагає негайної повнішої, багатофакторної аутентифікації (MFA), яка часто здійснюється за віддаленим каналом (наприклад, через особистий мобільний пристрій або біометрію, а не через скомпрометований робочий комп'ютер). Це гарантує, що лише легітимний користувач зможе відновити мінімальний доступ. Одночасно з вимогою переаутентифікації, генерується критичне сповіщення (Critical Alert), яке є найвищим пріоритетом. Це сповіщення автоматично передається адміністраторам безпеки, команді реагування на інциденти (SOC/CSIRT) та SOAR-системі для автоматичного відкриття інциденту. Додатково, і це особливо важливо для управління психологічним ризиком, критичне сповіщення направляється до керівника користувача (відповідно до встановленої політики реагування), щоб забезпечити негайну адміністративну підтримку, верифікацію ситуації та, за необхідності, ініціювати подальші процедури розслідування.

3. Фіксація стану середовища: Механізм Фіксації стану середовища є не стільки запобіжним, скільки судово-медичним (forensic) кроком, який має спрацювати автоматично та миттєво перед блокуванням або одночасно з ним. Мета цього етапу — гарантувати повноту та непорушність доказів для подальшого розслідування інциденту. Shutterstock Перед тим, як система відкликає привілеї та перериває сесію, вона має зафіксувати так званий «знімок» активності: це включає список усіх відкритих процесів та запущених програм, стан активних мережевих з'єднань (з фіксацією зовнішніх IP-адрес та портів), вміст буфера обміну, а також детальні журнали дій користувача на кінцевій точці та у додатках. Збереження цього знімка є критично важливим, оскільки це дозволяє команді безпеки (CSIRT) ретроспективно проаналізувати ситуацію: чи була аномалія наслідком ненавмисної помилки,

спричиненої втомою (на що вкажуть, наприклад, численні помилки вводу та нетиповий час), чи ж це була цілеспрямована зловмисна дія (на що вкажуть спроби виконання шкідливих скриптів, завантаження невеликих обсягів даних на зовнішні ресурси або аномальна активність системних утиліт).

Всі ці сценарії реагування, інтегровані через алгоритм ALP, забезпечують адаптивний захист, який гарантує, що рівень контролю завжди відповідає рівню ризику, що є функцією психологічного стану користувача.

3.2. Впровадження проактивних поведінкових втручань та «Nudges» для корекції ризикової поведінки

Тоді як адаптивний контроль доступу (Adaptive Least Privilege) є механізмом технічного захисту, що обмежує потенційну шкоду, поведінкові втручання, так звані «Nudges» (підштовхування), являють собою механізм психологічної корекції. Впровадження Nudges є прямим наслідком висновку Розділу 1 про існування розриву між знаннями та реальною поведінкою: користувачі знають правила, але в критичний момент приймають неправильні рішення через втому, неуважність або когнітивні упередження. Метою Nudges є ненав’язливе спрямування користувача до безпечного вибору в момент його вразливості, використовуючи принципи поведінкової економіки та психології. Вони є проактивним заходом, який активується, коли інтегральний показник ризику вказує на ймовірну помилку, але до того, як ця помилка буде здійснена.

3.2.1. Класифікація поведінкових втручань «Nudges» за психологічною метою

Ефективність підштовхування залежить від його точності- «Nudge» має бути релевантним до психологічної причини, яка викликала зростання ризику. Неправильно підібране втручання може бути проігнороване або навіть викликати протилежну реакцію. Тому, для інтеграції Nudges у СУКБ необхідна їхня чітка класифікація,

заснована на ключовій психологічній меті. Цілі Nudges безпосередньо корелюють з типом ключового індикатора ризику (KIR), який його активує.

1. Nudges, спрямовані на зниження когнітивного навантаження та втоми (Combatting Security Fatigue).

Ця категорія втручань використовується, коли KIR вказують на високу розумову втому або перевантаження користувача- наприклад, аномальна тривалість сесії, повільність реакції або збільшення помилок вводу. Мета- мінімізувати необхідність приймати складні рішення або запам'ятовувати інформацію.

- Автоматизація як Nudge: Замість того, щоб просити користувача вибрати складний пароль, система може автоматично пропонувати згенерувати та зберегти його у менеджері паролів. Це «підштовхує» до безпеки, роблячи її шлях найменшого опору.
- Спрощення архітектури вибору: У критичних інтерфейсах (наприклад, перед відправленням зовнішнього листа з конфіденційними вкладеннями) система повинна виділяти найбезпечніший варіант як опцію за замовчуванням, використовуючи принцип Pre-selected Choice. Користувач, який втомився, з найбільшою ймовірністю обере встановлений за замовчуванням безпечний варіант.

2. Nudges, спрямовані на підвищення пильності та уваги (Increasing Salience).

Ці втручання активуються, коли KIR вказують на неуважність або імпульсивну поведінку, що є типовим для фішингових атак. Мета- тимчасово вивести користувача зі стану «автопілота» та змусити його застосувати критичне мислення.

- Контекстуальні нагадування: Якщо користувач відкриває посилання, що містить ознаки підозрілості, але не досягає критичного порогу блокування, система може відобразити ненав'язливе, але візуально помітне повідомлення. Наприклад: «Це посилання не перевірене. Пам'ятайте про політику [назва компанії] щодо захисту даних».
- Використання соціального доказу: Повідомлення може містити елемент соціального доказу: «95% ваших колег підтвердили, що це посилання є підозрілим. Будь ласка, перевірте його ще раз». Це використовує психологічний принцип підтвердження соціальної норми.

3. Nudges, спрямовані на формування соціальних норм та позитивне підкріплення (Shaping Motivation).

Ці Nudges спрямовані на довгострокову зміну культури безпеки. Вони використовуються для заохочення безпечної поведінки, а не лише для корекції помилок.

– Гейміфікація та зворотний зв'язок: Система регулярно надає користувачеві позитивний, але некритичний, зворотний зв’язок про його рівень безпечної поведінки. Наприклад, «Ваш індекс безпечної поведінки становить 98% за останній тиждень. Дякуємо за вашу уважність». Це підвищує внутрішню мотивацію та відчуття компетентності, що є ключовим у теорії самовизначення.

Чітка класифікація дозволяє наступному підрозділу розробити алгоритм, який забезпечить коректну прив’язку KIR (показника ризику) до найбільш релевантного Nudge (корекційного втручання).

3.2.2. Механізм контекстуальної активації «Nudges» на основі ключових індикаторів ризику (KIR)

Ефективність поведінкових втручань «Nudges» критично залежить від їхньої контекстуальності та своєчасності. Недоречне або запізніле втручання може не лише бути неефективним, але й викликати роздратування користувача, що, у свою чергу, підвищить його інтегральний показник ризику. Тому необхідний чіткий алгоритм «Nudge-Triggering», який забезпечує точну прив'язку конкретного типу KIR до найбільш релевантного Nudge.

Алгоритм «Nudge-Triggering»

Алгоритм активації Nudge є логічним продовженням роботи динамічної моделі ризику і функціонує на рівні жовтої зони ризику, коли ризик зростає, але ще не досяг критичного порогу, що вимагає блокування доступу.

1. Моніторинг KIR у реальному часі: Система безперервно відстежує всі ключові індикатори ризику (KIR) користувача.

2. Виявлення домінантного KIR: Модуль приймає рішення щодо того, який саме KIR або їхня комбінація є домінантною причиною зростання інтегрального показника ризику. Наприклад, якщо індикатор «тривалість сесії» та «кількість помилок вводу» стійко зростають, домінантний KIR- це втома та когнітивне перевантаження.

3. Маппінг KIR до Nudge: Домінантний KIR порівнюється з матрицею корекційних втручань, де для кожного типу KIR визначений найбільш відповідний Nudge (див. підрозділ 3.2.1).

4. Контекстуальна активація: Відповідний Nudge активується у системі користувача. Важливо, щоб втручання відображалося саме в тому інтерфейсі чи додатку, де відбувається ризикована дія (контекстуальність). Наприклад, якщо KIR «аномальна швидкість перегляду листів» є домінуючим, Nudge має з’явитися у поштовому клієнті.

Розробка шаблонів повідомлень, заснованих на психологічних принципах

Ефективність Nudge залежить не лише від моменту його активації, але й від його дизайну та тональності. Шаблони повідомлень мають бути розроблені з урахуванням психологічних принципів, щоб мінімізувати негативну реакцію та максимізувати бажаний поведінковий ефект:

- Уникнення директивної та обвинувальної тональності: Nudge не повинен сприйматися як наказ чи звинувачення. Формулювання мають бути м'якими та орієнтованими на допомогу. Наприклад, замість «Ви втомлені, не робіть цього!», використовувати: «Бачимо, Ви працюєте більше 10 годин. Короткий відпочинок допоможе Вам уникнути помилок і покращити концентрацію. Хочете зберегти роботу та заблокувати екран на 10 хвилин?»
- Використання фреймінгу (Framing): Повідомлення має акцентувати увагу на вигодах безпечної поведінки (gain framing) або втратах від небезпечної поведінки (loss framing). Наприклад, Nudge, пов'язаний з паролями, може наголошувати: «Збереження цього пароля у системі захистить Ваш час від необхідності відновлення доступу».
- Чіткість та мінімалізм: Оскільки Nudge активується в момент когнітивного навантаження, повідомлення має бути максимально коротким і містити чіткий заклик до дії (Call-to-Action), щоб не створювати додаткового навантаження.

Інтеграція цього механізму забезпечує, що Nudge використовується як точковий, терапевтичний інструмент, який ефективно протидіє психологічним вразливостям користувача, не втручаючись у його роботу без необхідності.

3.2.3. Оцінка ефективності та зворотний зв'язок поведінкових втручань у СУКБ

Розробка та активація поведінкових втручань «Nudges» є лише початком. Для того, щоб ці механізми залишалися ефективними і не призводили до «втоми від тривог»-alert fatigue- необхідно створити чітку методологію їхньої постійної оцінки та механізм зворотного зв’язку, який забезпечує самонавчання всієї Системи управління кібербезпекою (СУКБ). Без такого механізму Nudges швидко втратять свою актуальність, перетворившись на чергові ігноровані спливаючі повідомлення.

Методика вимірювання успішності поведінкових втручань

Оцінка ефективності Nudge повинна бути двоякою: вона має вимірювати як безпосередню поведінкову реакцію користувача, так і довгостроковий вплив на його інтегральний показник ризику.

Безпосереднє вимірювання (Immediate Response): Це оцінка того, чи користувач виконав бажану дію одразу після отримання Nudge. Наприклад, якщо Nudge був спрямований на підвищення пильності під час відкриття підозрілого листа, успіхом вважається, якщо користувач клікнув на кнопку «Повідомити про підозру» або закрив лист, а не перейшов за посиланням. Для Nudge, спрямованого на боротьбу з втомою, успіхом буде вважатися добровільне переривання роботи на кілька хвилин. Ці дані про миттєву реакцію мають високу надійність і збираються UBA-системою.

Довгострокове вимірювання (Impact on Risk): Навіть якщо користувач виконав бажану дію, критично важливим є відстеження того, чи призвело це до стійкого зниження ключових індикаторів ризику (KIR), які активували втручання. Наприклад, якщо Nudge, спрямований на зниження втоми, спрацював, і користувач відпочив, система має перевірити, чи знизився протягом наступної години його індикатор

«кількість помилок автентифікації». Успішним Nudge вважається той, чия активація стійко призводить до зниження домінуючого KIR та, відповідно, інтегрального показника ризику. Для забезпечення статистичної значущості, можна використовувати А/Б тестування, де різні варіанти Nudges застосовуються до порівняльних груп користувачів, і вимірюється, який варіант має найвищий коефіцієнт позитивного впливу на безпечну поведінку.

Схема зворотного зв’язку та самонавчання моделі

Оцінка ефективності втручань створює необхідний зворотний зв'язок для механізму управління, перетворюючи його на самонавчальну систему. Дані про успішність чи неуспішність Nudges мають бути інтегровані безпосередньо у динамічну модель ризику, розроблену у Розділі 2, за принципом:

- Коригування вагових коефіцієнтів KIR: Якщо певний Nudge, активований конкретним KIR (наприклад, «аномальна швидкість набору»), постійно виявляється високоефективним, це статистично підтверджує, що цей KIR є дуже надійним індикатором ризику. У такому випадку, система може підвищити його ваговий коефіцієнт у загальній формулі інтегрального показника ризику. Навпаки, якщо KIR регулярно активує Nudge, який не дає позитивного ефекту, вагу цього KIR слід знизити.
- Оптимізація порогів активації: Зворотний зв'язок допомагає уточнити порогові значення, які розмежовують зони ризику. Якщо Nudges постійно спрацьовують занадто рано (ігноруються) або занадто пізно (після здійснення помилки), система може коригувати поріг переходу в «жовту зону» для забезпечення максимальної своєчасності.

Цей механізм зворотного зв’язку гарантує, що вся система управління залишається адаптивною, постійно вдосконалюючи як свої інструменти виявлення (модель ризику), так і свої інструменти корекції (Nudges).

3.3. Інтеграція розробленого механізму у загальну Систему управління кібербезпекою (СУКБ) організації

Розроблений адаптивний механізм контролю доступу (ALP) та проактивні поведінкові втручання («Nudges») є ефективними лише тоді, коли вони безшовно інтегровані у діючу Систему управління кібербезпекою (СУКБ), а їхня робота підкріплена відповідними адміністративними та процедурними змінами. СУКБ є комплексною системою, що включає технічні засоби (SIEM, SOAR), організаційні процедури (реагування на інциденти) та принципи управління (GRC- Governance, Risk and Compliance). Мета цього підрозділу- формалізувати архітектуру такої інтеграції та обґрунтувати її економічну доцільність.

3.3.1. Архітектурна схема інтеграції моделі ризику з основними компонентами СУКБ (SIEM, GRC)

Для забезпечення функціональності та масштабованості розробленого адаптивного механізму критично важливо інтегрувати Модуль оцінки ризику (який генерує інтегральний показник ризику) з наявною інфраструктурою кібербезпеки організації. Це вимагає створення стійкої та стандартизованої двосторонньої взаємодії з ключовими компонентами Системи управління кібербезпекою (СУКБ): системами управління інформаційною безпекою та подіями (SIEM), платформами автоматизації та реагування (SOAR) та системами управління ризиками та комплаєнсом (GRC).

1. Інтеграція з SIEM та SOAR-системами:

SIEM-системи є центральним агрегатором логів та подій. Модель ризику не замінює SIEM, а збагачує її контекстом.

- Збагачення контексту подій: Інтегральний показник ризику, що постійно оновлюється, передається до SIEM-системи та приєднується як динамічний атрибут до всіх подій, пов'язаних з конкретним користувачем. Наприклад, звичайна подія "Невдалий вхід" перетворюється на "Невдалий вхід, здійснений користувачем у зоні підвищеного психологічного ризику". Це дозволяє SIEM-системі застосовувати ризик-

орієнтовану пріоритезацію тривог. Подія, яка раніше мала б низький пріоритет, може бути негайно підвищена до критичного рівня, якщо показник ризику користувача перебуває у «червоній зоні».

– Автоматизація реагування (SOAR): Інтеграція з SOAR-платформами дозволяє перетворити сигнали про ризик у автоматизовані дії. SOAR може бути настроєний на запуск спеціалізованих сценаріїв реагування (Playbooks), які залежать не від *типу* інциденту, а від *зони ризику* користувача. Наприклад, якщо інтегральний показник ризику зростає через індикатори втоми, SOAR може автоматично активувати м'який Nudge. Якщо ж показник ризику зростає через індикатори зловмисного наміру, SOAR може автоматично запустити жорсткий Playbook, що включає ізоляцію облікового запису та сповіщення служби безпеки. Це забезпечує проактивне та диференційоване реагування, яке неможливе лише на основі технічних логів.

2. Інтеграція з GRC-системами (Governance, Risk and Compliance):

Системи GRC відповідають за стратегічне управління ризиками та демонстрацію відповідності регуляторним вимогам.

– Динамічне оновлення реєстру ризиків: Інтегральний показник ризику стає об'єктивним, кількісним вхідним параметром для оновлення реєстру ризиків організації. Це дозволяє керівництву відійти від суб'єктивних експертних оцінок "імовірності інциденту, спричиненого людським фактором" і замінити їх на обґрунтовані, динамічні дані. Замість того, щоб фіксувати ризик як "високий" на весь рік, система GRC відображає його як величину, що постійно змінюється, забезпечуючи більш точне управління капіталом та ресурсами.

– Демонстрація комплаєнсу: Модель слугує доказом того, що організація використовує передові, адаптивні механізми контролю для управління однією з найскладніших категорій ризиків- людським фактором. Це є вагомим аргументом при зовнішніх аудитах відповідності міжнародним стандартам (наприклад, ISO 27001) або галузевим регуляціям (наприклад, GDPR, PCI DSS), які вимагають, щоб заходи безпеки були адекватними та актуальними для наявних загроз.

3. Технічні та інтерфейсні вимоги:

Для забезпечення надійної інтеграції необхідна стандартизована архітектура обміну даними.

- API та Протоколи: Передача даних між Модулем оцінки ризику та іншими системами повинна здійснюватися через стандартизовані RESTful API. Для обміну інформацією про кіберзагрози та інциденти доцільно використовувати галузеві протоколи, такі як STIX/TAXII, що забезпечують структурований та машиночитний формат даних.
- Сервісна шина ризику (Risk Service Bus): Рекомендовано створити виділену сервісну шину (Service Bus) або брокер повідомлень (наприклад, Kafka) для надійної передачі інтегрального показника ризику. Це гарантує, що критичні дані про ризик будуть доставлені до всіх залежних систем (SIEM, SOAR, IAM) у режимі реального часу, навіть за високих навантажень або тимчасових збоїв.

Така інтеграційна архітектура перетворює СУКБ із набору ізольованих інструментів на цілісну, адаптивну систему управління, яка використовує психологічні дані для підвищення своєї загальної ефективності.

3.3.2. Розробка процедур реагування на інциденти з урахуванням психологічного контексту

Навіть найдосконаліша технічна модель є лише інструментом. Її ефективність у кінцевому підсумку визначається тим, як адміністративні процедури та персонал організації використовують отримані дані. Розробка процедур реагування на інциденти повинна принципово відрізнятися від традиційних підходів, оскільки інцидент, спричинений психологічним фактором (втомою, стресом), вимагає іншого управлінського та кадрового реагування, ніж інцидент, спричинений технічним збоєм чи зовнішньою атакою.

Впровадження принципу «Культури безпеки без провини» (No-Blame Security Culture)

Ключовим принципом, що забезпечує успіх цієї моделі, є формування Культури безпеки без провини. Якщо користувачі знатимуть, що система моніторингу

використовується для їхнього покарання за помилки, вони свідомо чи підсвідомо намагатимуться обійти моніторинг або приховати факти помилок, що значно підвищить загальний ризик. Натомість, фокус має бути зміщений на системну підтримку та навчання.

- Диференціація реагування: Процедури мають чітко розмежовувати інциденти, спричинені ненавмисною помилкою (високий показник KIR, пов'язаний зі втомою чи неуважністю), та інциденти, спричинені зловмисним наміром (високий показник KIR, пов'язаний з аномальним доступом до даних). Реакція на першу категорію має бути спрямована на допомогу та навчання, а на другу - на розслідування та санкції.
- Конфіденційність даних: Необхідно розробити суворі правила щодо доступу до інтегрального показника ризику та первинних психологічних індикаторів. Ці дані повинні бути доступні лише вузькому колу фахівців (CISO та UBA-аналітикам) і не повинні використовуватися для рутинного оцінювання продуктивності чи дисциплінарних стягнень, за винятком випадків доведеного зловмисного наміру.

Створення Плану комунікації та взаємодії у зоні ризику

План реагування має включати детальні інструкції для різних функціональних підрозділів щодо взаємодії з користувачем, який потрапив у «жовту» чи «червону» зону ризику:

1. Служба безпеки (SOC): Отримує критичний сигнал про перехід у «червону зону». Завдання SOC - підтвердити, чи було здійснено технічне блокування (за сценарієм 3.1.3), та розпочати розслідування. Пріоритет- мінімізація шкоди.

2. Відділ кадрів (HR): HR-спеціалісти, особливо у випадку стійких KIR, пов’язаних зі стресом або надмірною роботою, повинні бути проінформовані. Їхня роль полягає у підтримуючому втручанні: наприклад, пропозиція вихідного дня, консультація психолога (якщо це передбачено корпоративною програмою) або аналіз робочого навантаження. Це перетворює процедуру безпеки на процедуру захисту співробітника.

3. Керівник підрозділу: Керівник повинен бути сповіщений про перехід свого співробітника у зону підвищеного ризику, але це сповіщення має бути максимально деперсоналізованим («Ваш співробітник Х демонструє високі індикатори втоми, що підвищує ризик помилки. Рекомендовано перерозподілити його навантаження»).

Ці процедури гарантують, що інтеграція моделі ризику не лише посилює технічні захисні механізми, але й трансформує корпоративну культуру, роблячи її більш орієнтованою на людину та психологічно безпечною.

3.3.3. Критерії та методика оцінки економічної ефективності впровадження адаптивних механізмів

Впровадження будь-якої інноваційної системи управління кібербезпекою, особливо такої, що вимагає інтеграції нових технологій (UBA, динамічне моделювання), повинно бути обґрунтоване не лише з технічної, але й з економічної точки зору. Керівництво організації потребує чітких доказів того, що інвестиції у розробку та впровадження адаптивних механізмів управління, орієнтованих на психологічні фактори, забезпечать відчутне повернення інвестицій (Return on Investment- ROI).

Критерії економічної ефективності

Оцінка ефективності має базуватися на зниженні двох основних категорій втрат, пов'язаних з людським фактором:

1. Прямі фінансові втрати від інцидентів: Включають витрати на ліквідацію наслідків витоків даних, штрафи за недотримання комплаєнсу (наприклад, GDPR), витрати на судові процеси та відновлення систем після помилок.

2. Непрямі втрати та операційні витрати: Включають зниження продуктивності праці через "втому від безпеки" (security fatigue), вартість простою систем, витрати на розслідування та час, витрачений співробітниками на усунення помилок, спричинених їхнім психологічним станом.

Пропонована модель спрямована на зниження ймовірності та масштабу цих втрат.

Методика розрахунку показника ROI

Розрахунок ROI для адаптивних механізмів базується на порівнянні витрат на впровадження та річних запобіжних втрат.

$$ROI = \frac{(\text{Запобіжні витрати}) - (\text{Витрати на впровадження})}{\text{витрати на впровадження}} \times 100\% \qquad (3.1)$$

1. Оцінка витрат на впровадження: Ці витрати включають придбання або ліцензування UBA-платформ, розробку та інтеграцію Модуля оцінки ризику (якщо не використовується готове рішення), навчання персоналу (SOC, HR) та адміністративні витрати на розробку нових процедур.

2. Оцінка запобіжних втрат (Avoided Costs): Це найважливіший та найскладніший параметр. Запобіжні втрати розраховуються на основі:

– Зниження частоти інцидентів, спричинених людським фактором (Reduction in Incident Frequency): Методика передбачає використання результатів валідації моделі (підрозділ 2.3.3). Якщо модель доводить, що проактивні втручання (ALP та Nudges) дозволяють запобігти X% інцидентів на рік, то загальна очікувана річна втрата від інцидентів знижується на відповідну величину.

– Зниження середніх витрат на інцидент (Reduction in Cost Per Incident): Навіть якщо інцидент стався, ALP мінімізує його масштаб, оскільки привілеї користувача були обмежені. Наприклад, якщо інцидент, спричинений неуважністю, раніше коштував Y, то завдяки ALP, який тимчасово заблокував доступ до критичної бази даних, витрати знижуються до Y'.

Критерії вимірювання зниження частоти інцидентів:

Для кількісного підтвердження зниження частоти інцидентів використовуються такі метрики, які є прямим результатом роботи системи Nudges:

– Коефіцієнт клікабельності фішингових посилань (Phishing Click-Through Rate): Впровадження контекстуальних Nudges у поштових клієнтах повинно призвести до стійкого зниження цього показника в ході симуляційних атак.

– Частота повторних помилок (Recurrence Rate of Errors): Відстеження того, як часто користувач, який отримав Nudge, повторює ту саму ризиковану поведінку. Успішна система повинна демонструвати значне зниження цієї частоти, підтверджуючи ефективність психологічної корекції.

Успішна економічна оцінка, яка демонструє позитивний ROI, є фінальним аргументом на користь впровадження розробленого комплексу адаптивних механізмів, завершуючи обґрунтування практичної цінності дипломної роботи.

3.4 Висновки до розділу 3

У третьому розділі було розроблено та обґрунтовано комплекс адаптивних механізмів управління, що дозволяють трансформувати теоретичну модель психологічних ризиків у практичний інструментарій захисту інформаційних систем. Центральним елементом запропонованого підходу став механізм адаптивного принципу найменших привілеїв, який долає обмеження традиційних статичних моделей контролю доступу. На відміну від класичних систем, де права користувача закріплені за роллю, впроваджений алгоритм дозволяє динамічно звужувати або розширювати обсяг доступних функцій залежно від поточного психоемоційного стану суб'єкта. Це забезпечує створення гнучкого захисного шару, який автоматично ізолює критичні операції у моменти підвищеної вразливості користувача, спричиненої втомою або стресом.

Паралельно з технічними обмеженнями доступу було впроваджено систему проактивних поведінкових втручань, відомих як «Nudges». Цей механізм спрямований на м'яку корекцію поведінки через психологічне спрямування користувача до безпечного вибору в моменти когнітивного перевантаження. Розроблена класифікація Nudges дозволяє системі обирати найбільш релевантний тип впливу - від контекстуальних нагадувань до зміни архітектури вибору - залежно від домінуючого ключового індикатора ризику (KIR). Важливою особливістю цієї підсистеми є її здатність виводити користувача зі стану «автопілота», стимулюючи критичне мислення та підвищуючи пильність саме тоді, коли ймовірність помилки є найвищою.

Фінальна інтеграція розроблених рішень у загальну архітектуру Системи управління кібербезпекою (СУКБ) дозволила об'єднати модулі оцінки ризику з інфраструктурними компонентами, такими як SIEM та SOAR. Сформований механізм зворотного зв'язку забезпечує самонавчання системи через постійну оцінку ефективності вжитих заходів та коригування вагових коефіцієнтів моделі. Такий підхід забезпечує перехід від реактивного реагування на інциденти до проактивного управління кібербезпекою, де технологічні заходи захисту гармонійно поєднуються з

урахуванням людського фактора, мінімізуючи ризики без порушення неперервності бізнес-процесів.

ВИСНОВКИ

Проведене дослідження у рамках магістерської роботи досягло своєї головної мети, яка полягала у розробці та глибокому обґрунтуванні комплексу механізмів управління кібербезпекою, що інтегрують критично важливі психологічні аспекти поведінки користувачів. Фінальний результат роботи являє собою не просто теоретичну концепцію, а цілісну, архітектурно оформлену систему, здатну трансформувати підхід організації до мінімізації ризиків, спричинених людським фактором. Узагальнюючи виконану роботу, можна сформулювати низку ключових висновків, які підтверджують наукову новизну та практичну значущість одержаних результатів.

На першому етапі дослідження було однозначно встановлено та глибоко проаналізовано фундаментальний розрив між існуючими статичними підходами до кібербезпеки та динамічною природою людської вразливості. Традиційні механізми, такі як фіксовані політики доступу та загальноосвітні інструктажі, виявилися недостатньо дієвими, оскільки вони повністю ігнорують контекстуальний стан користувача - його втому, рівень стресу, когнітивне навантаження або емоційну дестабілізацію, які є прямими каталізаторами більшості внутрішніх інцидентів. Таким чином, теоретичне обґрунтування підтвердило життєву необхідність переходу від реактивного контролю, що фіксує наслідки, до проактивного управління, яке ґрунтується на постійному, об'єктивному прогнозуванні поведінки. Саме це усвідомлення необхідності динамічності заклало методологічну основу для всієї подальшої розробки.

Наступний, центральний етап роботи присвячено створенню динамічної моделі оцінки персонального кіберризику. Наукова новизна тут полягає у розробці та формалізації ключових індикаторів психологічного ризику (KIR), таких як аномалії швидкості набору, частоти помилок аутентифікації чи нетипова тривалість сесії, які є мостом між сирими технічними даними UBA-систем та психологічною інтерпретацією. Була успішно розроблена комплексна методика для кількісного

перетворення цих різнорідних даних у єдиний інтегральний показник ризику. Це досягалося шляхом нормалізації, ієрархічного ранжування та присвоєння вагових коефіцієнтів, що дозволило позбутися суб'єктивності у оцінці. Обґрунтування застосування складного математичного апарату- зокрема, динамічних Баєсових мереж- дозволило впровадити критично важливу прогностичну функцію, яка враховує часову залежність та нелінійну взаємодію факторів ризику, класифікуючи користувачів за чіткими зонами ризику (Низький, Підвищений, Критичний) у режимі реального часу.

Кульмінацією дослідження стала розробка двох взаємодоповнюючих адаптивних механізмів, які є прямим практичним втіленням моделі ризику. По-перше, був створений Адаптивний принцип найменших привілеїв (ALP), що є інноваційним рішенням для технічного захисту. ALP автоматично коригує привілеї користувача, не за роллю, а за його поточним інтегральним показником ризику: у разі переходу у «жовту зону» відбувається тимчасове звуження доступу до критичних функцій, а у «червоній зоні»- негайне блокування та ізоляція сесії. Це забезпечує проактивне мінімізацію збитку ще до того, як помилка чи намір зможуть реалізуватися. По-друге, розроблений механізм проактивних поведінкових втручань «Nudges» використовує принципи поведінкової психології для м’якої корекції ризикової поведінки. «Nudges» активуються лише в «жовтій зоні» відповідно до домінуючого KIR, наприклад, пропозиція відпочинку при високій втомі або контекстуальне нагадування про пильність. Цей психологічний інструмент дозволяє системі не карати користувача за його вразливість, а надавати йому своєчасну, ненав’язливу підтримку, сприяючи формуванню сталої безпечної поведінки.

Нарешті, для забезпечення життєздатності та масштабованості розроблених механізмів, була деталізована архітектура інтеграції моделі у загальну Систему управління кібербезпекою. Інтеграція з SIEM та SOAR дозволяє збагатити технічні тривоги психологічним контекстом, забезпечуючи автоматизовану та диференційовану реакцію. Передача інтегрального показника ризику до системи GRC переводить управління ризиком людського фактора на стратегічний, кількісно вимірюваний рівень. Вирішальне значення має обґрунтування адміністративних процедур, що базуються на «Культурі безпеки без провини», що є запорукою достовірності даних та

лояльності співробітників. Комплекс цих заходів, підкріплений розробленою методикою оцінки економічної ефективності (ROI), яка прямо корелює впровадження моделі зі зниженням частоти та масштабів інцидентів, підтверджує, що розроблена система є інвестиційно привабливою та необхідною інновацією для будь-якої сучасної організації, що прагне досягти найвищого рівня кіберстійкості.

ПЕРЕЛІК ПОСИЛАНЬ

1. Інформаційний портал «NSA та CISA розкривають 10 найбільш поширених помилок у забезпеченні кібербезпеки». URL:https://internetua.com/nsa-ta-cisa-rozkrivauat-10-naibilsh-poshirenih-pomilok-u-zabezpecsenni-kiberbezpeki?utm_source=ukrnet_news

2. Інформаційний портал «Top 5 Cyberattacks Caused By Human Error». URL:https://www.vumetric.com/blog/top-5-cyberattacks-caused-by-human-error

3. Інформаційний портал «CISOs list human error as their top cybersecurity risk». URL:https://www.ibm.com/think/insights/cisos-list-human-error-top-cybersecurity-risk

4. Wikipedia «Комп'ютерна безпека». URL:https://uk.wikipedia.org/wiki/Комп%27ютерна_безпека

5. Fortinet «What is the CIA Triad and Why is it important?». URL:https://www.fortinet.com/resources/cyberglossary/cia-triad

6. Fortinet «What is Defense in Depth? Defined and Explained». URL:https://www.fortinet.com/resources/cyberglossary/defense-in-depth

7. Fortinet «How to Enforce the Principle of Least Privilege to Reduce Security Risks». URL:https://www.fortinet.com/resources/cyberglossary/principle-of-least-privilege

8. Wikipedia «Цикл Деминга». URL:https://ru.wikipedia.org/wiki/Цикл_Деминга

9. ISO/IEC 27001:2013. Information technology — Security techniques — Information security management systems — Requirements. Geneva : ISO/IEC, 2013. 23 p.

10. ISO/IEC 27002:2013. Information technology — Security techniques — Code of practice for information security controls. Geneva : ISO/IEC, 2013. 80 p.

11. NIST «cybersecurity Framework»**. URL:**https://www.nist.gov/cyberframework

12. Wikipedia «Capability Maturity Model Integration». URL:https://en.wikipedia.org/wiki/Capability_Maturity_Model_Integration#:~:text=Capability%20Maturity%20Model%20Integration%20(CMMI,contracts%2C%20especially%20in%20software%20development.

13. ISACA «CMMI». URL:https://www.isaca.org/enterprise/cmmi-cybermaturity-platform

14. «КЛАСИФІКАЦІЯ ЗАГРОЗ І РИЗИКІВ СУЧАСНИХ ІНФОКОМУНІКАЦІЙНИХ СИСТЕМ» / Шматок О.С., Фіненко Ю.І., Єлізаров А.Б., Телющенко В.А. 2019. 9с.

15. Hostzealot «Що таке IPS/IDS і де застосовується». URL:https://www.hostzealot.com.ua/blog/about-solutions/shho-take-ipsids-i-de-zastosovujetsya

16. «СТРАТЕГІЇ КІБЕРСТІЙКОСТІ: УПРАВЛІННЯ РИЗИКАМИ ТА БЕЗПЕРЕРВНІСТЬ БІЗНЕСУ» / Легомінова С.В. та ін. Київ 2025. 305с.

17. SurveyMonkey «Что такое шкала Лайкерта?». URL:https://ru.surveymonkey.com

18. IBM «What is user behavior analytics (UBA)?». URL:https://www.ibm.com/think/topics/user-behavior-analytics

19. Ironscales «What Is Impossible Travel?». URL:https://ironscales.com/glossary/impossible-travel

20. PingIdentity «Why Security Fatigue Is a Huge Cybersecurity Risk» by Louise Watson. URL:https://www.pingidentity.com/en/resources/blog/post/why-security-fatigue-huge-cybersecurity-risk.html