УДК 004.3-185.4; 004.7-185.4, 330.4, 336; 336.01;
336.11; 336.741.28; 336.7
УКПП
№ держреєстрації 0121U109559
Інв. №

Міністерство освіти і науки України
Сумський державний університет (СумДУ)
40007, м. Суми, вул. Р.-Корсакова, 2; тел. 66-50-37
cyber@uabs.sumdu.edu.ua

ЗАТВЕРДЖУЮ
Проректор з наукової роботи
д-р. фіз.-мат. наук, професор

А.М. Чорноус

ЗВІТ
ПРО НАУКОВО-ДОСЛІДНУ РОБОТУ

НАЦІОНАЛЬНА БЕЗПЕКА ЧЕРЕЗ КОНВЕРГЕНЦІЮ СИСТЕМ ФІНАНСОВОГО МОНІТОРИНГУ ТА КІБЕРБЕЗПЕКИ: ІНТЕЛЕКТУАЛЬНЕ МОДЕЛЮВАННЯ МЕХАНІЗМІВ РЕГУЛЮВАННЯ ФІНАНСОВОГО РИНКУ

(остаточний)

Частина 1

Керівниця НДР
доцентка кафедри економічної кібернетики
д-рка екон. наук, доцентка Г.М. Яровенко

2023

Рукопис закінчений 22 грудня 2023 р.

Результати роботи розглянуті науковою радою СумДУ, протокол від 2023.12.28 №13

СПИСОК АВТОРІВ

Доцентка кафедри економічної кібернетики, д-рка екон. наук, доцентка (керівниця)	22.12.2023	Г.М. Яровенко (вступ, підрозділи 1.1.2, 1.1.3, 1.2-1.4, 2.1, 2.2.2-2.2.3, 2.3-2.4, розділ 3, висновки)
Професорка кафедри економічної кібернетики, д-рка екон. наук, професорка	22.12.2023	О.В. Кузьменко (підрозділи 1.1.1, 1.1.2, 1.3., 2.4.1, 2.4.2)
Професор кафедри економічної кібернетики, д-р екон. наук, професор	22.12.2023	С.В. Лєонов (підрозділ 1.1.1, 1.1.3, 2.4.2, 3.1.1)
Викладачка-стажистка кафедри економічної кібернетики (відповідальна виконавиця)	22.12.2023	О.В. Колотіліна (підрозділи 1.3.2, 2.3.3, 3.4)
Старший викладач кафедри економічної кібернетики, д-р філософії	22.12.2023	С.В. Миненко (підрозділи 2.2.1, 3.1.1)
Асистентка кафедри економічної кібернетики, д-рка філософії	22.12.2023	Т.В. Доценко (підрозділ 1.1.1)
Докторка філософії	22.12.2023	Н.Ю. Сідельник (підрозділ 3.3.2)
Аспірантка кафедри економічної кібернетики	22.12.2023	М.С. Рожкова (підрозділ 2.2.1)
Магістрантка кафедри економічної кібернетики	22.12.2023	В.В. Радько (підрозділ 1.2.2, 2.1)
Магістрантка кафедри економічної кібернетики	22.12.2023	А.О. Рапута (підрозділ 1.3.2, 2.2.2, 3.1.2)

РЕФЕРАТ

Звіт про НДР: 615 с., 304 рис., 54 табл., 132 формул, 356 джерел, 12 додатків.

КІБЕРБЕЗПЕКА, КОНВЕРГЕНЦІЯ, НАЦІОНАЛЬНА БЕЗПЕКА, ФІНАНСОВИЙ МОНІТОРИНГ, ІНТЕЛЕКТУАЛЬНЕ МОДЕЛЮВАННЯ.

Об’єкт дослідження – система економічних відносин між суб’єктами економіки та регуляторами фінансового ринку в процесі комплексного застосування засобів фінансового моніторингу та боротьби із кіберзлочинністю. Мета роботи – розвиток методології та міждисциплінарного методичного інструментарію пошуку оптимальної моделі інтеграції систем фінансового моніторингу та кібербезпеки, що дозволить напрацювати принципово нові, засновані на концептах поведінкової економіки та відокремлені від людського фактору, інтелектуальні алгоритмізовані регуляторні механізми, які уможливлять комплексне забезпечення економічної, фінансової та інформаційної складових національної безпеки держави, а також захисту прав споживачів фінансових послуг. Методи дослідження: фундаментальні положення економічної теорії, макро- і мікроекономіки, теорії стратегічного управління, державного регулювання економіки, економіко-математичного моделювання. Інформаційно-фактологічна база: законодавчі та нормативні документи Національного банку України, інструкції банків, наукові праці вітчизняних та зарубіжних фахівців.

Обґрунтовано концепції конвергенції системи фінансового моніторингу та кібершахрайств: розроблено фазові портрети «зрілості», «станів рівноваги» та «релаксаційних коливань втрати стійкості», ключові алгоритми систем фінансового моніторингу та кібербезпеки, сценарії взаємодії систем кібербезпеки та протидії фінансовим злочинам, оцінено ефект від конвергенції. Проведено модернізацію інструментарію протидії легалізації кримінальних доходів та кібершахрайствам: проведено структурний багатошаровий аналіз, побудовано моделі ризиків конвергенції, розроблено алгоритми розпізнавання поведінки кібершахраїв, зроблено прогнози кібератак. Сформовано управлінські засади забезпечення стійкості фінансового кіберпростору на мікро- та макрорівні: концепція експертної системи діагностики кібершахрайств, її інформаційне забезпечення, заходи ребілдингу на основі заходів кібербезпеки, модель стійкого розвитку країни.

ЗМІСТ

ВСТУП

За останні десятиліття стрімкий розвиток інформаційних технологій вніс свій внесок у впровадження їх у діяльність економічних суб'єктів для розв'язання різноманітних економічних завдань. Раніше було достатньо мати інформаційну систему обліку, яка повністю відповідала потребам суб'єктів у автоматизації їхньої діяльності. Проте сьогодні спектр завдань є значно ширшим і не обмежується лише бухгалтерським обліком. З іншого боку, впровадження автоматизації та діджиталізації процесів призвело до зростання рівня кібершахрайств, особливо в сфері фінансів. Це пов'язано зі збільшенням доступності програмних та технічних інструментів для здійснення кіберзлочинів та зростанням рівня інформаційної грамотності населення. Отже, сучасна діяльність фінансових установ спрямована на забезпечення необхідного рівня кіберзахисту в умовах поширення можливостей для вчинення кіберзлочинів.

Фінансові установи також є об'єктами фінансового моніторингу, що вимагає від них перевірки операцій, які можуть стати об'єктом легалізації кримінальних доходів або фінансування тероризму. Навіть при зростаючих вимогах до системи протидії відмиванню грошей, які передбачають постійний моніторинг для виявлення підозрілих операцій, спостерігається збільшення обсягів операцій, спрямованих на легалізацію кримінальних доходів та фінансування тероризму. Ці операції часто виконуються за допомогою втручання хакерів та інших кіберзлочинців. Зазнаається зростання поєднання кібер- та фінансової злочинності, що має наслідком втрати для фінансових інституцій та падіння рівня довіри їм з боку громадян та підприємств. На тлі цього також спостерігається зростання потоків інформації, стрімка зміна оточуючого середовища та вдосконалення програмних та технічних інструментів. Ці фактори ведуть до того, що фінансові установи неспроможні ефективно протистояти кібер- та фінансовим злочинам. Отже, ідея конвергенції кібербезпеки та фінансового моніторингу стає важливою і практично значущою для фінансових установ, оскільки це сприятиме спрощенню управлінських

процесів, пов'язаних із виявленням та запобіганням кіберзлочинам та легалізації кримінальних доходів.

Основною передумовою злиття системи фінансового моніторингу та кібербезпеки є зростання обсягу інформаційних потоків. Їх інтеграція на рівні інформації, в першу чергу, розширить набір критеріїв для перевірки та виявлення злочинних операцій і дій. По-друге, це дозволить охопити різноманітні бази вхідних даних. На організаційному рівні злиття дозволить відповідним відділам фінансової установи обмінюватися інформацією, сприяючи більш ефективному моніторингу операцій, не лише в аспекті їх відповідності законодавству, але й у виявленні можливого кібершахрайства. На технологічному рівні, з розвитком можливостей використання інтелектуальних методів моделювання, виникають перспективи модернізації технологій та інструментів, застосовуваних для виявлення злочинних схем та операцій. Це сприятиме їхньому розвитку на темпи, що відповідають темпам розвитку інструментарію кібер- та фінансових злочинців.

Злиття систем кібербезпеки та фінансового моніторингу призведе до отримання ряду переваг. По-перше, це призведе до зменшення витрат, пов'язаних з управлінням двома окремими системами, зокрема в сфері залучення персоналу та використання програмно-технологічного забезпечення. По-друге, впровадження інтегрованого інформаційного забезпечення підвищить якість аналітичної інформації, охоплюючи критерії та вимоги, щодо протидії як легалізації коштів, так і кібершахрайству. По-третє, синергетичний ефект виникне від інтегрованої взаємодії обох систем, що призведе до підвищення ефективності перевірок за рахунок впровадження комплексу різних методів та інструментів.

Таким чином, сучасні реалії зростання обсягів кібершахрайств та легалізації кримінальних доходів, потребують не тільки збільшення вимог до операцій, але впровадження більш дієвих заходів, реалізація яких можлива на інформаційному, програмно-технологічному та організаційному рівнях управління фінансовою установою. Відповідно забезпечення цих процесів

можливо тільки за рахунок конвергенції двох систем – кібербезпеки та фінансового моніторингу.

Окреслена проблема дозволила обрати об’єкт та предмет дослідження. Об’єкт дослідження – система економічних відносин, що виникають між суб’єктами господарювання та регуляторами фінансового ринку, що виникають в процесі комплексного застосування засобів фінансового моніторингу та боротьби із кіберзлочинністю.

Предмет дослідження – методологічні засади та методичний формування комплексних, упереджувальних інтелектуальних механізмів регулювання фінансового ринку, що сприятимуть підвищенню національної безпеки в умовах цифровізації фінансового простору.

Відповідно до об’єкта та предмета дослідження було сформовано мету. Так, метою дослідження є розвиток методології та міждисциплінарного методичного інструментарію пошуку оптимальної моделі інтеграції систем фінансового моніторингу та кібербезпеки, що дозволить напрацювати принципово нові, засновані на концептах поведінкової економіки та відокремлені від людського фактору, інтелектуальні алгоритмізовані регуляторні механізми, які уможливлять комплексне забезпечення економічної, фінансової та інформаційної складових національної безпеки держави, а також захисту прав споживачів фінансових послуг.

Для реалізації поставленої мети необхідно було вирішити наступні завдання:

- охарактеризувати поняття, цілі й задачі, напрями, моделі конвергенції систем фінансового моніторингу і кібербезпеки;
- здійснити попередній аналіз процесу конвергенції систем кібербезпеки та фінансового моніторингу країн;
- побудувати фазові портрети «зрілості», «станів рівноваги» та «релаксаційних коливань втрати стійкості» системи протидії фінансовим та кібершахрайствам;

- здійснити моделювання ключових алгоритмів конвергенції системи кібербезпеки та фінансового моніторингу у банках;
- розробити математичне забезпечення алгоритмів виявлення кібершахрайських операцій з кредитними картками;
- розробити чотириполюсні барицентричні моделі збалансованого розвитку національної економіки, що інтегрують композитні індикатори економічного, соціального й політичного розвитку країни, а також рівня її протидії фінансовим шахрайствам та кібербезпеки;
- розробити сценарії взаємодії систем кібербезпеки та протидії фінансовим злочинам для країн з різним рівнем економічного розвитку;
- оцінити синергетичний ефект від конвергенції моделей фінансового моніторингу та кібербезпеки;
- провести базовий, статистичний, кореляційний, декомпозиційний аналіз часових рядів;
- побудувати об’єднані регресії, регресії з детермінованими індивідуальними і випадковими ефектами для змінних «Mail Anti Virus», «Kaspersky Anti-Spam», «Intrusion Detection Scan»;
- спрогнозувати тренди трьох видів кібератак на основі рекурентної нейронної мережі Long short-term memory та об’єднані регресії;
- проаналізувати сутність поняття «протидія легалізації доходів отриманих незаконним шляхом» в умовах діджиталізації суспільства;
- здійснити оцінку ризику конвергенції системи протидії відмивання грошей та кібербезпеки;
- побудувати нейромережеву модель потенційної конвергенції системи кібербезпеки та протидії легалізації кримінальним доходам;
- сформовати кіберпрофілі жертв на основі гендерного аналізу;
- розробити кіберпрофілі сучасних фінансових кіберзлочинів на основі кластерного аналізу;
- розробити алгоритми розпізнавання поведінки кіберзлочинців на основі методів інтелектуального аналізу;

- розроблено моделі прогнозування кібершахрайських атак на комп’ютерні системи, мережеву та хмарну інфраструктуру фінансової установи;
- розробити ударно-хвильову модель впливу кібершахрайських атак на рівень фінансової безпеки;
- розробити імітаційну модель діяльності інсайдера у банку;
- здійснити моделювання ймовірної поведінки дій інсайдерів-кібершахраїв у банку;
- розробити структуру інформаційної бази експертної системи виявлення інсайдерських кіберзагроз у банках;
- сформувати підхід щодо розробки онтологічної моделі для формалізації інформаційного забезпечення фінансової установи;
- сформувати соціально-економічні профілі країн - жертв кіберзлочинів;
- розробити методику оцінювання потреб суб’єктів економіки у розвитку кібербезпеки;
- розробити модель стійкого розвитку країни, що інтегрує композитні таргети економічного, соціального, політичного та кібербезпекового регуляторних вимірів.

Методи дослідження – фундаментальні положення економічної теорії, макро- і мікроекономіки, теорії стратегічного управління, державного регулювання економіки, економіко-математичного моделювання (теорія біфуркації динамічних систем; канонічний аналіз; гендерний аналіз; статистичний аналіз; декомпозиційний аналіз часових рядів; регресійний аналіз панельних даних; сучасні концепції моделювання бізнес-процесів; методи експоненційного згладжування; методи інтелектуального аналізу даних: логіт-регресія, кластерний аналіз, асоціативний аналіз, класифікаційні дерева рішень, нейронні мережі, рекурентні нейронні мережі, самоорганізовані карти; модель Сєдова-Тейлора; метод переваг та функція Харрінгтона – Менчера; метод визначення центра мас; DEA-аналіз; системно-динамічне моделювання). Розрахунки та візуалізація в роботі проводилися із використанням мови

програмування Python та аналітичних пакетів STATISTICA, MS Excel, Deductor Academic, Frontier Analyst, Viscovery 8, AnyLogic 8 PLE, Bizagi Studio, VOSviewer.

Інформаційно-фактологічну базу дослідження сформували законодавчі та нормативні документи Національного банку України, інструкції банків, наукові праці вітчизняних та зарубіжних фахівців.

Отримані у роботі результати впроваджені у навчальний процес, а саме при викладанні дисциплін «Введення до бізнес-аналітики», «Інформаційні системи і технології в управлінні», «Ефективність інформаційних систем», «Моделювання бізнес-процесів», «Моделювання економіки», «Прогнозування соціально-економічних процесів», «Машинне навчання для забезпечення кібербезпеки у сфері фінансових послуг».

За результатами НДР опубліковано: 22 статті у журналах, що індексується у БД Scopus та WoS; 24 фахові статті у виданнях України категорії Б; 8 – монографій та розділів у закордонних монографіях англійською мовою та вітчизняних монографіях українською мовою; захищено 6 дисертацій на здобуття наукового ступеня доктора філософії та доктора наук та 7 бакалаврських та магістерських робіт; виконано 5 конкурсних робіт за тематикою НДР, за які отримано дипломи I, II та III ступеня за напрямами «Економічна аналітика і статистика» та «Економічна кібернетика» у 1-му та 2-му турах Всеукраїнського конкурсу студентських наукових робіт; також було подано та отримано 16 свідоцтв про реєстрацію авторського права на твір.

Звіт виконано на основі публікацій виконавців, перелік яких надано у списку літератури.

1 ОБҐРУНТУВАННЯ КОНЦЕПЦІЇ КОНВЕРГЕНЦІЇ СИСТЕМИ ФІНАНСОВОГО МОНІТОРИНГУ ТА КІБЕРШАХРАЙСТВ

1.1 Оцінювання зрілості діючої системи протидії фінансовим та кібершахрайствам та побудова фазових портретів їх «зрілості», «станів рівноваги» та «релаксаційних коливань втрати стійкості»

1.1.1 Конвергенція систем фінансового моніторингу і кібербезпеки: поняття, цілі й задачі, напрями, моделі

Стрімкий розвиток, зростання та накопичення сучасних технологій, інформаційного забезпечення, цифрових даних, створюють передумови застосування інформаційних технологій у різних сферах та напрямах діяльності. В Україні та світі наразі створено комфортні умови вільної роботи з новітніми технологіями фінансово-економічної галузі в online режимі. Так, фізичні особи та суб’єкти господарювання отримали майже необмежені можливості дистанційно користуватись фінансовими послугами, підключатись до online банкінгу, валютних бірж, фондового ринку, інших фінансово-кредитних установ та організацій. Наряду з цим, кожен інноваційний процес супроводжується певним загрозами. Так, в інформаційному, віртуальному кіберпросторі з’являться загроза зростання злочинів через його доступність як легальним учасникам, так і злочинцям. Але застосування в економічній сфері таких сучасних можливостей, формує потребу у забезпеченні належної економічної безпеки фінансових операцій, що проходять через відповідні системи.

За останні десять років в Україні було сформовано систему боротьби з відмиванням незаконних коштів, фінансуванням тероризму, розповсюдження зброї масового знищення, в тому числі з кіберзагрозами, у вигляді комплексу заходів з фінансового моніторингу та системи кібербезпеки, що передбачає перевірку клієнтів та їх фінансових транзакцій, з метою контролю за економічною чистотою та прозорістю фінансових транзакцій.

Тому, в сучасних ринкових умовах, питання конвергенції систем фінансового моніторингу та кібербезпеки є особливо актуальним та потребує детального вивчення та аналізу.

Загальні питання з фінансового моніторингу теоретичного характеру та його специфічні практичні особливості у своїх трактатах розкривають як вітчизняні, так і зарубіжні вчені, серед яких: Морс Дж. К. [1], який висвітлює глобальний режим боротьби з фінансуванням тероризму під впливом транснаціонального ринку; Радигін В. Ю., Купріянов Д. Ю., Бессонов Р. А., Іванов М. Н., Ослякова І. В. [2], які пропонують шляхи вирішення завдань первинного фінансового моніторингу; Яшина Н. І., Кашина О. І., Прончатова-Рубцова Н. Н., Яшин С. Н., Кузнєцов В. П. [3], які висвітлюють окремі аспекти фінансового моніторингу, фінансової стабільності та цифровізації; Грабчук О., Супрунова І. [4], які розкривають поняття, складові, етапи розвитку фінансового моніторингу як умови забезпечення державної безпеки країни; Першин В. Г. [5], який розглядає проблематику, визначає роль фінансового моніторингу в межах протидії легалізації доходів, одержаних злочинним шляхом, пропонує шляхи удосконалення системи фінансового моніторингу; Рисін В. В., Степанова А. В. [6], які описують інструменти протидії фінансуванню тероризму з використанням фінансових установ; та ін.

Питання вивчення, дослідження, використання, удосконалення поняття кібербезпека у сучасній науковій економічній літературі розглядаються такими вченими як: Шекелфорд С., Докері Р., Прабхакар Б., і Реймонд А. [7] досліджують кібербезпеку в умовах кризи; Ученду Б., Медсестра Дж. Р. К., Бада М. і Фернелл С. [8] узагальнюють розвиток культури кібербезпеки; Хан К. [**Ошибка! Источник ссылки не найден.**] пропонує використовувати напівкількісну оцінку ризиків кібербезпеки шляхом аналізу рівня блокування та захисту; Мохор В., Гончар С., Онискова А. [**Ошибка! Источник ссылки не найден.**] запроваджують оцінку ризиків кібербезпеки інформаційних систем; Гіменес-Агілар М., де Фуентес Ж. М., Гонсалес-Манцано Л., та Арройо Д. [**Ошибка! Источник ссылки не найден.**] висвітлюють практичні досягнення

кібербезпеки в системах на основі блокчейну; Репетто М., Стрікколі Д., Піро Г., Каррега А., Боггіа Г., і Болла Р. [**Ошибка! Источник ссылки не найден.**] аналізують автономну систему кібербезпеки для ланцюгів цифрових послуг нового покоління; та ін.

Результати впливу конвергенції на ефективність досліджуваних процесів у своїх працях розкривають наступні науковці: Мадейра П. М., Вале М., Мора-Алиседа Дж. [**Ошибка! Источник ссылки не найден.**] пропонують комбінування стратегії розумної спеціалізації та регіональної конвергенція в економіці; Ібрагім А. Е. А., Еламер А. А., і Езат А. Н. [**Ошибка! Источник ссылки не найден.**] описують конвергенцію великих даних та бухгалтерського обліку при прогнозування; Донг Ф., Лі Ю., Цінь К. і Сан Дж. [**Ошибка! Источник ссылки не найден.**] досліджують вплив промислової конвергенція на ефективність екологічного розвитку; Гілбо Д., Баронселлі А., і Ченгола Д. [**Ошибка! Источник ссылки не найден.**] висвітлюють експериментальні докази конвергенції аналізованих категорій; та ін.

Отже, питаннями пошуку методів та шляхів упередження, протидії та боротьби з фінансовими злочинами, при здійснені яких застосовуються інформаційні, технологічні, комунікаційні, технічні системи, опікуються і національні державні органи, і міжнародне світове співтовариство.

Розглядаючи особливості фінансового моніторингу, в першу чергу необхідно розглянути його визначення. Так згідно Закон України «Про запобігання та протидію легалізації (відмиванню) доходів, одержаних злочинним шляхом, фінансуванню тероризму та фінансуванню розповсюдження зброї масового знищення» «фінансовий моніторинг - сукупність заходів, що вживаються суб’єктами фінансового моніторингу у сфері запобігання та протидії, що включають проведення державного фінансового моніторингу та первинного фінансового моніторингу» [17]. У загальному трактуванні, фінансовий моніторинг представляє собою комплексну систему принципів, заходів, методів, методик, які проводяться суб’єктами та учасниками фінансового моніторингу для виявлення, протидії, запобігання використання

фінансової та банківської системи для легалізації доходів, отриманих незаконним шляхом, фінансування терористично діяльності, розповсюдження зброї масового знищення; сукупність дій по виявленню операцій, що можуть бути пов'язані з відмиванням коштів, їх досконале вивчення та вчинення відповідних дій; це система контролю за фінансовими транзакціями для недопущення легалізації та відмивання шахрайських коштів. А легалізація доходів, здобутих незаконним шляхом – це такі дії, заходи, операції коштами, що мають відношення до незаконної діяльності, мають нелегальні джерела походження, приховують незаконне володіння коштами, передбачають незаконне їх переміщення, використання, зміну форми.

Враховуючи вищезазначене, формується трактування поняття системи фінансового моніторингу – сукупності заходів (державного та первинного рівнів), що вчиняються суб'єктами первинного фінансового моніторингу та Спеціально уповноваженим органом з питань фінансового моніторингу в частині ідентифікації, вивчення та аналізу фінансових транзакцій, додаткової інформації про них, про клієнтів, на предмет виявлення відношення до незаконної діяльності, фінансування тероризму, здійснення обліку таких фінансових операцій, оцінювання ризиків від подібних транзакцій. Визначальна роль в системі фінансового моніторингу відводиться суб'єктам перинного фінансового моніторингу, до яких належать [**Ошибка! Источник ссылки не найден.**]: банківські установи, страхові компанії, перестрахові організації, організації – надавачі фінансових послуг, платіжні установи, організації, що виступають членами платіжних систем, ломбарди, кредитні спілки, учасники ринку цінних паперів, товарні біржі, оператори поштових переказів, установи, що проводять валютні транзакції, аудиторські компанії, організації – надавачі бухгалтерських послуг, нотаріуси, адвокати, надавачі юридичних послуг, надавачі послуг по утворенню та управлінню суб'єктів господарювання, агентства нерухомості, суб'єкти господарювання - продавці лотерей, організатори азартних ігор, продавці віртуальних активів, інші надавачі фінансових послуг. Провідна роль належить суб'єктам державного фінансового моніторингу [**Ошибка! Источник**

ссылки не найден.]: Національний банк України, Міністерство юстиції України, Національна комісія з цінних паперів та фондового ринку, Міністерство цифрової трансформації України, Спеціально уповноважений орган.

Організація ефективного фінансового моніторингу передбачає визначення та досягнення стратегічних пріоритетних цілей, тобто: протидія легалізації доході, отриманих незаконним шляхом; боротьба з фінансуванням терористичної діяльності; протидія фінансування розповсюдження зброї масового знищення; належна організація, реалізація та контроль внутрішньобанківської системи боротьби з відмиванням нелегальних коштів, фінансуванням терористичної активності; функціонування належної системи управління системою ризиків у сфері фінансового моніторингу; захист державних, суспільних, громадянських інтересів від збитків процесу легалізації незаконних доходів, фінансування тероризму; дотримання ризик-орієнтовного підходу при здійсненні фінансового моніторингу; забезпечення координованості співпраці учасників системи фінансового моніторингу; вжиття каральних заходів до порушників законодавства у сфері фінансового моніторингу.

Із зазначених цілей фінансового моніторингу формуються задачі фінансового моніторингу, які поставлені для виконання для досягнення ефективності фінансового моніторингу: розробляти та виконувати вимоги законодавчих актів, інших правових документів та внутрішньої нормативної бази згідно фінансового моніторингу; організовувати відповідні структурні органи та підрозділи з фінансового моніторингу на державному рівні, на рівні суб’єктів первинного фінансового моніторингу, на рівні суб’єктів господарювання; організовувати внутрішньобанківську систему фінансового моніторингу; забезпечувати достатність ресурсної бази для належної роботи системи фінансового моніторингу; забезпечувати ефективне функціонування системи заходів управління ризиками відмивання коштів та фінансування тероризму; забезпечувати достатню інформованість, обізнаність, компетентість фізичних осіб, суб’єктів господарювання, працівників суб’єктів первинного фінансового моніторингу та державних органів щодо ризиків відмивання коштів;

організовувати та на постійній основі вдосконалювати внутрішній та зовнішній контроль за протидією легалізації злочинних ресурсів; розробляти та реалізовувати комплекс заходів з належної перевірки клієнтів - проводити поточні та додаткові перевірки клієнта, здійснювати перевірки фінансових операцій клієнта, виявляти невідповідності між фінансовими транзакціями клієнтів та інформацією, що є у банка про таких клієнтів, про специфіку їх бізнесу, суть діяльності клієнта, цілі та очікувань від ділової співпраці клієнтом, досліджувати джерела походження коштів, визначати статки клієнтів, визначати та встановлювати кінцевих бенефіціарів клієнтів, оцінювати економічну та фінансову доцільність транзакцій клієнта; виявляти незвичність дій та операцій клієнта; запобігати використанню банківських послуг для відмивання нелегальних коштів; виявляти порогові фінансові транзакції; ідентифікувати підозрілі фінансові операції; заморожувати активи неблагонадійних клієнтів, шахраїв, терористів; оперативно та своєчасно повідомляти про ризикові та підозрілі фінансові операції Спеціально уповноваженому органу; проводити належний обмін відповідними даними між суб’єктами первинного фінансового моніторингу та Спеціально уповноваженим органом; забезпечувати роботу дієвої системи ескалації певних підозр та загрозливих питань у сфері запобігання відмивання незаконних коштів; своєчасно розглядати підозри та факти порушень сфері боротьби з легалізацію шахрайських коштів; повідомляти про злочинні дії клієнтів правоохоронним органам; своєчасно надавати потрібні дані, роз’яснення, документи, інформацію на запити Національного банку України в частині дотримання фінансового моніторингу; розробляти, використовувати та постійно удосконалювати автоматизовані системи учасників фінансового моніторингу; проводити замороження активів особам, пов’язаним терористичною діяльністю; здійснювати захист даних, відомостей, інформації у сфері фінансового моніторингу; забезпечувати необхідний доступ суб’єктам фінансового моніторингу до потрібної інформації для проведення фінансового моніторингу; організовувати своєчасний обмін інформацією між учасниками системи фінансового моніторингу; організовувати та забезпечувати міжнародне

співробітництво з питань протидії відмивання незаконних коштів а фінансування терористичної діяльності; та ін.

В свою чергу, при формуванні належної системи фінансового моніторингу, виділяються ключові напрями фінансового моніторингу:

- належна перевірка клієнтів (проводиться при встановленні з клієнтом дідових відносин, у разі появи підозр та сумнівів у правдивості поданих клієнтом даних, достовірності інформації щодо суті діяльності та фінансових транзакцій клієнта, у разі проведення разової фінансової транзакції без заключення з клієнтом ділових стосунків, проведенні переказу без відкриття рахунку у сумі 30тис.грн. та більше, здійснення транзакцій по віртуальним активам у сумі 30тис.грн. та більше): ідентифікація (заходи та дії, що проводяться банком для визначення особи клієнта через ідентифікаційні клієнтські дані, до встановлення з потенційним клієнтом ділових відносин), верифікація (заходи та дії банку щодо клієнта для підтвердження відповідності наявних у банку відомостей про клієнта такому клієнту, а також для визначення кінцевих бенефіціарів клієнта, до та під час встановлення з потенційним клієнтом ділових відносин) та відеоверифікація клієнта (верифікація клієнтів шляхом відеотрансляції) – при ідентифікації, верифікації, відеоверифікації клієнтів встановлюються ідентифікаційні дані щодо його найменування паспортних даних, ідентифікаційних кодів, установчих даних, місця реєстрації та знаходження, банківських реквізитів; визначення кінцевих бенефіціарів клієнта (тобто осіб, які мають пріоритетний, вирішальний контроль та вплив на функціонування клієнта та його фінансові транзакції у вигляді прямого володіння часткою 25 та більше відсотків у статутному капіталі, чи у правах голосу, або непрямого впливу володіння менше 25 відсотків у статутному капіталі, чи у правах голосу, але зберігаючи при цьому вирішальну силу голосу); визначення мети та особливостей ділової співпраці з клієнтом (визначити вид банківських послуг, які цікавлять клієнта, особливі умови заключення угод, тарифи, масштаби транзакцій, репутацію клієнта); посилена перевірка клієнта (це заходи з мінімізації рівня ризику, що може бути притаманний клієнту, діловим відносинам з ним; застосовуються до клієнтів з

високим ризиком, клієнтів з підозрами проведення операцій з метою відмивання коштів чи фінансування тероризму; проводиться наступними способами: збільшення ряду дій по перевірці клієнта, актуалізації інформації про нього, та частоти таких дій; затребування у клієнта додаткових даних, роз'яснень, підтверджуючих документів щодо структури власності клієнта, джерел походження коштів, доходів, про наявність ліцензій та дозволів на певну діяльність; пошук додаткових даних про клієнта у відкритих офіційних інформаційних джерелах, тому числі про його господарську діяльність, його кінцевих бенефіціарних власників, про порушені кримінальні провадження, про наявні фінансово-економічні та господарські зв'язки з іншими особами та суб'єктами господарювання; виїзди на місце реєстрації та ведення діяльності клієнта, та ін.); спрощена перевірка клієнта (це заходи з мінімізації рівня ризику, що може бути притаманний клієнту, діловим відносинам з ним; застосовуються до клієнтів з низьким ризиком, до яких належать фізичні особи, які проводять оплату житлово-комунальних послуг на маленькі розміри сум; фізичні особи, у яких відкрито рахунки з виплати соціальних виплат, пенсійних виплат, виплат з оплати праці, стипендій; які проводять раціонально обґрунтовані, звичайні, типові фінансові транзакції у невеликих обсягах; підприємства – надавачі житлово-комунальних послуг, послуг телебачення, інтернет послуг, з якими заключено угоду про приймання платежів; об'єднання співвласників багатоквартирних будинків; юридичні особи та фізичні особи-підприємці при сплаті обов'язкових податкових платежів; державні органи влади; органи місцевого самоврядування; фонди соціального страхування; органи ЄС, дипломатичні представництва членів Організації економічного співробітництва та розвитку; до таких заходів належать: скорочення ряду дій по перевірці клієнта, актуалізації інформації про нього, та частоти таких дій; застосування спрощених методик верифікації клієнтів; скорочення затребування у клієнта додаткових даних, роз'яснень, підтверджуючих документів; застосування відомостей Єдиного державного реєстру юридичних осіб та фізичних осіб підприємців); актуалізація клієнтських даних (тобто актуалізація попередньо одержаних

даних, документації, відомостей, а також теперішньої інформації та документів; актуалізація проводиться з різною періодичністю для різних типів ризику клієнтів – один раз на рік для високого ризику, один раз на три роки для середнього ризику, один раз на п’ять років для низького ризику, або за потребою в певних випадках; актуалізація відбувається у такий спосіб – шляхом заповнення клієнтом при відвіданні банківської установи анкети-опитувальника; шляхом відправки клієнту поштового листа про актуалізацію інформації та документів з анкетою-опитувальником; шляхом відправки клієнту електронного листа про актуалізацію інформації та документів з анкетою-опитувальником; актуалізовані відомості фіксуються у автоматизованій системі баку картці клієнта, а документи підшиваються в особову справу клієнта); відмова банком клієнт від встановлення чи продовження з ним ділової взаємодії, відмова від проведення фінансових транзакцій клієнта (банківській установі необхідно відмовити клієнту у встановленні чи підтриманні ділової співпраці у наступних випадках: якщо клієнт не надав необхідні для ідентифікації, верифікації чи належної перевірки даних; надання клієнтом неправдивих даних; подання клієнтом інформації, яка заплутує банк, вводить його в оману; коли неможливо визначити кінцевих бенефіціарів клієнта; якщо у банківської установи з’являються сумніви стосовно проведення фінансових операцій клієнтом не від власного імені; визначення по клієнту неприйнятно високого рівня ризику; якщо учасник фінансової транзакції банк-оболонка; якщо клієнт підтримує певні відносини з банком-оболонкою; банківській установі заборонено заключати ділову співпрацю з наступними клієнтами: клієнти з Переліку осіб, до яких застосовані спеціальні санкції Міністерства економічного розвитку та торгівлі України; клієнти, які проводять фінансові транзакції, кінцевими вигодоодержувачами яких є особи з Переліку осіб, до яких застосовані спеціальні санкції Міністерства економічного розвитку та торгівлі України; банки-оболонки; особи, які підтримують певні відносини з банками-оболонками; клієнтами, що розташовані чи належать до держав, які не виконують FATF

рекомендацій; клієнти, банки яких знаходяться у країнах, які не виконують FATF рекомендацій);

- управління ризиками фінансового моніторингу: визначення ризиків легалізації незаконних коштів та фінансування тероризму (оцінювання та переоцінювання ризик-портфелю банківської установи (ідентифікація та оцінка ризиків відмивання кошті та фінансування тероризму, що характерні роботі банківської установи шляхом визначення для кожної банківської послуги наявних по ним ризикам в залежності від специфіки, направленості, масштабу функціонування банку в часині направленості на обслуговування фізичних осіб роздрібного бізнесу, суб'єктів господарювання мікро, малого, середнього бізнесу, великих корпоративних клієнтів, послуг, що надаються банком, типів клієнтів, їх ризик-портфелів, географічного признаку банку, способів надання банківських послуг, цільового використання банківських послуг, специфічних можливостей використання певних послуг банку при відмиванні коштів, цільового сегменту для різних послуг, можливих обсягів обігу коштів, інших факторів, важливих для банку; аналіз наявних у банківської установи заходів та методик управління ризиками для їх мінімізації); оцінювання та переоцінювання ризик-портфелю клієнтів (визначення критеріїв ризику; ідентифікація первинного ризику встановлення ділових відносин з клієнтом; застосування скорингової ризик-моделі для оцінювання рівня ризику по клієнту; аналіз наявних у банківської установи заходів та методик управління ризиками для їх мінімізації; оцінювання залишкового ризику від встановлення з клієнтом ділових відносин); розрахунок ризик-апетиту банківської установи (визначення ризиків, що банк може прийняти; ризики, що банк готовий прийняти тільки після їх мінімізації; ризики, що банк прийняти не може – клієнти з неприйнятно високим рівнем ризику, злочинна діяльність, клієнти із санкційних та заборонених списків, клієнти з індикаторами підозрілості, визначеними банком)); заходи по мінімізації ризиків легалізації незаконних коштів та фінансування тероризму (дослідження та аналіз нових послуг на наявність ризиків; введення лімітів та обмежень по послугах; надання дозволу на взаємовідносини з публічними та

пов’язаними з ними особами; надання дозволу на певні ризикові фінансові операції; використання автоматизованих систем для визначення ризиків по клієнтам та їх фінансовим транзакціям згідно критеріїв ризиків; проведення належної перевірки клієнтів для усвідомлення та розуміння суті та особливостей господарської діяльності клієнтів; здійснення постійного вивчення та аналізу інформації про клієнта; постійне дослідження відповідності фінансових транзакцій клієнта суті його роботи; дослідження джерел походження фінансових ресурсів клієнта; особливе вивчення та моніторинг клієнті високого рівня ризику; окреме вивчення неприбуткових та благодійних організацій на предмет можливості їх використання в незаконних цілях для відмивання коштів); належне управління ризиками (проведення комплексної оцінки ризиків відмивання незаконних коштів та фінансування тероризму, та його періодичної переоцінки; здійснення належної перевірки клієнтів; організація належної оцінки та переоцінки ризиків, які можуть виникнути при встановленні чи підтриманні ділових відносин з клієнтами; відповідне усвідомлення банком ризиків легалізації незаконних коштів клієнтів; вжиття диференційованих заходів для клієнтів з різним рівнем ризику; проведення певних дій по приведенню ризиків до прийнятного для банку рівня; розробка та використання дієвих інструментів для перешкоджанню систематичному та великомаштабному проведенню підозрілих фінансових транзакцій; налагодження дієвого внутрішньобанківського контролю, аудиту, ревізій з питань фінансового моніторингу; наявність прозорої системи своєчасного виявлення політично значущих, та пов’язаних з ними осіб; наявність ефективної системи ідентифікації та вивчення кінцевих бенефіціарних власників клієнтів).

- міжнародне співробітництво (передбачає співпрацю за принципом взаємності між різними країнами у сфері попередження, перешкоджання, боротьби з відмиванням незаконних коштів, фінансування тероризму, розповсюдженням зброї масового знищення, в тому числі по питанням: надання пропозицій щодо внесення фізичних осіб та суб’єктів господарювання, та повної інформації про них, до санкційних списків іноземних держав; надання

пропозицій щодо виключення фізичних осіб та суб'єктів господарювання із санкційних списків іноземних держав; приведення виконання рішень судів стосовно конфіскації незаконних доходів; зарахування конфіскованих коштів до державного бюджету; дотримання принципу конфіденційності та таємності інформації; забезпечення дозволу спеціалізованим органам іноземних країн до розкриття певної інформації, та ін.);

- організація та забезпечення відповідальності за порушення нормативно-правових та законодавчих актів з питань фінансового моніторингу в частині протидії легалізації незаконних коштів, фінансуванню тероризму, розповсюдженню зброї масового знищення (визначення відповідальності по певним видам правопорушень, такі як письмові застереження та попередження, відкликання та анулювання ліцензій, відсторонення від робіт, штрафні санкції, ліквідація; врахування певних обставин вчинених порушень щодо характеру порушень, їх тривалості, фінансового стану банківської установи чи іншого суб'єкта первинного фінансового моніторингу, характеру та обсягів вигід від протиправних дій, розмірів отриманих збитків третіх осіб, повторності вчинення однотипного порушення, ступеня відповідальності осіб, готовності співпраці з питань фінансового моніторингу).

Ефективність фінансового моніторингу досягається шляхом застосування дієвих моделей, розроблених сучасними науковими діячами та фахівцями досліджуваного напрямку. Такими моделями є наступні:

- скорингова ризик-модель - бальна модель оцінки ризик-портфелю клієнтів, що використовується банківськими установами з застосуванням на базі наявних програмних модулів; реалізовується шляхом проведення оцінки клієнта банку за визначеними критеріями ризиків. Банком визначаються критерії ризику (визначаються відповідальними працівниками банку згідно вимог законодавства та внутрішньої нормативної бази банку; вносяться або вилучаються з відповідної вкладки критеріїв ризику автоматизованої системи банку та кожному з них присвоюється відповідний код); кожному з критеріїв ризику присвоюються відповідні бали від 0 до 101 одиниці; визначається рівень ризику для кожного з

критеріїв ризику (0 балів - низький рівень ризику, коли клієнт не співпав ні з одним з критеріїв ризику; від 1 до 50 балів - середній рівень ризику; від 51 до 100 балів - високий рівень ризику; від 101 - неприйнятно високий рівень ризику); згідно результатів проведеної оцінки, бали по кожному критерію ризику, що притаманний клієнту, сумуються в автоматичному режимі; визначається сумарна кількість балів для кожного клієнта банку; сумарна кількість балів по клієнту співставляється зі шкалою визначених рівнів ризиків та клієнту присвоюється автоматично рівень відповідного ризик-портфелю клієнта, тобто рівень ризику ділових відносин з клієнтом. Відповідальні працівники банку вносять зміни у картку клієнта інформацію по критеріям ризиків в день появи обставин, які визначають рівень ризика клієнта, або в день одержання відповідної інформації від підрозділу фінансового моніторингу банку, а також один а на квартал за допомогою відповідних програмних комплексів проводять процедури виявлення притаманних клієнтам критеріїв ризиків;

- автоматизована модель фінансового моніторингу - модель автоматизації процесів фінансового моніторингу, таких як: автоматизація ідентифікації, верифікація, відеоверифікація осіб, які здійснюють фінансові транзакції, що підлягають під фінансовий моніторинг; автоматизація бізнес-процесів інтеграції результатів проведення перевірок внутрішнього моніторингу банків із системою держфінмоніторингу; автоматизація бізнес-процесу внутрішніх перевірок фінансових транзакцій, що підлягають під фінансовий моніторинг; автоматизована розробка структури шаблонів вхідної та вихідної документації, пов’язаної із ідентифікацією, верифікацією та відеоверифікацією осіб, які проводять фінансові транзакції, що підлягають фінансовому моніторингу; автоматизована розробка структури баз даних для внутрішнього фінансового моніторингу банків як схему даних з урахуванням ключових елементів та взаємозв’язків, структури нормативно-довідкової інформації, що є необхідною для проведення основних процедур фінансового моніторингу; автоматизоване розроблення структури шаблонів вхідних та вихідних

документів, повідомлень, пов’язаних із початком перевірок та отриманими результатами моніторингу; та ін.

- бізнес-моделі процесів проведення фінансового моніторингу банків та інших економічних агентів – комплексна модель, що включає ряд етапі, на кожному з яких використовується певне моделювання: модель бізнес-процесу автоматизованого внутрішнього моніторингу, що реалізуються безпосередньо самими економічними агентами; модель бізнес-процесу автоматизованого здійснення моніторингу платежів, що проводить фільтрацію фінансових операцій без наявного фінансового підтвердження джерела коштів через систему Інтернет-Клієнт-Банк; бізнес-модель автоматизованого проведення внутрішньобанківського фінансового моніторингу операцій для визначення ризику, що пов'язаний із використанням послуг банку для відмивання коштів;
- та ін.

В загальному розумінні категорія кібербезпека представляє собою комплекс оптимальних заходів, стратегій попередження, захисту, мінімізації загроз, ризиків, втрат від впливу та скоєння кіберзлочинів, кібератак, цифрових нападів на фінансову систему та суспільний добробут, нівелювання шкідливих, несприятливих, небезпечних наслідків для громадян, суб’єктів господарювання, економічної системи, ефективного керівництва, розвитку потенціалу правоохоронних та кримінальних органів, інформаційно-просвітницької суспільної діяльності, національного та міжнародного співробітництва.

Основними стратегічними цілями кібербезпеки виділяють: забезпечення безпечності кіберпростору, дієвості кібероборони, ефективної протидії кіберзлочинам, розробка інструментів кібербезпеки, підтримання кіберстійкості, забезпечення надійного кіберзахисту, підтримання кіберготовності до кібератак, забезпечення безпеки цифрового фінансового ринку, забезпечення інтеграції, координації та співробітництва щодо кібербезпеки.

Поставлені вище цілі визначають важливі до виконання задачі кібербезпеки, такі як: удосконалення законодавчої бази щодо збереження електронної інформації, відомостей про рух електронних даних, збирання,

перехоплення, акумулювання даних, надання та розкриття відомостей відповідним органам, стосовно арешту комп'ютерної інформації; захист від несанкціонованого втручання у функціонування автоматизованих систем, комп'ютерної техніки, інформаційних та комп'ютерних мереж, мереж зв'язку; захист від розробки, використання, купівлі-продажу, розповсюдження шкідливих програмних комплексів та технічних винаходів; захист від несанкціонованої купівлі-продажу, поширення таємної інформації та даних з обмеженим правом доступу; захист від несанкціонованих, протиправних дій посадових осіб з відповідною інформацією; захист від неправильної експлуатації автоматизованих систем, комп'ютерної техніки, інформаційних та комп'ютерних мереж, мереж зв'язку; захист від неправильного користування інформаційними ресурсами; захист від порушення та перешкоджання функціонування автоматизованих систем, комп'ютерної техніки, інформаційних та комп'ютерних мереж, мереж зв'язку; боротьба з незаконними, протиправними діями з фінансовими документами, електронними платіжними засобами, електронними коштами, спеціалізованою технікою для їх виготовлення; прозоре, компетентне та результативне розслідування фінансових online-злочинів, шахрайств електронного, дистанційного банкінгу; криміналізація нападів на інформаційні, комп'ютерні системи, дані, бази, нелегального доступу до них, незаконного втручання у їх функціонування, зловживань ними; криміналізація правопорушень авторських прав; криміналізація комп'ютерних підробок; криміналізація правопорушень незаконного змістовного характеру; розвиток, підтримка та розширення міжнародної інтеграції, співробітництва, партнерства шляхом взаємної допомоги, обміну інформацією, розкриття певних даних, підтримки цілодобових інформаційних мереж, видачі злочинців, надання юридичної та правової допомоги, взаємовизнання судових рішень, неофіційно співпраці органів правопорядку країн світу, згоди на спеціалізовані слідчі дії; ін.

Далі слід зазначити основні напрями кібербезпеки, які розподіляються на певні групи в залежності від категорій спрямованості їх дій: згідно кіберзлочинів – боротьба з посяганнями на конфіденційність, тайну, цілісність, недоступність

комп’ютерної інформації, систем, мереж (боротьба з незаконним доступом, зломом, перехопленням інформації; боротьба зі злочинним втручанням у комп’ютерну систему, створенням перепон її функціонуванню; боротьба з протиправним втручанням у відомості, приховуванням, пошкодженням, погіршенням, порушенням, зміною, видаленням, знищенням даних; боротьба із злочинним використанням комп’ютерних систем, технологій); боротьба з незаконним застосуванням комп’ютерної техніки; боротьба з незаконним змістом даних та відомостей; боротьба з посяганнями на авторські та суміжні права на програмне забезпечення, інформаційні ресурси, бази даних, цифрові продукти та послуги; згідно кримінальних кіберзлочинних правопорушень – боротьба з використанням автоматизованих систем, комп’ютерної техніки, інформаційних та комп’ютерних мереж, мереж зв’язку для підготовки, скоєння, приховування кіберзлочинів; боротьба з передачею даних незаконного характеру шляхом використання автоматизованих систем, комп’ютерної техніки, інформаційних та комп’ютерних мереж, мереж зв’язку; боротьба з незаконною господарською діяльністю, нелегальними фінансовими транзакціями, що проводяться з використанням комп’ютерних мереж; згідно мотивації кібершахраїв – боротьба з кібершахрайством, метою якого є привласнення коштів; боротьба з кібершахрайством, метою якого є привласнення інформації; боротьба з кібершахрайством, метою якого є отримання доступу до автоматизованої системи для нанесення збитків, пошкодження, дезорганізації; згідно кіберзлочинів у банківському секторі – боротьба з банкоматним шахрайством (боротьба зі скімингом, тобто встановленням на банкоматах зчитувальних та копіювальних механізмів для отримання пін-кодів, а також інформації з магнітної смуги чи чіпу; боротьба з «білим пластиком», тобто використанням пустого неідентифікованого пластику для дублювання клієнтських карт; боротьба з шахрайством при відміні транзакції, тобто нібито з технічних причин кошти зараховуються назад на карту, але фізично встигають бути вилученими з банкомату; боротьба з кеш-трепінгом, тобто крадіжкою готівки за допомогою додатково встановлених шахраями спеціальних пристроїв

вилучення готівки); боротьба з кібершахрайством торгівельно-сервісних мереж (боротьба з використанням викрадених, підроблених електронних платіжних засобів, в тому числі пластикових карт; боротьба з привласненням шахраями реквізитів платіжних карт; боротьба з умисним дробленням фінансових операцій і проведенням фінансових транзакцій та суми, менші за граничні з уникненням авторизації та ідентифікації; боротьба фіктивними еквайринговими угодами, метою яких є здійснення операцій з використанням викрадених, підроблених платіжних карт); боротьба з інтернет шахрайством (боротьба з привласненням шахраями через інтернет ресурси реквізитів електронних платіжних засобів та проведенням з їх використанням фінансових транзакцій; боротьба з розробкою, створенням програмного забезпечення, призначеного для викрадення реквізитів електронних платіжних засобів); боротьба з шахрайством у системах банківського дистанційного обслуговування (боротьба з відкриттям рахунків для проведення нелегальних безготівкових та готівкових фінансових транзакцій за допомогою web-банкінгу, mobile-банкінгу, інших систем дистанційного обслуговування; боротьба з розробкою та створенням комп'ютерних вірусів з метою заволодіння системою управління комп'ютером інших осіб для здійснення від їх імені несанкціонованих фінансових транзакцій; боротьба із втручанням в роботу систем міжнародних грошових переказів та міжнародних платіжних систем для отримання несанкціонованих переказів із-за кордону); згідно нормативно-правового забезпечення – забезпечення змін у законодавстві щодо посилення відповідальності осіб за скоєння кіберзлочинів; затвердження законодавчим порядком визнання законної сили електронних доказів по кіберзлочинам; визначення чіткої схеми відносин та відповідальності клієнт-банк, відправник отримувач, у разі протиправного списання клієнтських коштів; законодавчо регламентувати ідентифікацію користувачів мережі Інтернет; законодавчо затвердити необхідність відбирати ідентифікаційні дані надавачем Інтернет послуг при заключенні угод на такі послуги; закріплення інтернет-магазинів за конкретними платниками податків; затвердити необхідність забезпечення захисту систем дистанційного обслуговування декількома рівнями

захисту – логіни, паролі, смс-повідомлення та ін.; закріпити як обов'язкову безкоштовну послугу – надання надавачами фінансових послуг обов'язкового онлайн інформування про всі фінансові транзакції та спроби їх провести не залежно від сум та видів транзакцій; зобов'язання банківським установам здійснювати вихідні платежі лише у межах залишку на рахунку; затвердження незмінного ліміту на зняття готівкових коштів з банкоматі в післяопераційний час; обов'язкова сертифікація усіх електронних платіжних засобів; обладнання банкоматної мережі банків антискімінговими засобами.

Моделі кібербезпеки – це комплексні організаційні, технічні, правлінські, контрольні заходи з забезпечення кібербезпеки, а саме: затвердження певних типових правил та схем для ідентифікації типових, нетипових, підозрілих, сумнівних операцій в системі дистанційного обслуговування; затвердження лімітів операцій в системі дистанційного обслуговування та в мережі інтернет; ведення бази підозрілих та сумнівних клієнтів; ведення бази клієнтів «чорного списку»; оформлення клієнтам пластикових карт чіпом, що має вищий рівень захисту; двоканальна аутентифікація; додаткове підтвердження дистанційних платежі через фінансовий номер телефону; використання захищених спеціальних токенів для електронних цифрових підписів співробітників та клієнтів; забезпечення можливості здійснення генерації електронного ключа особисто клієнтом без участі співробітників банківської установи; забезпечення повідомлення клієнтів про всі фінансові операції та спроби їх проведення; прив'язка електронного цифрового підпису до конкретного переліку серійних номерів комп'ютерної техніки; періодичний аналіз трафіку; систематичний огляд банкоматної мережі з метою ідентифікації сторонніх приладів.

Також зазначимо, що розглядаючи одночасно різні об'єкти та системи, варто виділити вагомість поняття конвергенція, яке означає процес знаходження компромісів, поєднання, зближення відмінних понять. А конвергенція систем передбачає злиття систем у єдине нероздільне ціле; процес їх універсалізації, шляхом поєднання спільних елементів, а в результаті збільшення кількості функцій таких систем, їх можливостей та переваг; виконання елементами систем

різних, але подібних задач за єдиними принципами для отримання додаткового ефекту. Так, результатами конвергенції визначають формування компромісних рішень, досягнення рівноважної позиції, спільний розвиток, загальна стійкість та стабілізація.

При чому конвергенція систем може реалізовуватись на основі різних моделей, таких як: ситуаційно-імітаційно-експертне моделювання (заснована на використання декількох типів підмоделей, передбачає мультиаспектний розгляд досліджуваного питання, дозволяє моделювати як реальну ситуацію, так і штучно створену); модель абсолютної конвергенції (зближення та збільшення рівнів розвитку однорідних об’єктів дослідження без введення додаткових умов); модель умовної конвергенції (передбачає від’ємну залежність між середніми темпами зростання за появи контролюючих факторів); модель прямої конвергенції (процес зближення та компромісів на основі наявних традицій, під впливом визначених факторів); модель непрямої конвергенції (процес зближення на основі нових запозичених понять, інтеграцію об’єктів); та ін.

В свою чергу, конвергенція систем фінансового моніторингу і кібербезпеки передбачає розбудову системи фінансового моніторингу для кібернетичних фінансових операцій; забезпечення кіберзахисту фінансових транзакцій банківської системи; підвищення кіберстійкості фінансової системи до відмивання незаконних коштів та фінансування тероризму. Завдяки конвергенції систем фінансового моніторингу і кібербезпеки формуються наступні новітні заходи, що допомагають у профілактиці, боротьбі та прогнозуванні сучасних фінансових та кіберзлочинів: затвердження на законодавчому рівні правил, інструкцій, протоколів, нормативів, вимог, стандартів, важелів, відповідальності, цінової політики, державних компенсаційних програм з питань фінансової кібербезпеки; затвердження обов’язковості проведення самооцінки стану фінансового кіберзахисту банківським установами; заборона чи обмеження використання іноземних програмних комплексів та систем захисту в національній фінансовій системі; цифрова трансформація, а також перехід на сучасну хмарову середу;

впровадження засобів інформаційної безпеки даних в банківському секторі; запровадження інструментів зі штучним інтелектом; використання багаторівневої системи захисту та боротьби з фінансовими кіберзагрозами; використання новітніх програмних комплексів для захисту операційних систем банків; запровадження міжмережевих екранів; впровадження систем виявлення сторонніх вторгнень в автоматизовану банківську систему; забезпечення посиленого захисту в'язку з віддаленими структурними елементами банківських установ; забезпечення посиленої прозорості, ідентифікації та авторизації при роботі дистанційних онлайн сервісів банків; впровадження новітніх розробок автоматизованих систем реагування на інциденти інформаційної, фінансової та кібербезпеки банків; запровадження роботизації виконання чергових банківських процедур у режимі реального часу; протидія соціальній інженерії, психологічному маніпулюванню людьми щодо здійснення неусвідомлених чи протиправних дій; використання послуг антифрода, послуги для боротьби з фінансовим шахрайством; використання способу збагачення заголовку онлайн запитів користувачів web-ресурсів; а ін.

Питання боротьби з відмиванням нелегальних коштів, протидії фінансування тероризму, розповсюдження зброї масового знищення для України та світу буде гострим ще протягом тривалого часу. Поряд з цим не менш актуальною є проблема кібернебезпеки, коли виникає потреба забезпечувати в цьому напрямку конфіденційність, захищеність, цілісність, доступність та автентичність інформаційних ресурсів. Поглиблені дослідження цих двох векторів передбачають узагальнення, структурування теоретичних надбань світової та вітчизняної літератури в частині визначення основних понять, цілей, задач, напрямів та моделей досліджуваних питань, а також розробки авторами власних висновків по зазначеним аспектам. При чому основний акцент дослідження робиться на те, що обидва комплекси заходів, як фінансовий моніторинг, так і кібербезпека, стають одними з головних завдань світового співтовариства, керівництва країн, державних органів, суспільства.

В загальному підсумку, запропонована в роботі конвергенція систем фінансового моніторингу та кібербезпеки, може бути взята за основу, адаптована та пристосована для розв’язання широкого кола питань як економічної та фінансової безпеки, боротьби з відмивання нелегальних коштів, так і інших проблемних питань фінансового ринку.

1.1.2 Попередній аналіз процесу конвергенції систем кібербезпеки та фінансового моніторингу країн

Забезпечення надійної та потужної системи національної безпеки є одним з пріоритетних завдань для будь-якої країни світу, особливо це є актуальним в умовах зростання рівня цифровізації різних сфер діяльності суспільства та її впливу на економічні, соціальні, політичні та інші процеси. Досягнення відповідного рівня безпеки можливо за рахунок системної взаємодії різних напрямів, одним з яких є формування оптимальної моделі системи державного фінансового моніторингу, що можливо за рахунок посилення її функцій у контексті її конвергенції із системою кібербезпеки. Даний процес є необхідним в умовах зростання інформаційних та кібер-ризиків, які є наслідками інформатизації та цифровізації, а особливо ця потреба відчувається у фінансовій сфері, що є одним з гарантів забезпечення умов надійності фінансово-економічної безпеки країни.

Здійснення конвергенції системи фінансового моніторингу та кібербезпеки повинно відбуватися на всіх рівнях управління економікою, тобто на рівні економічного об’єкта – суб’єкта внутрішнього фінансового моніторингу, так й на рівні держави – суб’єкта обов’язкового фінансового моніторингу. Даний процес повинен передбачати інформаційну, програмну, технічну, організаційну, правову, методичну інтеграції функцій моніторингу та кібербезпеки, результатом чого повинна бути потужна система протидії фінансовим та кібер-злочинам.

У даному контексті виникає потреба оцінити рівень існуючих передумов, сформованих суб’єктами моніторингу та кібербезпеки, до початку імплементації

процесів інтеграції у практичну їх діяльність. Першим кроком є визначення факторів, які ідентифікують рівень країн протидіяти фінансовим та кіберзагрозам. На наступному кроці необхідно провести попередній аналіз процесу конвергенції систем фінансового моніторингу і кібербезпеки, який полягатиме у здійсненні: статистичного аналізу обраних факторів для оцінки однорідності вибірки емпіричних даних; канонічного аналізу для визначення рівня взаємовпливів системи кібербезпеки та фінансового моніторингу; кореляційного аналізу для оптимізації даних. Це сприятиме визначенню факторів, найбільш релевантних для забезпечення означеного процесу.

За останнє десятиріччя зросла кількість наукових праць, присвячених актуальним питанням кібербезпеки та її реалізації у фінансово-економічній сфері. Зацікавленість даною тематикою обумовлена потребами практики щодо посилення захисту інформації та знань фінансового характеру. В цьому контексті слід виділити вектор публікацій, які вирішують проблеми, пов’язані із управлінням кібер-ризиками в банківській сфері, що розглядалися такими науковцями, як Скотт Б.Ф. [**Ошибка! Источник ссылки не найден.**], Ан Дж., Дуань Т., Хоу В., Лю Х. [**Ошибка! Источник ссылки не найден.**], Чен Дж., Чжу К., Башар Т. [**Ошибка! Источник ссылки не найден.**] та іншими.

Наступним актуальним напрямом досліджень є вивчення загроз та вразливостей систем кіберзахисту, що можуть виступити слабкими місцями для кіберзлочинців та кібершахраїв. В цій сфері можна виділити праці таких фахівців, як Бердибаєв Р., Гнатюк С., Євченко Ю., Кіщенко В. [**Ошибка! Источник ссылки не найден.**], Комаров М., Давидюк А., Онискова А., Ткаченко В., Гончар С. [**Ошибка! Источник ссылки не найден.**], Уддін М.Х., Алі М.Х., Хасан М.К. [**Ошибка! Источник ссылки не найден.**] та інші. Застосування сучасних технологій, таких як штучний інтелект та блокчейни, є значним блоком наукових досліджень, спрямованих на вирішення проблеми протидії фінансових та кібер-шахрайств. Цей напрям досліджують Коучоро М.К., Содокін К., Коріко М. [**Ошибка! Источник ссылки не найден.**], Карпуніна Є.К., Михайлов А.М., Бондарева Н.А., Любименко О.А., Федотова Є.В. [**Ошибка! Источник ссылки**

не найден.], Мхланга Д. [**Ошибка! Источник ссылки не найден.**], Сміт К.Дж., Діллон Г. [**Ошибка! Источник ссылки не найден.**], Картер Д. [**Ошибка! Источник ссылки не найден.**] та інші.

Правові та організаційні аспекти, пов'язані із здійсненням процесів кібербезпеки та фінансового моніторингу, є також напрямом наукових досліджень, де розкриваються питання: забезпечення конфіденційності фінансової інформації в праці К. Атта Уль Хак [**Ошибка! Источник ссылки не найден.**]; правових аспектів технологічної нейтральності в статті Г. Гальяні [**Ошибка! Источник ссылки не найден.**]; системного поєднання сфери освіти, технологій та політики для підвищення ефективності системи кібербезпеки в роботі М. Доусон [**Ошибка! Источник ссылки не найден.**]; вимог до організації та функціонування відповідних підрозділів кібербезпеки у фінансовій сфері Т.П. Августінос [**Ошибка! Источник ссылки не найден.**]; тощо. Не дивлячись на широко коло проблем, які вирішуються науковцями – фахівцями в сфері кібербезпеки та фінансового моніторингу, питання конвергенції цих двох систем є ще не розкритим, що потребує подальших досліджень.

Мета дослідження полягає у проведенні попереднього аналізу процесу конвергенції систем кібербезпеки та фінансового моніторингу країн для виявлення найбільш релевантних факторів для їх інтеграції.

Для реалізації поставленої мети даного наукового дослідження було проведено збір та систематизацію статистичних даних в розрізі 76 країн світу за 2018 рік за двома групами показників. Перша група характеризує спроможність країн протидіяти кіберзагрозам за рахунок створення відповідних умов розвитку інформаційних, комп'ютерних та мережевих технологій, а також умов організації ефективної системи кібербезпеки. Дані було узято з офіційного джерела компанії «e-Governance Academy Foundation». Сюди увійшли п'ять індексів: глобальний індекс кібербезпеки (Global Index Cybersecurity); індекс розвитку інформаційно-комунікаційних технологій (ICT Development Index); індекс мережевої готовності (Networked Readiness Index); національний індекс

кібербезпеки (National Index Cybersecurity); рівень цифрового розвитку (Digital Development Level).

Другу групу показників було сформовано з урахуванням існуючих можливостей країн світу щодо формування системи фінансового моніторингу, спроможної протидіяти процесам легалізації кримінальних доходів та фінансування тероризму. Дані було узято з офіційного джерела Світового банку. Сюди увійшли 7 індексів: індекс політичної стабільності (Political Stability Index); індекс ефективності уряду (Government Effectiveness Index); легкість ведення бізнесу (Ease of Doing Business); індекс злочинності (Crime Index); індекс сприйняття корупції (Corruption Perceptions Index); глобальний індекс тероризму (Global Terrorism Index); індекс фінансової таємниці (Financial Secrece Index).

Проведемо за допомогою аналітичного пакету “STATISTICA” статистичний аналіз обраних груп показників, який полягає у визначенні базових статистичних характеристик: середнього значення, медіани, модального значення, мінімального та максимального рівнів, стандартного відхилення та коефіцієнта варіації. Так, його результати в розрізі показників, що ідентифікують систему кібербезпеки, представлені на рисунку 1.1.

	Descriptive Statistics (Spreadsheet1.sta)						
Variable	Mean	Median	Mode	Minimum	Maximum	Std.Dev.	Coef.Var.
Global Cybersecurity Index	66,0789	75,0000	89,0000	2,0000	93,0000	24,2897	36,7587
ICT Development Index	65,0789	69,5000	72,0000	0,0000	90,0000	18,0715	27,7686
Networked Readiness Index	61,8947	63,5000	63,0000	0,0000	86,0000	19,9048	32,1591
National Cyber Security Index	54,2550	57,1400	57,1400	3,9000	96,1000	23,1957	42,7531
Digital Development Level	65,5761	66,8150	58,0000	28,1000	85,1300	13,9731	21,3082

Рисунок 1.1 – Описові статистики групи показників кібербезпеки

Отримані результати дозволяють констатувати, що серед показників кібербезпеки однорідну вибірку мають лише індекс розвитку інформаційно-комунікаційних технологій, індекс мережевої готовності та рівень цифрового розвитку, оскільки значення їх коефіцієнту варіації не перевищує гранично допустимого рівня 33%. В той же час, за показниками глобального індексу кібербезпеки та національного індексу кібербезпеки спостерігається нерівномірність розподілу значень в межах розглянутих 76 країн світу.

Переходячи до аналізу модального значення в розрізі показників (див. рис. 1.1), можна стверджувати, що найбільш поширене значення, яке є найбільшим і незначним чином відрізняється від максимуму, сягає рівня 89 і належить глобальному індексу кібербезпеки. Це свідчить про досить високий рівень даного показника для більшості країн світу. В розрізі інших чотирьох показників кібербезпеки модальне значення приймає значення від 57 до 72 і в усіх випадках перевищує відповідні середні рівні.

Аналогічно, як і для модального значення, глобальний індекс кібербезпеки вирізняється найбільшим середнім рівнем, набуваючи значення 66,08. Найменше значення серед медіанних рівнів, тобто рівнів, що ділять множину розглянутих країн світу навпіл, набуває значення 57 в розрізі національного індексу кібербезпеки.

Проведемо аналіз базових статистик в розрізі показників, які ідентифікують спроможність країн протидіяти процесам легалізації кримінальних доходів. Його результати представлені на рисунку 1.2.

	Descriptive Statistics (Spreadsheet1.sta)						
Variable	Mean	Median	Mode	Minimum	Maximum	Std.Dev.	Coef.Var.
Political stability index	0,322	0,465	,750000	-1,8600	1,54	0,779	241,374
Government effectiveness index	0,633	0,495	Multipl	-1,5800	2,23	0,848	133,954
Ease of doing business	70,199	71,825	Multipl	30,8500	86,59	10,234	14,578
Crime Index	42,055	40,170	Multipl	13,1000	83,60	14,360	34,146
Corruption Perceptions Index	55,342	55,000	Multipl	18,0000	88,00	18,745	33,872
Global Terrorism Index	2,143	1,011	0,00000	0,0000	7,56	2,317	108,107
Financial Secrece Index	284,696	208,255	Multipl	27,8607	1589,57	279,032	98,010

Рисунок 1.2 – Описові статистики показників, які ідентифікують спроможність країн протидіяти процесам легалізації кримінальних доходів

Отримані результати середнього значення, медіани, модального значення, мінімального та максимального рівнів, стандартного відхилення та коефіцієнта варіації дозволяють констатувати, що серед досліджуваних показників тільки в розрізі одного – легкість ведення бізнесу, спостерігається однорідність вибірки для 76 країн світу. Для всіх інших показників виявлено сильно виражену нерівномірність, оскільки коефіцієнт варіації приймає значення від 33,87% (за

показником індекс злочинності) до 241,37% (за показником індекс політичної стабільності).

Отримані модальні значення (див. рис. 1.2) свідчать, що найпоширене значення спостерігається лише за індексом політичної стабільності та глобальним індексом тероризму. В розрізі інших п’яти показників виявлено досить різнорідні значення характеристик спроможності країн протидіяти фінансовим злочинам.

Для виявлення причинно-наслідкових зв’язків між групами показників кібербезпеки та спроможності країн протидіяти процесам легалізації кримінальних доходів проведено канонічний аналіз із використанням аналітичного пакету “STATISTICA”. Його результати представлені на рисунку 1.3.

	Canonical Analysis Summary (Spreadsheet1.sta) Canonical R: ,91259 Chi?(35)=196,50 p=0,0000	
N=76	Left Set	Right Set
No. of variables	5	7
Variance extracted	100,000%	86,6707%
Total redundancy	65,5082%	49,3947%
Variables: 1	Global Cybersecurity Inde	Political stability inde
2	ICT Development Ind	Government effectiveness inc
3	Networked Readiness Inde	Ease of doing busines
4	National Cyber Security Ind	Crime Inde
5	Digital Development Le	Corruption Perceptions Ind
6		Global Terrorism Ind
7		Financial Secrece Inde

Рисунок 1.3 – Результати канонічного аналізу причинно-наслідкових зв’язків між групами показників кібербезпеки та спроможності країн протидіяти фінансовим злочинам

Так, виявлено, що варіативність у множині показників кібербезпеки пояснюється на 65,51% множиною показників спроможності країн протидіяти процесам легалізації кримінальних доходів. В той же час, варіативність у множині показників спроможності країн протидіяти фінансовим загрозам лише на 49,39% пояснюється множиною показників кібербезпеки. Таким чином, результати виявлення причинно-наслідкових зв’язків за допомогою канонічного

аналізу дозволяють констатувати, що показники спроможності країн протидіяти процесам легалізації кримінальних доходів виступають причиною, а множина показників кібербезпеки, відповідно, наслідком.

Крім того, аналіз рисунку 1.3 свідчить також про те, що частка дисперсії (варіативності), яка пояснюється множиною показників кібербезпеки складає 100%, а частка дисперсії (варіативності), яка пояснюється множиною показників спроможності країн протидіяти фінансовим загрозам приймає значення 86,67%. Це говорить про те, що у першому випадку 100% дисперсії будуть пояснювати усі вилучені корені, у другому випадку – на 86,67%.

Канонічна кореляція R=0,91 (див. рис. 1.3), яка відповідає кореляції між першими канонічними змінними, дорівнює максимальному канонічному кореню. Її значення свідчить про наявність сильної лінійної залежності між групами змінних. Статистична значущість коефіцієнта канонічної кореляції підтверджується високим значеннями Хі-квадрату (196,5) та рівнем ймовірності менше ніж 0,05 (p=0,00). Візуалізацію шматково-лінійного графіку спадаючих власних значень, що відповідають канонічним кореням, представимо на рисунку 1.4, а результати тестів Хі-квадрату для статистичної значущості канонічних коренів – на рисунку 1.5.

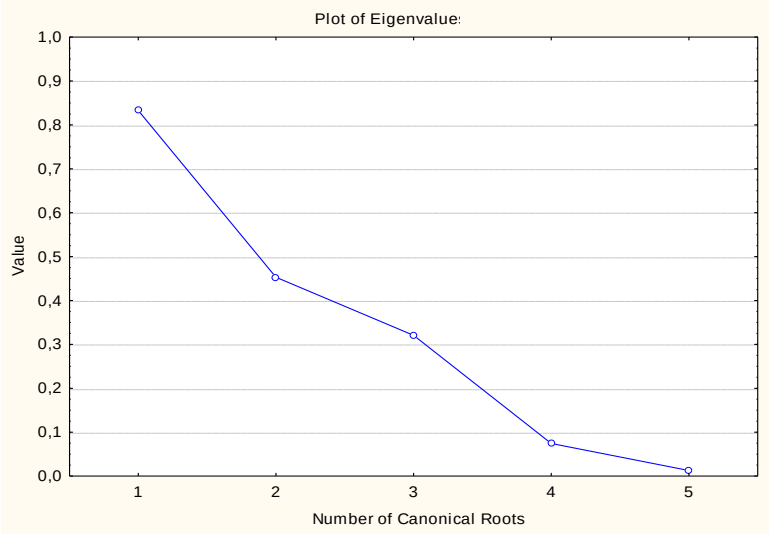


Рисунок 1.4 – Шматково-лінійний графік спадаючих власних значень, що відповідають канонічним кореням

Отже, на основі рисунків 1.4 і 1.5 можна зробити висновок, що статистично значущими є перші три канонічні корені, оскільки p-значення для них не перевищують гранично допустимого рівня 0,05. Саме зазначені канонічні корені пропонується розглядати на наступному етапі при оптимізації вхідного масиву даних.

	Chi-Square Tests with Successive Roots Removed (Spreadsh					
Root Removed	Canonicl R	Canonicl R-sqr.	Chi-sqr.	df	p	Lambda Prime
0	0,91258	0,83281	196,498	35	0,00000	0,05677
1	0,67304	0,45299	73,974	24	0,00000	0,33962
2	0,56621	0,32059	32,648	15	0,00526	0,62087
3	0,27271	0,07437	6,170	8	0,62814	0,91385
4	0,11276	0,01271	0,876	3	0,83108	0,98728

Рисунок 1.5 – Тести Хі-квадрат для статистичної значущості канонічних коренів

З метою оптимізації масиву вхідних даних проведемо із використанням аналітичного пакету “STATISTICA” кореляційний аналіз двох груп показників кібербезпеки та спроможності країн протидіяти фінансовим загрозам. Розглянемо спочатку кореляційну матрицю лівої множини – множини показників кібербезпеки (рисунок 1.6). За результатами аналізу можна зробити висновок про наявність значної кореляційної залежності між такими показниками як індекс розвитку інформаційно-комунікаційних технологій та рівень цифрового розвитку. Підтвердженням даного факту виступає високе значення коефіцієнту кореляції, що дорівнює 0,96. Для оптимізації множини вхідних показників кібербезпеки рекомендується один із колінеарних індикаторів видалити з подальших обчислень.

Root Removed	Correlations, left set (Spreadsheet1.sta)				
	Global Cybersecurity Index	ICT Development Index	Networked Readiness Index	National Cyber Security Index	Digital Development Level
Global Cybersecurity Index	1,00000	0,53583	0,71135	0,70943	0,57919
ICT Development Index	0,53583	1,00000	0,58341	0,64298	0,96073
Networked Readiness Index	0,71135	0,58341	1,00000	0,68127	0,64674
National Cyber Security Index	0,70943	0,64298	0,68127	1,00000	0,65470
Digital Development Level	0,57919	0,96073	0,64674	0,65470	1,00000

Рисунок 1.6 – Кореляційна матриця лівої множини – множини показників кібербезпеки

Для прийняття рішення щодо показника, який варто залишити в масиві вхідних даних, а який треба видалити, розглянемо отриману в результаті проведення канонічного аналізу факторну структуру за першими трьома статистично значущими канонічними коренями (рисунок 1.7).

Variable	Factor Structure, left set (Spreadsheet1.sta)				
	Root 1	Root 2	Root 3	Root 4	Root 5
Global Cybersecurity Index	0,79354	-0,57377	0,03246	-0,19620	-0,03894
ICT Development Index	0,87121	0,17211	-0,39099	0,23550	-0,05499
Networked Readiness Index	0,80257	-0,24082	0,37935	0,35375	0,16974
National Cyber Security Index	0,72570	-0,29619	-0,21889	0,03413	0,58011
Digital Development Level	0,94284	0,25738	-0,19766	0,07562	0,00148

Рисунок 1.7 – Факторна структура лівої множини – множини показників кібербезпеки

Аналіз рисунку 1.7 дозволяє констатувати, що більш значущий вплив здійснює показник рівень цифрового розвитку (0,9428, 0,2573, -0,1977), а ніж індекс розвитку інформаційно-комунікаційних технологій (0,8712, 0,1721, -0,3901) . Відповідно, пропонується залишити індекс рівня цифрового розвитку для проведення подальших досліджень щодо конвергенції систем фінансового моніторингу та кібербезпеки.

Перейдемо до розгляду кореляційної матриці правої множини –показників спроможності країн протидіяти фінансовим злочинам (рисунок 1.8).

Root Removed	Correlations, right set (Spreadsheet1.sta)						
	Political stability index	Government effectiveness index	Ease of doing business	Crime Index	Corruption Perceptions Index	Global Terrorism Index	Financial Secrece Index
Political stability index	1,00000	0,65751	0,45574	-0,49525	0,75031	-0,64890	0,13537
Government effectiveness index	0,65751	1,00000	0,80293	-0,62157	0,90372	-0,04766	0,43522
Ease of doing business	0,45574	0,80293	1,00000	-0,58260	0,64647	0,00231	0,26867
Crime Index	-0,49525	-0,62157	-0,58260	1,00000	-0,55707	0,17321	-0,22725
Corruption Perceptions Index	0,75031	0,90372	0,64647	-0,55707	1,00000	-0,18089	0,34492
Global Terrorism Index	-0,64890	-0,04766	0,00231	0,17321	-0,18089	1,00000	0,21433
Financial Secrece Index	0,13537	0,43522	0,26867	-0,22725	0,34492	0,21433	1,00000

Рисунок 1.8 – Кореляційна матриця правої множини – множини показників характеристики спроможності країн протидіяти фінансовим і кіберзагрозам

За отриманими результатами аналізу можна зробити висновок про наявність значної кореляційної залежності між такими двома показниками, як індекс ефективності уряду та індекс сприйняття корупції. Підтвердженням даного факту виступає високе значення коефіцієнту кореляції, що дорівнює 0,904. Для оптимізації вхідних показників в розрізі спроможності країн протидіяти процесам легалізації кримінальних доходів рекомендується один із колінеарних індикаторів видалити для подальших досліджень.

Для прийняття рішення щодо показника, який варто залишити в масиві вхідних даних, а який треба видалити, розглянемо факторну структуру за статистично значущими канонічними коренями, отриманими в результаті проведення канонічного аналізу (рисунок 1.9).

Variable	Factor Structure, right set (Spreadsheet1.sta)				
	Root 1	Root 2	Root 3	Root 4	Root 5
Political stability index	0,43140	0,61309	-0,53189	0,11795	0,02536
Government effectiveness index	0,95447	0,19149	-0,18007	0,04055	-0,10269
Ease of doing business	0,85417	-0,21703	-0,24630	0,10144	0,26600
Crime Index	-0,55691	0,01299	0,67239	-0,11977	-0,18779
Corruption Perceptions Index	0,81616	0,50482	-0,22105	-0,12459	-0,00193
Global Terrorism Index	0,14948	-0,62097	0,32072	-0,60256	-0,27448
Financial Secrece Index	0,50548	0,08991	0,32266	-0,31996	0,28286

Рисунок 1.9 – Факторна структура правої множини – множини показників характеристики спроможності країн протидіяти фінансовим і кіберзагрозам

Результати аналізу, представленого на рисунку 1.9, показують, що більш значущий вплив здійснює показник індекс ефективності уряду (0,9545 за першим

канонічним коренем), ніж індекс сприйняття корупції (0,8162 за першим канонічним коренем), який і пропонується залишити для проведення подальших досліджень.

Процес конвергенції систем фінансового моніторингу та кібербезпеки є одним з напрямів формування ефективної системи протидії фінансовим та кіберзлочинам в країні. Його реалізація потребує зваженого підходу, оскільки передбачається складна та системна інтеграція багатьох процесів, функцій та механізмів. Тому необхідно оцінити умови, сформовані в країні, які характеризують поточний рівень її кібербезпеки та фінансового моніторингу. Відповідно в даному дослідженні було сформовано дві групи показників, які характеризують для 76 країн світу рівень розвитку окреслених систем за 2018 рік. Сформована база статистичних даних дозволила провести попередній аналіз процесу конвергенції систем фінансового моніторингу і кібербезпеки. В результаті статистичного аналізу проведено оцінку однорідності вибірки емпіричних даних, що дозволило виявити їх неоднорідність для ряду показників. Це обумовлюється нерівномірністю розвитку країн в напрямку забезпечення ефективної системи кіберзахисту та фінансового моніторингу. Проведення канонічного аналізу дозволило встановити, що між групами обраних показників існує тісний зв'язок, при цьому рівень кібербезпеки виступає наслідком, а рівень фінансового моніторингу – причиною. На основі кореляційного аналізу проведено оптимізацію даних, в результаті чого такі показники, як індекс розвитку інформаційно-комунікаційних технологій та індекс сприйняття корупції, слід виключити для проведення подальших досліджень як нерелевантних для розглянутих наборів даних.

В подальшому отримані результати планується використати для проведення біфуркаційного аналізу зрілості діючої системи протидії фінансовим та кібершахрайствам, а також побудови фазових портретів їх «зрілості», «станів рівноваги» та «релаксаційних коливань втрати стійкості».

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**].

1.1.3 Побудова фазових портретів «зрілості», «станів рівноваги» та «релаксаційних коливань втрати стійкості» системи протидії фінансовим та кібер-шахрайствам

Для оцінювання зрілості діючої системи протидії фінансовим та кібер-шахрайствам та побудова фазових портретів їх «зрілості», «станів рівноваги» та «релаксаційних коливань втрати стійкості» побудуємо інтегральний індекс кібербезпеки на основі застосування методу згортки Сундаровського. Зведення індикативних показників до єдиного інтегрального індексу кібербезпеки за допомогою методики Сундаровського передбачає застосування наступної формули (1.1):

$$IS_j = \prod_{i=1}^{n} \left[a_{ij} - a_i^*\right]^{\alpha} \tag{1.1}$$

де IS_j – інтегральний індекс кібербезпеки для j-ої країни;

a_{ij} – фактичне значення і-го показника кібербезпеки для j-ої країни;

a_i^*- рівноважне значення і-го показника кібербезпеки для розглянутої множини країн;

α – константа, показник ступеня.

З метою практичного застосування формули (1.1) для обчислення інтегрального індексу кібербезпеки введемо наступні припущення:

1) в якості рівноважних рівнів складових індикаторів оберемо абсолютне значення різниці середньоквадратичного відхилення та мінімально допустимого рівня:

$$a_i^* = \left|\min_j a_{ij} - \sigma_i\right| = \left|\min_j a_{ij} - \sqrt{\frac{\sum_{j=1}^{m}\left(a_{ij} - \bar{a}_\iota\right)}{n-1}}\right| \quad (1.2)$$

де σ_i – середньоквадратичне відхилення і-го показника кібербезпеки;

$\bar{a}_\iota$ – середнє арифметичне значення і-го показника кібербезпеки;

2) в якості постійного значення показника ступеня функціональної залежності (1.1) оберемо співвідношення одиничного значення та кількості релевантних показників характеристики кібербезпеки. Враховуючи зазначені припущення, формула (1.1) набуває вигляду (1.3):

$$IS_j = \prod_{i=1}^{n}\left[a_{ij} - \left|\min_j a_{ij} - \sqrt{\frac{\sum_{j=1}^{m}\left(a_{ij} - \bar{a}_\iota\right)}{n-1}}\right|\right]^{1/n} \quad (1.3)$$

де n - кількість релевантних показників характеристики кібербезпеки.

Проведемо обчислення за допомогою формули (1.3), представивши поетапно проведені розрахунки в таблиці 1.1: рядок «Рівноважні значення» в розрізі 4 показників кібербезпеки – абсолютні значення $\left|\min_j a_{ij} - \sigma_i\right|$; значення на перетині рядків (країн) та граф (показників кібербезпеки) – значення $\left[a_{ij} - a_i^*\right]$; значення граф IS – результативні величини інтегрального індексу кібербезпеки, визначені методом Сундаровського, в розрізі кожної країни із розглянутої множини 76 країн світу.

Визначення релевантних показників характеристики спроможності країн протидіяти фінансовим і кіберзагрозам за допомогою методу сигма-обмеженої параметризації та Парето-оптимізації. Для реалізації даного етапу використовується інструментарій Statistics, Advanced Linear/Nonlinear Models,

GRM Results. В якості результативної ознаки пропонується обрати інтегральний індекс кібербезпеки, визначений методом Сундаровського, в якості факторів впливу – відповідно, показники характеристики спроможності країн протидіяти фінансовим і кіберзагрозам. Представимо отримані результати на рисунку 1.10.

Таблиця 1.1 – Проміжні та кінцеві розрахунки інтегрального індексу кібербезпеки, визначені методом Сундаровського, за 76 країнами світу

Country	Global Cybersecurity Index	Networked Readiness Index	National Cyber Security Index	Digital Development Level	IS	Country	Global Cybersecurity Index	Networked Readiness Index	National Cyber Security Index	Digital Development Level	IS
Рівноважні значення	22,29	19,90	19,30	14,13							
Australia	2,86	2,77	2,52	2,85	57,03	Liberia	1,07	2,12	0,66	2,26	3,33
Austria	2,79	2,75	2,65	2,83	57,70	Lithuania	2,88	2,66	2,88	2,75	60,61
Bahrain	2,46	2,70	1,61	2,79	29,76	Luxembourg	2,86	2,80	2,56	2,88	58,97
Barbados	1,52	2,11	1,39	2,77	12,32	Malaysia	2,86	2,66	2,70	2,70	55,41
Belgium	2,77	2,75	2,85	2,82	61,32	Malta	2,25	2,65	2,37	2,78	39,17
Bolivia	1,70	2,28	1,75	2,36	15,94	Mauritius	2,85	2,56	2,32	2,61	44,16
Botswana	2,16	2,34	1,29	2,41	15,75	Mexico	2,53	2,47	2,03	2,52	31,91
Brazil	2,44	2,47	2,29	2,59	35,78	Montenegro	2,54	2,53	1,95	2,64	33,17
Brunei Darussalam	2,51	2,11	1,99	2,70	28,56	Netherlands	2,86	2,82	2,81	2,89	65,46
Bulgaria	2,66	2,50	2,39	2,65	42,09	New Zealand	2,74	2,77	2,46	2,86	53,48
Canada	2,86	2,78	2,48	2,84	55,98	North Macedonia	2,76	2,56	2,14	2,62	39,64
Chile	2,16	2,61	2,48	2,68	37,39	Norway	2,86	2,82	2,56	2,89	59,60
China	2,79	2,52	1,99	2,57	36,02	Panama	1,96	2,53	2,32	2,53	29,08
Costa Rica	0,73	2,58	2,41	2,66	12,15	Paraguay	2,48	2,32	2,48	2,36	33,70
Croatia	2,80	2,53	2,83	2,70	54,06	Philippines	2,54	2,47	1,91	2,48	29,63
Cyprus	2,56	2,61	2,17	2,75	39,86	Poland	2,78	2,58	2,67	2,69	51,48
Czech Republic	2,43	2,62	2,92	2,73	50,66	Portugal	2,71	2,66	2,69	2,74	53,06
Denmark	2,81	2,78	2,81	2,89	63,60	Romania	2,43	2,50	2,69	2,63	42,83
Dominica	2,12	2,11	1,98	2,56	22,71	RF	2,80	2,58	2,60	2,70	50,74
Dominican Republic	2,13	2,36	2,17	2,42	26,45	Saudi Arabia	2,85	2,65	2,50	2,70	50,99
Estonia	2,88	2,75	2,91	2,84	65,41	Seychelles	1,39	2,47	1,73	2,51	14,84
Finland	2,83	2,85	2,81	2,87	65,08	Singapore	2,87	2,85	2,80	2,88	65,94
France	2,89	2,74	2,83	2,84	63,45	Slovakia	2,67	2,56	2,78	2,69	51,23

Germany	2,81	2,78	2,80	2,87	62,90	Slovenia	2,63	2,62	2,48	2,74	46,79
Ghana	2,16	2,34	1,86	2,36	22,17	South Africa	2,56	2,52	1,68	2,53	27,30
Greece	2,35	2,50	2,96	2,68	46,64	Spain	2,87	2,65	2,88	2,77	60,68
Grenada	1,70	2,11	1,50	2,57	13,80	Sweden	2,77	2,82	2,48	2,89	55,84
Guatemala	1,28	2,34	1,75	2,29	12,02	Switzerland	2,74	2,82	2,75	2,90	61,78
Hungary	2,77	2,56	2,60	2,68	49,49	Tanzania	2,54	2,14	1,58	1,99	17,16
Iceland	2,18	2,77	2,29	2,89	40,08	Thailand	2,76	2,52	2,20	2,58	39,41
India	2,66	2,42	2,52	2,30	37,28	Trinidad and Tobago	1,35	2,50	1,10	2,60	9,65
Indonesia	2,73	2,47	2,11	2,45	34,80	Turkey	2,81	2,56	2,44	2,63	46,17
Ireland	2,73	2,74	2,58	2,83	54,54	Ukraine	2,57	2,52	2,58	2,58	43,00
Israel	2,73	2,75	2,60	2,83	55,18	UK	2,90	2,80	2,77	2,89	64,85
Italy	2,80	2,56	2,75	2,69	53,19	US	2,90	2,82	2,78	2,87	65,35
Japan	2,85	2,78	2,56	2,87	58,31	Uruguay	2,60	2,58	2,32	2,71	42,02
Kenya	2,69	2,42	1,99	2,29	29,73	Vanuatu	1,87	2,11	1,73	1,93	13,21
Latvia	2,69	2,65	2,69	2,74	52,54	Venezuela	1,89	2,32	1,91	2,45	20,47

Univariate Tests of Significance for (Spreadsheet1.sta)
Sigma-restricted parameterization
Effective hypothesis decomposition

Effect	SS	Degr. of Freedom	MS	F	p
Intercept	15,932	1	15,932	0,1991	0,65683
Political stability index	60,262	1	60,262	0,7531	0,38850
Government effectiveness index	651,47	1	651,47	8,1415	0,00570
Ease of doing business	1068,59	1	1068,59	13,3543	0,00049
Crime Index	197,79	1	197,79	2,4718	0,12047
Global Terrorism Index	185,39	1	185,39	2,3169	0,13254
Financial Secrece Index	63,668	1	63,668	0,7956	0,37549
Error	5521,27	69	80,019		

Рисунок 1.10 - Одномірний тест значущості впливу показників характеристики спроможності країн протидіяти фінансовим і кіберзагрозам на інтегральний індекс кібербезпеки, визначений методом Сундаровського

На основі даних рисунку 1.10, в якій приведені одномірні результати для оцінки ступеня та характеру взаємозв'язку відгука та впливів, можна стверджувати, що статистично значущими виступають лише два впливи: індекс ефективності уряду та легкість ведення бізнесу, оскільки рівні значущості p критерія Фішера для них менше 0,05. Найбільший вклад в загальну модель вносить показник легкість ведення бізнесу, оскільки сума квадратів відхилень SS, яка приймає значення 1068,59, має найбільше значення, а p-значення приймає найменше значення 0,000499. Наступним статистично значущим впливом виступає індекс ефективності уряду, для якого SS=651.48, а p-рівень 0,0057. Наступним за пріоритетністю показником характеристики спроможності країн

протидіяти фінансовим і кіберзагрозам виступає індекс злочинності, хоча для даного показника p-рівень приймає значення 0,12. Візуальним підтвердженням значущості розглянутих факторів виступає діаграма Парето t-значень значущості впливу показників характеристики спроможності країн протидіяти фінансовим і кіберзагрозам на інтегральний індекс кібербезпеки, визначений методом Сундаровського (рисунок 1.11).

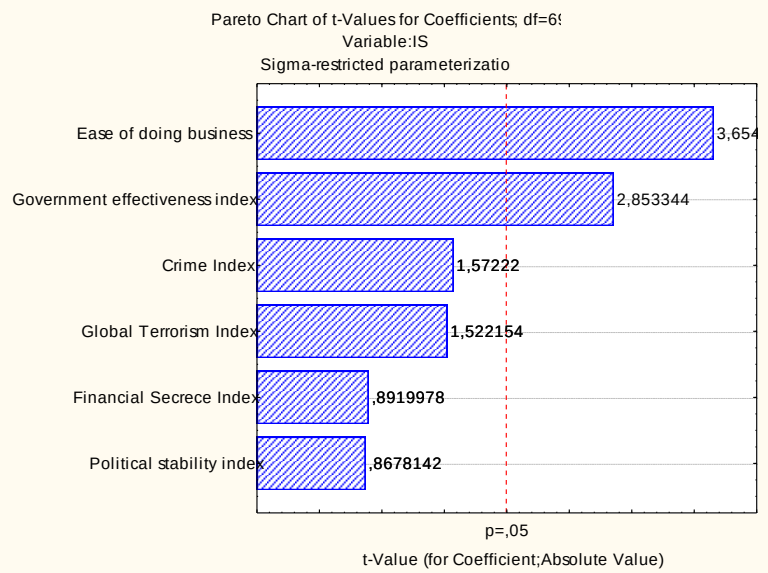


Рисунок 1.11 - Діаграма Парето t-значень значущості впливу показників характеристики спроможності країн протидіяти фінансовим і кіберзагрозам на інтегральний індекс кібербезпеки, визначений методом Сундаровського

Представлена на рисунку 1.11 діаграма Парето дозволяє не просто визначити статистично значущі впливи (регресори) інтегрального індексу кібербезпеки, визначеного методом Сундаровського, але й впорядкувати їх від найбільшого на найменшого впливу. Даний статистичний інструментарій дозволяє графічно проінтерпретувати правило 80 на 20, виділяючи 80% впливових факторів зовнішнього середовища, зокрема: індекс ефективності уряду; легкість ведення бізнесу; індекс злочинності, які і виступають релевантними і пропонується обрати для проведення подальшого дослідження.

Побудова нелінійної регресії з покроковим виключенням залежності інтегрального індексу кібербезпеки від релевантних предикторів: індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності на основі

комбінації логарифмічної та квадратичної функцій, та мультиплікативної залежності трьох показників з метою подальшого проведення біфуркаційного аналізу зрілості діючої системи протидії фінансовим та кібершахрайствам та побудовані фазових портретів їх «зрілості» та «релаксаційних коливань втрати стійкості».

Для реалізації даного етапу пропонується скористатись можливостями програмних пакетів Statistica (Statistics/Advanced Linear/Nonlinear Models/Fixed Nonlinear Regression) та MS Excel (Аналіз даних/Регресія).

Для реалізації даного етапу виникає необхідність визначення специфікації нелінійної регресійної залежності інтегрального індексу кібербезпеки від релевантних предикторів: індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності, для чого скористаємось можливостями програмного пакетіву Statistica, зокрема інструментарію Fixed Nonlinear Regression, який за допомогою методу покрокового включення дозволяє констатувати наявність статистично значущої залежності у вигляді квадратного кореня для індексу ефективності уряду, логарифмічної залежності для показника легкість ведення бізнесу, квадратичної залежності для показника індекс злочинності (рисунок 1.12). В розрізі індексу ефективності уряду, у зв'язку із наявністю від'ємних значень вхідної статистичної бази, пропонується розглянути залежність інтегрального індексу кібербезпеки від даного показника лише у складі мультиплікативної залежності.

Regression Summary for Dependent Variable: IS (Spreadsheet1.sta)
R= ,80929115 R?= ,65495216 Adjusted R?= ,62891081
F(4,53)=25,150 p<,00000 Std.Error of estimate: 9,2027

N=58	Beta	Std.Err. of Beta	B	Std.Err. of B	t(53)	p-level
Intercept			-229,78	67,4183	-3,4084	0,00125
LN-V9	0,42181	0,11168	61,72	16,3445	3,7767	0,00040
SQRV8	0,40110	0,12686	17,08	5,4048	3,1616	0,00259
V10**2	-0,18768	0,08862	-0,003	0,0012	-2,1178	0,03889
1/V8	0,15013	0,09380	0,172	0,1072	1,6005	0,11542

Рисунок 1.12 – Результати регресійної статистики залежності інтегрального індексу кібербезпеки від релевантних предикторів: індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності

Враховуючи результати специфікації залежності інтегрального індексу кібербезпеки від релевантних предикторів: індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності як логарифмічної, квадратичної функцій та мультиплікативної залежності трьох показників, на наступному другому кроці даного етапу проведемо формалізацію зазначеної нелінійної залежності. Так, скористаємось можливостями MS Excel (Аналіз даних/Регресія), обираючи в якості змінних ln(EDI), CI^2, GEI*EDI*CI. Представимо отримані результати у вигляді таблиці 1.2.

Таблиця 1.2 – Результати статистичного аналізу залежності інтегрального індексу кібербезпеки від релевантних предикторів: індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності

Показники	Коефіцієнти	Стандартна похибка	t-статистика	P-Значення	Нижні 95%	Верхні 95%
Y-перетин	-108,69291	56,58893	-1,92074	0,05872	-221,50089	4,11507
ln(EDI)	35,37739	13,25911	2,66816	0,00942	8,94585	61,80893
CI^2	-0,00188	0,00117	-1,60802	0,11221	-0,00420	0,00045
GEI*EDI*CI	0,00277	0,00077	3,58477	0,00061	0,00123	0,00432

На основі даних таблиці 1.2, побудуємо регресійну модель у вигляді формули (1.4):

$$IS = -108.69 + 35.3774 \cdot \ln(EDI) - 0.00188 \cdot CI^2 + 0.00277 \cdot GEI \cdot EDI \cdot CI \qquad (1.4)$$

де IS – інтегральний індекс кібербезпеки;

GEI - індекс ефективності уряду,

EDI – легкість ведення бізнесу,

CI - індекс злочинності.

Статистичну значущість показників ln(EDI) та GEI*EDI*CI підтверджено з рівнем p менше рівня 0,05 та показника CI^2 з рівнем p=0,11. Коефіцієнт детермінації для даної моделі складає 62,73%, фактичне значення критерію Фішера на рівні 40,40 перевищує критично допустимий рівень.

Проведення біфуркаційного аналізу зрілості діючої системи протидії фінансовим та кібершахрайствам та побудовані фазових портретів їх «зрілості» та «релаксаційних коливань втрати стійкості».

Для реалізації даного етапу виникає необхідність попереднього здійснення проміжних обчислень за допомогою застосування апарату диференціального числення, що включає визначення часткових похідних функції залежності інтегрального індексу кібербезпеки від релевантних предикторів: індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності, формування системи диференційних рівнянь, які виступають основою подальшого дослідження динамічної стійкості розглянутої системи (формули 1.5-1.8). Для реалізації даного етапу пропонується застосування пакету прикладних програм MathCAD.

Основою подальшого дослідження динамічної стійкості системи протидії фінансовим та кібершахрайствам та побудові фазових портретів їх «зрілості» та «релаксаційних коливань втрати стійкості» виступає нелінійна функція (1.5):

$$f(gei, edi, ci) := -108.693 + 35.37739 \ln(edi) - 0.00188 ci^2 + 0.002774 gei \cdot edi \cdot ci \qquad (1.5)$$

На основі розглянутої функції (1.5), змоделюємо систему диференціальні рівнянь, які характеризують поводження динамічної системи протидії фінансовим та кібершахрайствам:

$$\frac{d}{dgei} f(gei, edi, ci) \rightarrow 0.002774 ci \cdot edi \qquad (1.6)$$

$$\frac{d}{dedi} f(gei, edi, ci) \rightarrow \frac{35.37739}{edi} + 0.002774 ci \cdot gei$$

$$\frac{d}{dci} f(gei, edi, ci) \rightarrow -0.00376ci + 0.002774edi \cdot gei$$

Представлені три диференційні рівняння (1.6) дозволяють встановити взаємозв'язки між змінними GEI (індекс ефективності уряду), EDI (легкість ведення бізнесу), CI (індекс злочинності) та їх першими частковими похідними $\frac{d}{dgei} f(gei, edi, ci), \frac{d}{dedi} f(gei, edi, ci), \frac{d}{dci} f(gei, edi, ci).$

Ґрунтуючись на нелінійному підході, який лежить в основі теорії біфуркації, побудуємо «фазові портрети» показника інтегральний індекс кібербезпеки у вигляді відображення фазових траєкторій як проекцій на попарно розглянуті площини: індекс ефективності уряду - легкість ведення бізнесу, легкість ведення бізнесу - індекс злочинності, індекс ефективності уряду - індекс злочинності. Побудуємо фазові портрети «зрілості» та «релаксаційних коливань втрати стійкості» системи протидії фінансовим та кібершахрайствам на базі системи диференційних рівнянь (1.6) на основі застосування математичного пакету програмного забезпечення математичного аналізу MathCad:

$$\begin{aligned}
&Faza(gei0, edi0, ci0, dt, N) := \\
&\left|
\begin{array}{l}
\left(gei_0 \leftarrow gei0 \quad edi_0 \leftarrow edi0 \quad ci_0 \leftarrow ci0\right) \\
\text{for} \quad k \in 0..N \\
\quad \left|
\begin{array}{l}
fff \leftarrow f\left(gei_k, edi_k, ci_k\right) \\
gei_{k+1} \leftarrow \left[gei_k + dt \cdot \left(0.002774ci_k \cdot edi_k\right)\right] \\
edi_{k+1} \leftarrow \left[edi_k + dt \cdot \left(\frac{35.37739}{edi_k} + 0.002774ci_k \cdot gei_k\right)\right] \\
ci_{k+1} \leftarrow \left[ci_k + dt \cdot \left(-0.00376ci_k + 0.002774edi_k \cdot gei_k\right)\right]
\end{array}
\right. \\
(gei \quad edi \quad ci)
\end{array}
\right.
\end{aligned} \tag{1.7}$$

З метою надання візуалізації представленого за допомогою формули (1.7) фазового портрету системи протидії фінансовим та кібершахрайствам та подальшої ідентифікації його типу як однієї із можливих альтернатив – різновидів у вигляді сідла, вузла чи фокуса, розглянемо різні варіанти можливих значень як факторів (індекс ефективності уряду, легкість ведення бізнесу, індекс

злочинності), так і значення функції, яка описує інтегральний індекс кібербезпеки із заданим рівнем точності на основі зазначеній кількості точок реалізації:

$$(gei1\ \ edi1\ \ ci1) := Faza(1.6, 80, 42, 0.01, 100) \quad (1.8)$$

$$(gei2\ \ edi2\ \ ci2) := Faza(1.45, 78, 20, 0.01, 100)$$

$$(gei3\ \ edi3\ \ ci3) := Faza(0.18, 68, 36, 0.01, 100)$$

$$(gei4\ \ edi4\ \ ci4) := Faza(0.43, 56, 51, 0.01, 100)$$

$$(gei5\ \ edi5\ \ ci5) := Faza(-0.32, 50, 52, 0.01, 100)$$

$$(gei6\ \ edi6\ \ ci6) := Faza(-0.45, 57, 70, 0.01, 100)$$

Таким чином, підставляючи фактичні значення вхідних даних (формули 1.8) у співвідношення, які дозволяють формалізувати фазові портрети (1.7), зобразимо для прикладу (перше співвідношення формул (1.8)) нелінійну залежність інтегрального індексу кібербезпеки від релеванитиних факторів у площинах «індекс ефективності уряду - легкість ведення бізнесу» (лівий фрагмент рисунку 1.13) та «легкість ведення бізнесу - індекс злочинності» (правий фрагмент рисунку 1.13).

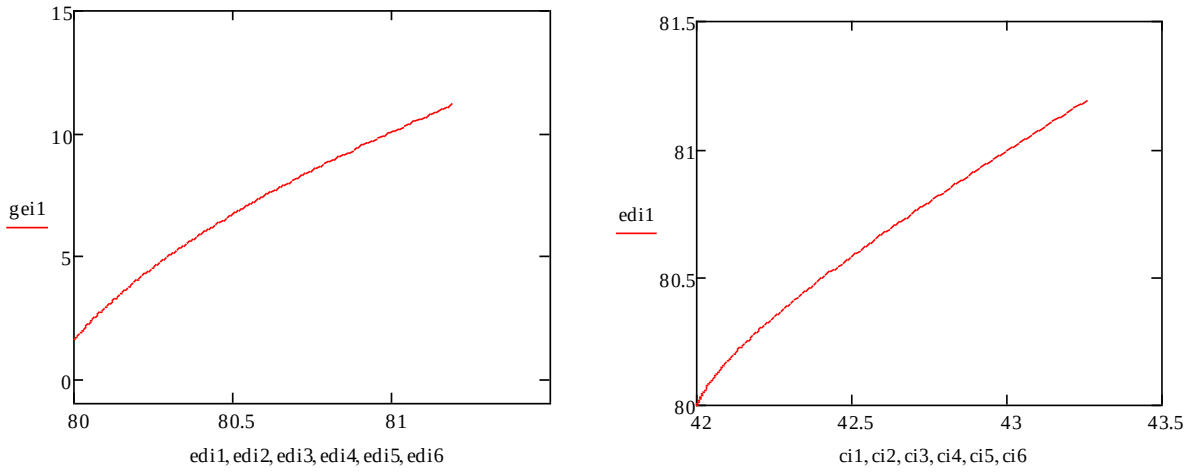


Рисунок 1.13 – Криві нелінійної залежності інтегрального індексу кібербезпеки від релеванитиних факторів у площинах «індекс ефективності уряду - легкість ведення бізнесу» (лівий фрагмент) та «легкість ведення бізнесу - індекс злочинності» (правий фрагмент)

Проведемо дослідження фазового портрету динамічної системи протидії фінансовим та кібершахрайствам на всій множині значень вхідних показників (формули (1.8)). Розглянемо спочатку фрагмент фазового портрету даної системи протидії фінансовим та кібершахрайствам в розрізі площини «індекс ефективності уряду - легкість ведення бізнесу» (рисунок 1.14).

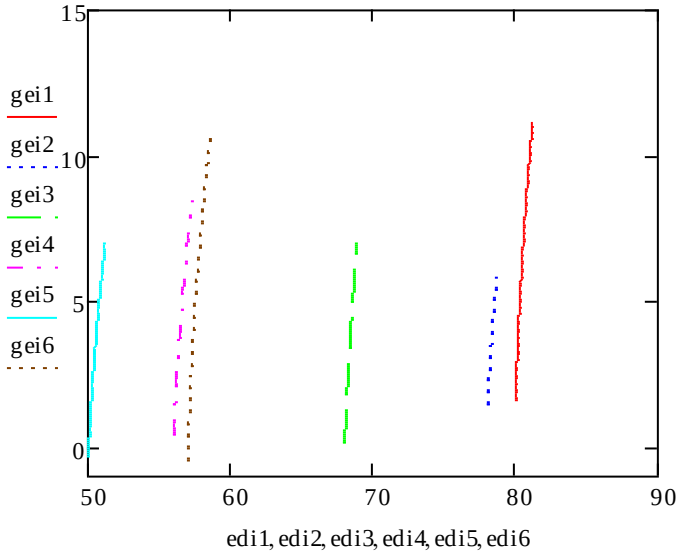


Рисунок 1.14 – Фазовий портрет «нестійкий фокус» системи протидії фінансовим та кібершахрайствам, що знаходиться в нерівноважному стані, в розрізі площини «(індекс ефективності уряду - легкість ведення бізнесу»

Даний «фазовий портрет» демонструє тип біфуркації «нестійкий фокус». Даний тип біфуркації свідчить про нестійкий стан системи, тобто при суттєвій зміні одного параметра і фіксованому значенні іншого параметра розглянута система знаходиться в нерівноважному стані.

Переходячи до розгляду фрагменту фазового портрету динамічної системи протидії фінансовим та кібершахрайствам, в розрізі площини «легкість ведення бізнесу - індекс злочинності» (рисунок 1.15) спостерігаємо, що вона знаходиться в нерівноважному стані, який характеризується як «нестійкий фокус».

Нерівноважний стан динамічної системи протидії фінансовим та кібершахрайствам у вигляді фазового портрету «нестійкий вузол»

спостерігається і в розрізі площини «індекс ефективності уряду - індекс злочинності», що представлено на рисунку 1.16.

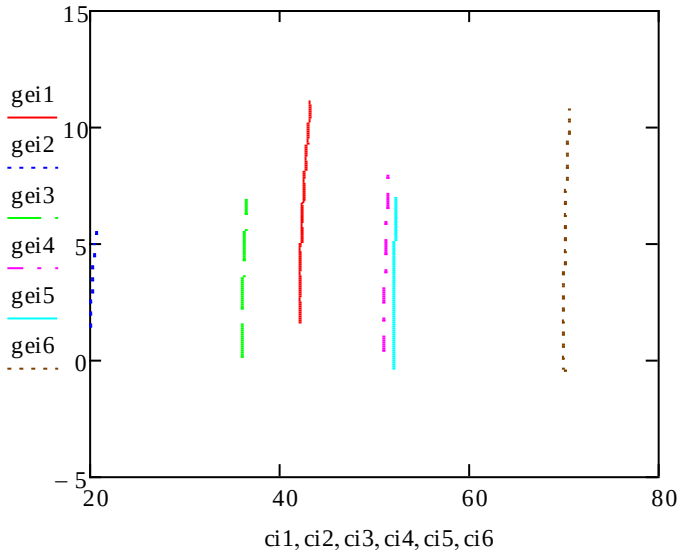


Рисунок 1.15 – Фазовий портрет «нестійкий фокус» системи протидії фінансовим та кібершахрайствам, що знаходиться в нерівноважному стані, в розрізі площини «легкість ведення бізнесу - індекс злочинності»

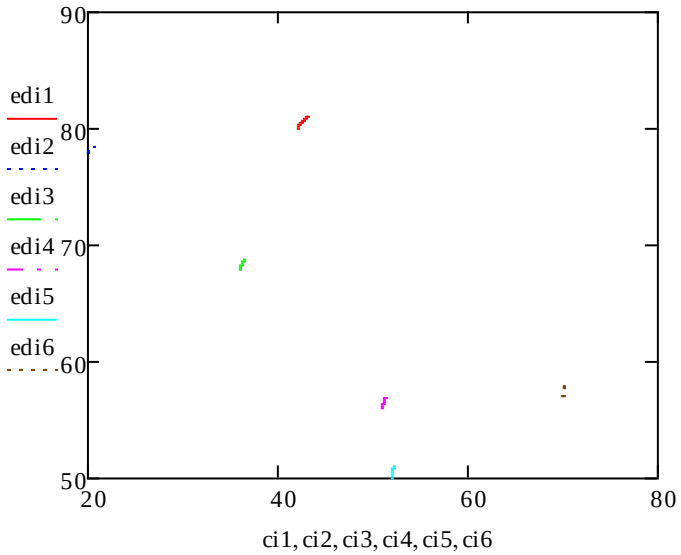


Рисунок 1.16 – Фазовий портрет «нестійкий вузол» системи протидії фінансовим та кібершахрайствам, що знаходиться в нерівноважному стані, в розрізі площини «індекс ефективності уряду - індекс злочинності»

Таким чином, на основі аналізу рисунків 1.14-1.16, які дозволяють провести біфуркаційний аналіз зрілості діючої системи протидії фінансовим та кібершахрайствам та побудувати фазові портрети їх «зрілості» та «релаксаційних коливань втрати стійкості» за допомогою портретів типу «нестійкий фокус» та «нестійкий вузол» в залежності від розглянутої проекції, що свідчать про нерівноважний стан розглянутої системи.

Побудова нелінійної регресії залежності інтегрального індексу кібербезпеки від релевантних предикторів: індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності на основі комбінації степеневої, тригономеричної, та мультиплікативної залежності трьох показників з метою подальшого проведення біфуркаційного аналізу зрілості діючої системи протидії фінансовим та кібершахрайствам та побудовані фазових портретів їх «станів рівноваги». Для реалізації даного етапу пропонується скористатись можливостями програмних пакетів Statistica (Statistics/Advanced Linear/Nonlinear Models/Fixed Nonlinear Regression) та MS Excel (Аналіз даних/Регресія).

Для реалізації даного етапу виникає необхідність визначення специфікації нелінійної регресійної залежності інтегрального індексу кібербезпеки від релевантних предикторів: індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності, для чого скористаємось можливостями програмного пакетіву MS Excel, зокрема інструментарію Аналіз даних/Регресія.

Визначимо специфікацію нелінійної залежності інтегрального індексу кібербезпеки від першої релевантної ознаки індексу ефективності уряду. Для цього розглянемо в якості результативної ознаки інтегральний індекс кібербезпеки, визначений методом Сундаровського, а в якості факторних: поліноміальну (другого і третього ступеня), обернену, тригонометричну залежності індексу ефективності уряду. Застосувавши інструментарій регресійного аналізу отримаємо результат, представлений в таблиці 1.3.

На основі даних таблиці 1.3 (графи p-значення) можна стверджувати, що статистично значущою є змінна sin(GEI), оскільки p-рівнеь приймає значення 0,0298, що менше ніж гранично допустимий рівень 0,05. Саме тому в якості

специфікації залежності інтегрального індексу кібербезпеки від індексу ефективності уряду пропонується в подальших обчисленнях обрати синусоїду.

Таблиця 1.3 – Результати статистичного аналізу залежності інтегрального індексу кібербезпеки від індексу ефективності уряду

Показники	Коефіцієнти	Стандартна похибка	t-статистика	P-Значення	Нижні 95%	Верхні 95%
Y-перетин	47,85679	43,45225	1,10137	0,27456	-38,82807	134,54166
GEI*EDI*CI	0,00035	0,00145	0,24021	0,81088	-0,00255	0,00325
gei2	-5,11919	18,25602	-0,28041	0,78000	-41,53895	31,30057
gei3	1,76326	2,04495	0,86225	0,39154	-2,31630	5,84282
1/gei	0,08037	0,09631	0,83447	0,40690	-0,11177	0,27250
Singei	14,12290	6,36372	2,21929	0,02976	1,42763	26,81816
Cosgei	-16,85691	43,99535	-0,38315	0,70278	-104,62523	70,91140

Визначимо специфікацію нелінійної залежності інтегрального індексу кібербезпеки від другої релевантної ознаки легкості ведення бізнесу. Для цього, як і в попередньому випадку, розглянемо в якості результативної ознаки інтегральний індекс кібербезпеки, визначений методом Сундаровського, а в якості факторних: поліноміальну (другого і третього ступеня), обернену, логарифмічну, квадратний корінь, тригонометричну залежності легкості ведення бізнесу. Застосувавши інструментарій регресійного аналізу отримаємо результат, представлений в таблиці 1.4.

Таблиця 1.4 – Результати статистичного аналізу залежності інтегрального індексу кібербезпеки від легкості ведення бізнесу

Показники	Коефіцієнти	Стандартна похибка	t-статистика	P-значення	Нижні 95%	Верхні 95%
Y-перетин	-215579,91	167427,76	-1,28760	0,20225	-549676,79	118516,97
edi2	5,40	4,03	1,33864	0,18515	-2,64699	13,44
edi3	-0,02	0,01	-1,36340	0,17725	-0,04614	0,01
1/edi	890050,46	692798,38	1,28472	0,20325	-492407,17	2272508,1
ln(edi)	99794,06	77102,97	1,29430	0,19994	-54062,512	253650,63
edi^0,5	-28814,34	22124,27	-1,30239	0,19718	-72962,638	15333,95
sinedi	0,24	1,78	0,13587	0,89233	-3,31319	3,80
cosedi	0,86	1,62	0,53233	0,59623	-2,36443	4,08

На основі даних таблиці 1.4 (графи p-значення) можна стверджувати, що відсутня жодна статистично значуща змінна з рівнем не більше 0,05, але p-рівень приймає найменше 0,1773 значення для кубічної залежності результативної ознаки від змінної легкість ведення бізнесу. Саме тому в якості специфікації залежності інтегрального індексу кібербезпеки від легкості ведення бізнесу пропонується в подальших обчисленнях обрати кубічну залежність.

Визначимо специфікацію нелінійної залежності інтегрального індексу кібербезпеки від третьої релевантної ознаки індексу злочинності. Для цього розглянемо в якості результативної ознаки інтегральний індекс кібербезпеки, визначений методом Сундаровського, а в якості факторних: поліноміальну (другого і третього ступеня), обернену, тригонометричну залежності індексу злочинності. Застосувавши інструментарій регресійного аналізу отримаємо результат, представлений в таблиці 1.5.

Таблиця 1.5 – Результати статистичного аналізу залежності інтегрального індексу кібербезпеки від індексу злочинності

Показники	Коефіцієнти	Стандартна похибка	t-статистика	P-Значення	Нижні 95%	Верхні 95%
Y-перетин	7828,56244	5651,85	1,38513	0,17054	-3449,52	19106,65
ci2	-0,56852	0,40	-1,41177	0,16258	-1,37	0,24
ci3	0,00268	0,00	1,39037	0,16895	-0,00	0,01
1/ci	-24362,11353	18011,51	-1,35259	0,18067	-60303,52	11579,30
ln(ci)	-4624,84426	3329,88	-1,38889	0,16940	-11269,52	2019,83
ci^0,5	1679,55972	1203,49	1,39558	0,16738	-721,97	4081,09
sinci	-3,24185	2,36	-1,37646	0,17320	-7,94	1,46
cosci	5,72062	2,37	2,41124	0,01861	0,99	10,46

На основі даних таблиці 1.5 (графи p-значення) можна стверджувати, що статистично значущою є змінна cos(CI), оскільки p-рівнеь приймає значення 0,0186, що менше ніж гранично допустимий рівень 0,05. Саме тому в якості специфікації залежності інтегрального індексу кібербезпеки від індексу злочинності пропонується в подальших обчисленнях обрати косинусоїду.

Таким чином, визначивши специфікацію залежності інтегрального індексу кібербезпеки від релевантних предикторів (індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності) у вигляді синусоїди, кубічної залежності, косинусоїди відповідно, а також ввівши додатково змінну мультиплікативного впливу на результативну ознаку усіх трьох релевантних факторів, побудуємо відповідну регресійну залежність. Представимо отримані результати в табличному вигляді (таблиця 1.6).

Таблиця 1.6 – Результати статистичного аналізу залежності інтегрального індексу кібербезпеки від релевантних предикторів: індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності

Показники	Коефіцієнти	Стандартна похибка	t-статистика	P-значення	Нижні 95%	Верхні 95%
Y-перетин	10,89788	4,12827	2,63982	0,01019	2,66634	19,12942
singei	9,97709	5,09017	1,96007	0,05391	-0,17241	20,12659
edi3	0,00008	0,00001	5,63831	0,00000	0,00005	0,00010
cosci	3,40130	1,60508	2,11909	0,03758	0,20087	6,60174
GEI*EDI*CI	-0,00057	0,00121	-0,47375	0,63713	-0,00299	0,00184

На основі даних графи «Коефіцієнти» таблиці 1.6 побудуємо регресійну залежність інтегрального індексу кібербезпеки від релевантних предикторів: індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності у вигляді наступного співвідношення (1.9):

$$IS = 10{,}8989 + 9{,}9771 \cdot \sin(GEI) + 0.00008 \cdot EDI^3 + 3.4013 \cdot \cos(CI) - 0.00057 \cdot GEI \cdot EDI \cdot CI \quad (1.9)$$

Достовірність та точність рівняння (9) підтверджено на основі наступних критеріїв. Статистично значущими є значеннями коефіцієнтів перед змінними за допомогою значень p-рівня менше 0,05, окрім коефіцієнту перед змінною мультиплікативного впливу трьох факторів. Але дану змінну пропонується залишити в моделі з метою подальшого проведення біфуркаційного аналізу

зрілості діючої системи протидії фінансовим та кібершахрайствам та побудовані фазових портретів їх «станів рівноваги», оскільки наявність даної змінною мультиплікативного впливу трьох факторів є необхідною умовою проведення якісного біфуркаційного аналізу. Коефіцієнт детермінації приймає значення 70,59%, це свідчить про те, що варіація результативної ознаки інтегрального індексу кібербезпеки на 70,59% пояснюється варіацією обраних факторних ознак.

Проведення біфуркаційного аналізу зрілості діючої системи протидії фінансовим та кібершахрайствам та побудовані фазових портретів їх «станів рівноваги».

Для реалізації даного етапу виникає необхідність попереднього здійснення проміжних обчислень за допомогою застосування апарату диференціального числення, що включає визначення часткових похідних функції залежності інтегрального індексу кібербезпеки від релевантних предикторів: індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності, формування системи диференційних рівнянь, які виступають основою подальшого дослідження динамічної стійкості розглянутої системи (формули 1.10-1.13). Для реалізації даного етапу пропонується застосування пакету прикладних програм MathCAD.

Основою подальшого дослідження динамічної стійкості системи протидії фінансовим та кібершахрайствам та побудові фазових портретів їх «зрілості» та «релаксаційних коливань втрати стійкості» виступає нелінійна функція:

$$f(gei, edi, ci) \coloneqq 10.8978783 + 9.977087\sin(gei) + 7.643510 \cdot 10^{-5} \cdot (edi^3) + 3.40130281\cos(ci) - 0.00057478 gei \cdot edi \cdot ci \quad (1.10)$$

На основі розглянутої функції (10), змоделюємо систему диференціальні рівнянь, які характеризують поводження динамічної системи протидії

фінансовим та кібершахрайствам з метою подальшої побудови фазових портретів «станів рівноваги»:

$$\frac{d}{dgei} f(gei, edi, ci) \to 9.97708769\cos(gei) + -0.00057478ci \cdot edi \tag{1.11}$$

$$\frac{d}{dedi} f(gei, edi, ci) \to 0.000229305edi^2 + -0.00057478ci \cdot gei$$

$$\frac{d}{dci} f(gei, edi, ci) \to -3.40130281\sin(ci) + -0.00057478edi \cdot gei$$

Представлені три диференційні рівняння (1.11) дозволяють встановити взаємозв'язки між змінними GEI (індекс ефективності уряду), EDI (легкість ведення бізнесу), CI (індекс злочинності) та їх першими частковими похідними $\frac{d}{dgei} f(gei, edi, ci), \frac{d}{dedi} f(gei, edi, ci), \frac{d}{dci} f(gei, edi, ci).$

Грунтуючись на нелінійному підході, який лежить в основі теорії біфуркації, побудуємо фазові портрети «станів рівноваги» показника інтегральний індекс кібербезпеки у вигляді відображення фазових траєкторій як проекцій на попарно розглянуті площини: індекс ефективності уряду - легкість ведення бізнесу, легкість ведення бізнесу - індекс злочинності, індекс ефективності уряду - індекс злочинності. Побудуємо фазові портрети «станів рівноваги» системи протидії фінансовим та кібершахрайствам на базі системи диференційних рівнянь (1.12) на основі застосування математичного пакету програмного забезпечення математичного аналізу MathCad:

$$\begin{aligned}
&Faza(gei0, edi0, ci0, dt, N) := \left| \begin{array}{l}
\left(gei_0 \leftarrow gei0 \quad edi_0 \leftarrow edi0 \quad ci_0 \leftarrow ci0\right) \\
\text{for} \quad k \in 0..N \\
\quad \left| \begin{array}{l}
fff \leftarrow f\left(gei_k, edi_k, ci_k\right) \\
gei_{k+1} \leftarrow \left[gei_k + dt \cdot \left(9.97708769\cos\left(gei_k\right) + -0.00057478ci_k \cdot edi_k\right)\right] \\
edi_{k+1} \leftarrow \left[edi_k + dt \cdot \left[0.000229305\left(edi_k\right)^2 + -0.00057478ci_k \cdot gei_k\right]\right] \\
ci_{k+1} \leftarrow \left[ci_k + dt \cdot \left(-3.40130281\sin\left(ci_k\right) + -0.00057478edi_k \cdot gei_k\right)\right]
\end{array} \right. \\
(gei \quad edi \quad ci)
\end{array} \right.
\end{aligned} \tag{1.12}$$

З метою надання візуалізації представленого за допомогою формули (1.12) фазового портрету «станів рівноваги» системи протидії фінансовим та кібершахрайствам та подальшої ідентифікації його типу як однієї із можливих альтернатив – різновидів у вигляді сідла, вузла чи фокуса, розглянемо різні варіанти можливих значень як факторів (індекс ефективності уряду, легкість ведення бізнесу, індекс злочинності), так і значення функції, яка описує інтегральний індекс кібербезпеки із заданим рівнем точності на основі зазначеній кількості точок реалізації:

$$(gei1\ \ edi1\ \ ci1) := Faza(1.6, 80, 42, 0.01, 100) \tag{1.13}$$

$$(gei2\ \ edi2\ \ ci2) := Faza(1.45, 78, 20, 0.01, 100)$$

$$(gei3\ \ edi3\ \ ci3) := Faza(0.18, 68, 36, 0.01, 100)$$

$$(gei4\ \ edi4\ \ ci4) := Faza(0.43, 56, 51, 0.01, 100)$$

$$(gei5\ \ edi5\ \ ci5) := Faza(-0.32, 50, 52, 0.01, 100)$$

$$(gei6\ \ edi6\ \ ci6) := Faza(-0.45, 57, 70, 0.01, 100)$$

Таким чином, підставляючи фактичні значення вхідних даних (формули 1.13) у співвідношення, які дозволяють формалізувати фазові портрети (1.12), визначимо рівноважні точки, представлені у площині «індекс ефективності уряду - легкість ведення бізнесу» рисунку 1.17 для різних значень вхідних даних. Отже, рівноважному стану системи протидії фінансовим та кібершахрайствам відповідають наступні значення її параметрів (точки перетину графіків, зображені на рисунку 1.17): індекс ефективності уряду – 1,4838, легкість ведення бізнесу –80,183.

Проведемо дослідження фазового портрету динамічної системи протидії фінансовим та кібершахрайствам на всій множині значень вхідних показників (формули (1.13)). Розглянемо спочатку фрагмент фазового портрету даної системи протидії фінансовим та кібершахрайствам в розрізі площини «індекс

ефективності уряду - легкість ведення бізнесу» (рисунок 1.18). Даний «фазовий портрет» демонструє наявність рівноважної точки.

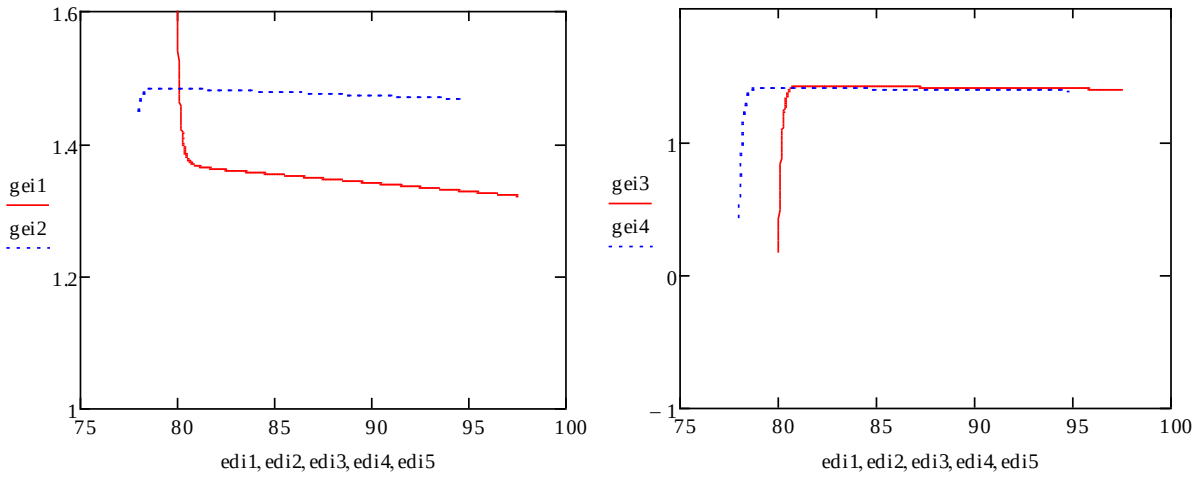


Рисунок 1.17 – Зображення рівноважних точок системи протидії фінансовим та кібершахрайствам у площині «індекс ефективності уряду - легкість ведення бізнесу»

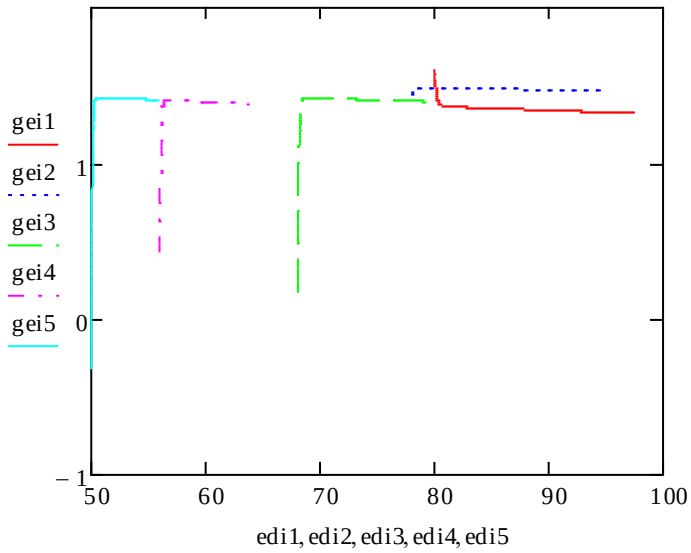


Рисунок 1.18 – Фрагмент фазового портрету «стан рівноваги» динамічної системи протидії фінансовим та кібершахрайствам в розрізі площини «індекс ефективності уряду - легкість ведення бізнесу»

Переходячи до розгляду фрагменту фазового портрету динамічної системи протидії фінансовим та кібершахрайствам, в розрізі площини «легкість ведення

бізнесу - індекс злочинності» (рисунок 1.19) спостерігаємо, що вона знаходиться в нерівноважному стані, який характеризується як «сідло». Даний тип біфуркації свідчить про нестійкий стан системи, тобто при суттєвій зміні одного параметра і фіксованому значенні іншого параметра розглянута система знаходиться в нерівноважному стані.

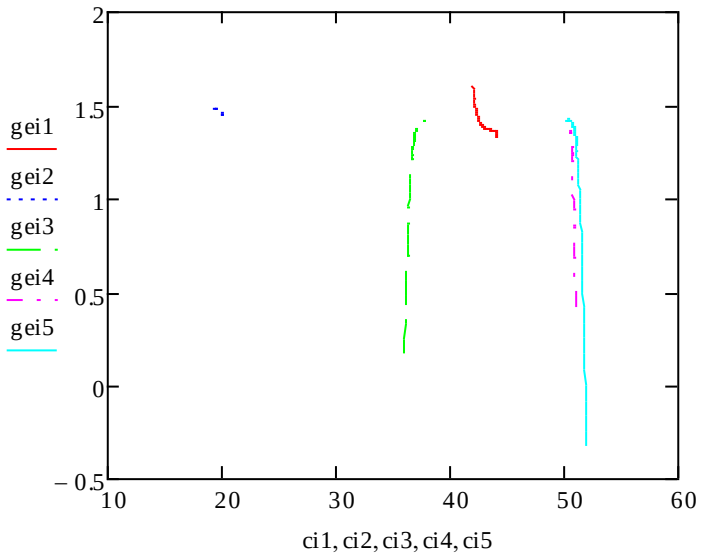


Рисунок 1.19 – Фрагмент фазового портрету «сідло» динамічної системи протидії фінансовим та кібершахрайствам, що знаходиться в нерівноважному стані, в розрізі площини «легкість ведення бізнесу - індекс злочинності»

Нерівноважний стан динамічної системи протидії фінансовим та кібершахрайствам у вигляді фазового портрету «сідло» спостерігається і в розрізі площини «індекс ефективності уряду - індекс злочинності», що представлено на рисунку 1.20.

Таким чином, на основі аналізу рисунків 1.18-1.20, які дозволяють провести біфуркаційний аналіз зрілості діючої системи протидії фінансовим та кібершахрайствам визначено фазовий портрет «станів рівноваги» у площині «індекс ефективності уряду - легкість ведення бізнесу» та нерівноважні фазові портрети типу «сідло» в розрізі інших проекцій.

Метою вищезазначеного аналізу було оцінити зрілість глобальної системи боротьби з фінансовим та кібершахрайством з метою визначення її готовності до інтеграції на різних рівнях державного управління. Оскільки досліджувана

система є динамічною, тобто змінюється під впливом різноманітних зовнішніх і внутрішніх факторів, було проведено біфуркаційний аналіз із побудовою фазових портретів її зрілості та рівноваги.

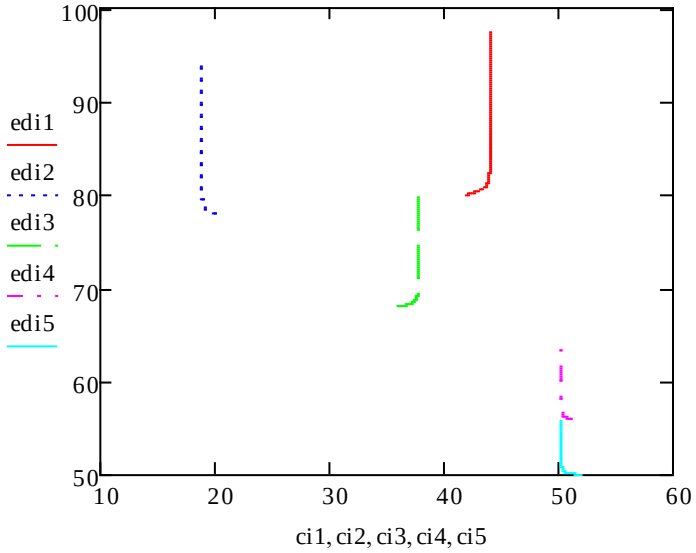


Рисунок 1.20 – Фрагмент фазового портрету «сідло» динамічної системи протидії фінансовим та кібершахрайствам, що знаходиться в нерівноважному стані, в розрізі площини «індекс ефективності уряду - індекс злочинності»

Отримані фазові портрети «зрілості» системи CFCF були класифіковані як «нестабільний фокус» у площинах «Легкість ведення бізнесу – Індекс злочинності» та «Індекс ефективності уряду – Легкість ведення бізнесу». Для площини «Government Efficiency Index – Crime Index» отримано фазовий портрет, класифікований як «нестабільний вузол». Результати показують, що глобальна система CFCF є досить зрілою згідно з отриманими фазовими портретами вузлів і фокусів, але нестабільною. Тобто на нього істотно впливає рівень злочинності в тій чи іншій країні, неефективність державних рішень, відсутність можливостей для розвитку та організації бізнесу. Однак такі фактори, як фінансова таємниця, політична стабільність, рівень корупції та тероризму не викликають коливань. Вони не призводять до істотних змін у системі кібербезпеки. Система CFCF, заснована на інтеграції систем кібербезпеки та боротьби з фінансовими шахрайствами, потребує перш за все законодавчих змін, які мають покращити рівень життя населення та знизити

рівень злочинності в цілому та фінансового та кібершахрайства зокрема. Ще одним стратегічним фактором, який необхідно враховувати, є створення можливостей для розвитку бізнесу, який також позитивно впливає на економічні процеси в країнах та сприяє їх економічному зростанню. Запропонована методологія дозволяє визначити точки, де буде досягнута рівновага системи, але фазові портрети, класифіковані як «сідлові» точки, вказують на те, що система CFCF не може досягти стану рівноваги. (побудований у контексті відповідних трьох площин). Зміна лише одного з параметрів вплине на цей стан за умови, що інший фактор має фіксоване значення. Тим самим підтверджуються попередні висновки про нестійкість системи та її нерівноважність. Підсумовуючи, виявлені стани зрілості та рівноваги системи ФКФК свідчать про її достатній рівень зрілості, але водночас про її неспроможність відновитися в площинах «Індекс ефективності державного управління – легкість ведення бізнесу», «Легкість ведення бізнесу». Doing Business – Crime Index» та «Government Efficiency Index – Crime Index». Тобто існує потреба у вдосконаленні процесів регулювання бізнесу та формування державної політики в країнах світу щодо протидії фінансовим та кіберзлочинам. У майбутньому такий підхід можна рекомендувати відповідним державним органам для формування ініціатив щодо розвитку державного фінансового моніторингу та національної кібербезпеки, а також міжнародним організаціям для вдосконалення стратегії глобальної протидії фінансовим та кіберзлочинам.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**].

1.2 Визначення ключових алгоритмів систем фінансового моніторингу та кібербезпеки

1.2.1 Моделювання ключових алгоритмів конвергенції системи кібербезпеки та фінансового моніторингу у банках

В умовах зростання кількості та різновидів інформаційних та кіберзагроз для

забезпечення функціонування ефективної системи інформаційної безпеки будь-якого суб'єкту економіки є потреба у конвергенції систем. Так, можливим напрямом є інтеграція системи фінансового моніторингу та кібербезпеки, що може здійснюватися на алгоритмічному, програмно-технічному, інформаційному та організаційному рівнях функціонування інформаційної системи банківської установи. Тільки системне поєднання кібербезпеки та фінансового моніторингу дозволить сформувати надійну систему захисту, яка буде не тільки виявляти наслідки, але й попереджувати загрози. Тому вкрай важливим є розуміння сутності та структури процесів забезпечення безпеки інформації, особливо тих, що стосуються заходів перевірок стосовно виявлення порушень цілісності, конфіденційності даних або наслідків кібершахрайств та кіберзагроз.

Пропонуємо розробку трирівневої системи попередження фінансових кіберзагроз, яку буде реалізовано для банківських установ та яка буде охоплювати організаційний, інформаційний та алгоритмічний рівні, заходи кожного з яких будуть спрямовані на виявлення ознак кіберзагроз на етапі, що передує здійсненню зовнішніх та внутрішніх загроз. Концептуальна модель даної системи представлено на рисунку 1.21. Концепція моделі полягає в тому, що операції, які відбуваються у фронт-офісі банку (безпосередньо у банку, за допомогою програмних та мобільних додатків) проходять перевірку щодо наявності властивостей шахрайських операцій. Це вимагає наявності модулю моніторингу – складової інформаційної системи банку, який складатиметься з двох рівнів. Перший передбачає формування бази знань із накопиченням статистичних даних кіберзлочинів. Другий базується на базі моделей – логічних правил, які відслідковують властивості кіберзлочину. Основною ідеєю модуля є використання методів моделювання, а саме інтелектуального аналізу (рисунок 1.21).

Основною метою модуля є знаходження ймовірних фінансових кібершахрайств незалежно від того, хто їх здійснює – клієнт чи співробітник банку. Транзакції перевіряються на їх відповідність кібершахрайським ознакам, зазначеним у базі знань та моделей з урахуванням попереднього досвіду. Означені процеси перевірки відбуваються з урахуванням заходів організаційного рівня, на якому

відбувається оптимізація бізнес-процесів інформаційного захисту, що дозволяє виявляти слабкі місця в системі захисту інформації. Виходячи з окреслених завдань трьох рівнів, розробимо конкретні пропозиції для їх реалізації.

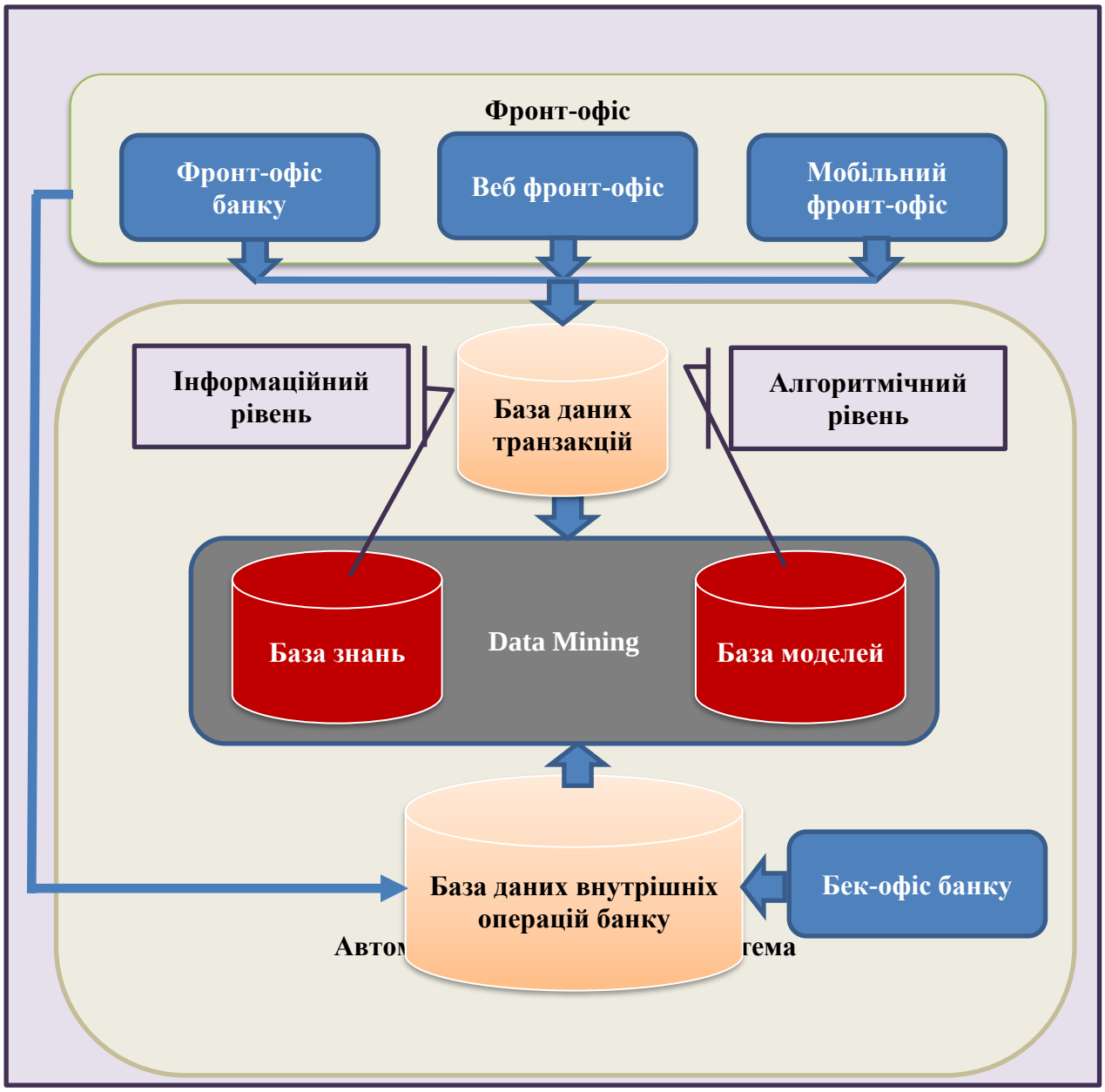


Рисунок 1.21– Концептуальна модель трирівневої системи попередження фінансових кіберзагроз [49, 51, 52]

Для забезпечення організаційного рівня трирівневої системи попередження фінансових кіберзагроз застосуємо методику моделювання бізнес-процесів, яка дозволить побудувати наочну модель будь-якого процесу та провести симуляцію його здійснення на практиці. Як результат, такий підхід сприятиме виявленню

слабких місць та оптимізації з урахуванням різних варіантів.

Методика передбачає побудову та оптимізацію процесів інформаційної безпеки банківської установи, які будуть змодельовані виходячи із можливої інтеграції системи протидії легалізації кримінальних доходів (первинного фінансового моніторингу) та системи інформаційної безпеки (попередження кібершахрайств із зовнішніх та внутрішніх джерел).

Так, на першому кроці будується модель процесу на основі нотації BPMN 2.0, яка є стандартом бізнес-моделювання, що враховує попроцесний підхід. Тобто будь-яка діяльність компанії розглядається не з позиції функцій, з якими вона пов'язана, а з позицій учасників та їх дій, які вони здійснюють протягом певного періоду часу. Це дозволяє бачити – хто виконує, що робить, по відношенню до чого (кого) діє, протягом якого періоду, чим керується. Відповідно в процесі побудови моделі повинні визначатися:

– учасники процесу або його виконавці, які виступатимуть ресурсами компанії, оскільки від їх кількості залежатимуть витрати, пов'язані із процесом. Це можуть бути працівники різних відділів з різними посадами, які приймають рішення, оформлюють документи, здійснюють видачу коштів, вносять дані в систему, контролюють тощо. Також сюди відносяться постачальники, клієнти, банківські установи, та інші, тобто ті, хто є зовнішнім учасником бізнес-процесу. Окремо можна виділити автоматизовані інформаційні системи та їх модулі, які можуть бути також виконавцями за умови автоматизації діяльності. В рамках одного бізнес-процесу може бути задіяно декілька різних учасників;

– операції, тобто конкретні дії виконавців, які здійснює учасник в рамках бізнес-процесу. На практиці вони стосуються конкретного об'єкта та виконуються особою, якій відповідає конкретна посада, а також здійснюються у відповідності з інструкціями установи. Наприклад, дії банківського працівника щодо укладення кредитного договору із клієнтом: вияснити мету отримання кредиту клієнтом; перевірити наявність клієнта в базі даних; ввести дані клієнта, якщо він відсутній у базі даних; перевірити дані клієнта, якщо він є у базі даних; відкоректувати дані; сформувати договір; узгодити умови із клієнтом; роздрукувати та підписати договір;

передати його клієнту, тощо;

– події, які представляють собою дії, що відбуваються з метою ініціалізації конкретної операції процесу. Їх безпосередньо не здійснюють виконавці, оскільки вони можуть відбуватися автоматично або проявлятися у якості певного сигналу, щоб почати або закінчити операцію. Наприклад, початок та кінець бізнес-процесу є основними подіями будь-якого процесу; отримання повідомлення складської системи щодо оприбуткування матеріалів, яке запускає операцію оплати постачальнику, є також подією; відміна операції в результаті помилкового її виконання учасником – це подія, яка буде переривати процес, тощо;

– потоки управління, які дозволяють формувати логіку переходів від однієї операції до іншої. Це відбувається у випадку існування альтернативних варіантів дій учасників, якщо застосовується певна умова, сформована на основі нормативно-правового базису економічного агента (інструкцій, стандартів, законів, положень, тощо). На практиці потоки управління визначаються доволі складно. Це пов’язано із тим, що процес моделювання повинен передбачати різні варіанти дій, а за часту умови їх переходів важко формалізувати. Тому деякі компанії надають перевагу функціональному моделюванню, яке базується суто на посадових інструкціях, де чітко визначені функціональні обов’язки персоналу, та інших документах, пов’язаних із функціональною структурою. Але такий підхід як раз і не дає можливості виділяти дії, які можуть виконуватися в межах однієї функції;

– дані, тобто весь той базис нормативно-правової документації або інформації, що міститься у базі чи сховищі даних, які використовуються для забезпечення виконання певних операцій, подій процесу чи потоків управління, або є їх прямим результатом. Як правило, сюди відносяться бухгалтерські документи, постанови, інструкції, стандарти, закони, положення, масиви, бази, сховища даних, тощо.

Для реалізації моделі застосовується спеціальне програмне забезпечення. Первинна її побудова, яка відображає реальний процес, що відбувається на практиці, називається моделлю “ЯК Є”.

На другому етапі задаються параметри моделі: час на виконання операцій,

вартість ресурсів та ймовірності для потоків управління. Як правило, дана інформація береться, виходячи із наявних даних, що відповідають даному бізнес-процесу. Тобто час задається на основі заміру його фактичних значень, що витрачаються учасниками в процесі виконання ними операцій. Вартість фіксується, виходячи з тарифної сітки учасників або вартісних показників, які символізують витрати, понесені на здійснення тієї чи іншої операції. Ймовірність виставляється також з урахуванням статистичних даних або персональної оцінки учасника процесу.

Для підвищення ефективності моделювання доцільно накопичувати статистику часу та ймовірності для потоків управління. Це дозволить відслідковувати саме ті операції, здійснення яких є найбільш вірогідним та результат несприятливим. У випадку бізнес-процесів банківської інформаційної безпеки, це якраз можуть бути саме ті транзакції, які за певний проміжок часу були відхилені завдяки наявності ознак кіберзагроз або не пройшли первинний фінансовий моніторинг. В подальшому, в процесі оптимізації процесу цей результат може бути враховано для побудови моделі “ЯК БУДЕ”.

На третьому етапі проводиться симуляція за різними типами – “Аналіз часу” та “Аналіз ресурсів”. Результати “Аналіз часу” надають інформацію щодо мінімального, максимального та середнього часу по кожній операції, а також загального часу, витраченого на заданий обсяг симуляції. Так, отримане значення середнього часу по кожній з операцій дозволить виявити слабку ланку, пов’язану із відхиленням від показників по типовим операціям, що сприятиме в подальшому оптимізації даної ділянки процесу.

Також можна отримати інформацію щодо кількості операцій, отриманих на виході та здійснених на кожному вузлі моделі. Так, дана кількість визначатиметься за формулою (1.14):

$$NO_{out} = (((N_0 - \lceil p_1^- \times N_0 \rceil) - \lceil p_2^- \times (N_0 - \lceil p_1^- \times N_0 \rceil)\rceil) - \lceil p_3^- \times ((N_0 - \lceil p_1^- \times N_0 \rceil) - \lceil p_2^- \times (N_0 - \lceil p_1^- \times N_0 \rceil)\rceil)\rceil) - \cdots - \lceil p_n^- \times ((((N_0 - \lceil p_1^- \times N_0 \rceil) - \lceil p_2^- \times (N_0 - \lceil p_1^- \times N_0 \rceil)\rceil) - \lceil p_3^- \times ((N_0 - \quad (1.14)$$

$$\lceil p_1^- \times N_0 \rceil) - \lceil p_2^- \times (N_0 - \lceil p_1^- \times N_0 \rceil) \rceil) \rceil) - \cdots - \lceil p_{n-1}^- \times ((((N_0 - \lceil p_1^- \times N_0 \rceil) - \lceil p_2^- \times (N_0 - \lceil p_1^- \times N_0 \rceil) \rceil)) - \lceil p_3^- \times ((N_0 - \lceil p_1^- \times N_0 \rceil) - \lceil p_2^- \times (N_0 - \lceil p_1^- \times N_0 \rceil) \rceil) \rceil) - \cdots \rceil \Big]$$

де NO_{out} – кількість операцій, отриманих після здійснення симуляції;

N_0 – кількість операцій на початку симуляції;

p_i^- $(\overline{1,n})$ – ймовірність негативного (альтернативного) випадку для потоків управління, коли відбувається розгалуження у моделі;

n – кількість розгалужень у моделі, які позначаються у вигляді шлюзів;

$\lceil \quad \rceil$ – округлення кількості операцій до найближчого цілого у більший бік.

У випадку моделювання основних бізнес-процесів системи інформаційної безпеки рекомендується визначати коефіцієнт результативності за формулою (1.15), який відображатиме якість її роботи:

$$KR = \frac{NO_{out}}{N_0}, \tag{1.15}$$

де KR – це коефіцієнт результативності окремих модулів системи інформаційної безпеки. Якщо KR = 1, то можна сказати, що операцій, які отримали статус загрозливих, шахрайських або підлягають фінансовому моніторингу, не виявлено. Або дійсно не відбувалися такі випадки, або система пропускає всі операції, оскільки має неефективні налаштування перевірок. Якщо KR = 0, то всі операції мають статус загрозливих. На практиці, якщо відбуватиметься така ситуація, то можна сказати, що система є неефективною, оскільки не пропускає всі операції. Граничні значення даного показника свідчать про неефективність роботи системи інформаційної безпеки. Якщо 0<KR<1, то це говорить про те, що деякі операції було заблоковано системою у зв’язку із знаходженням в них ознак загроз.

Системи, які використовують банки для фінансового моніторингу, можуть

блокувати операції, які при подальшій їх перевірці не виявляють ознаки відмивання кримінальних доходів. Це можна пояснити тільки тим, що використовуються непрозорі критерії перевірки, тому система автоматично відносить такі транзакції в категорію підозрілих. Використання індексу (2.2) дозволить накопичувати статистику результативності системи та у випадку помилкового відбору операцій здійснювати коригування критеріїв перевірок.

Результати симуляції “Аналіз ресурсів” надають інформацію щодо завантаженості кожного з виду задіяних ресурсів та їх вартості. Це дозволяє сформувати уявлення щодо фінансових витрат, пов’язаних із виконанням даного процесу. Якщо задіяно декілька учасників (ресурсів), то можна визначити відповідні витрати на кожного з них окремо та порівняти вартісні показники у випадку вибору альтернатив, що дозволить визначити шляхи економії.

На четвертому етапі проводиться оптимізація бізнес-процесу шляхом внесення змін та коректувань у модель, які будуть враховувати слабкі місця, виявлені в результаті здійснення симуляції на попередньому кроці. Тобто будується модель “ЯК БУДЕ”, яка буде відображати бажані елементи процесу. Далі здійснюється процес налаштування симуляції (другий етап) та сама симуляція (третій етап). Отримані результати порівнюються із результатами для моделі “ЯК Є”. Це стосується даних часу та вартості ресурсів. Якщо значення показників покращилися, то отримана модель буде вважатися придатною для практичного використання. Якщо показники після оптимізації не змінилися у найкращий бік, то оптимізацію проводимо ще раз. Це відбуватиметься доти, доки результати моделювання не будуть придатними для практичного застосування.

Отримані моделі бізнес-процесів впроваджуються у діяльність банку або іншого економічного агента. Тобто внесені корективи запроваджуються до тих операцій та учасників, які було оптимізовано у моделі.

Дану методику використаємо для побудови моделей бізнес-процесів, які сьогодні є найбільш критичними для системи інформаційної безпеки банків: процес ідентифікації та верифікації клієнта; процес перевірки транзакцій на наявність ознак кібершахрайств; автоматизованого фінансового моніторингу;

перевірки дій інсайдерів на ознаки кібершахрайств. Для моделювання було використано програмне забезпечення Bizagi Modeler. Проводимо моделювання тих процесів, які задіяні безпосередньо у системі банківської безпеки. На рисунку 1.22 представлено бізнес-модель процесу ідентифікації та верифікації клієнта, яка є придатною у випадку дистанційних операцій. Вона вже є результатом "ЯК БУДЕ".

Оскільки на практиці проводиться ідентифікація клієнтів, а верифікація здійснюється тільки для окремих операцій, то побудова моделі "ЯК Є" буде недоцільною в даному випадку, оскільки вона не враховуватиме багатьох параметрів та порівняння покаже неефективність моделі "ЯК БУДЕ" за рахунок її більш складної структури. Ця проблема буде стосуватися й інших запропонованих моделей, тому аналіз та порівняння буде проводитися для повністю автоматизованого процесу та процесу, де частина операцій виконується людиною, що є характерним для багатьох українських банків.

В моделі зазначено два етапи, які повинен пройти клієнт. На першому відбувається його ідентифікація, коли він входить у систему через мобільний додаток, або веб-банкінг, або термінал. Це здійснюється шляхом виконання запиту на підтвердження особи клієнта шляхом використання відбитка пальця, сітківки ока, або підтвердженням через СМС-повідомлення або телефонний дзвінок. Зараз в Україні використовується тільки два останні види підтвердження. У випадку, якщо шахрай намагається увійти до системи, використовуючи чужі дані, то ідентифікацію буде не пройдено, а операцію заблоковано.

Після успішного підтвердження, починається другий етап – верифікація, тобто здійснюється перевірка клієнта на наявність у (рисунок 1.22): «чорному списку» банку, де він є клієнтом, та у «чорних списках» інших банків; реєстрі судових рішень по клієнту; реєстрі осіб, які переховуються від органів української влади; базі розшуку ФБР; Єдиному реєстрі боржників; базі даних втрачених паспортів; базі даних офшорних компаній; Єдиному державному реєстрі підприємств, щодо яких порушено впровадження у справі про банкрутство; базі даних осіб, до яких застосовано спеціальні санкції Мінеконом розвитку України.

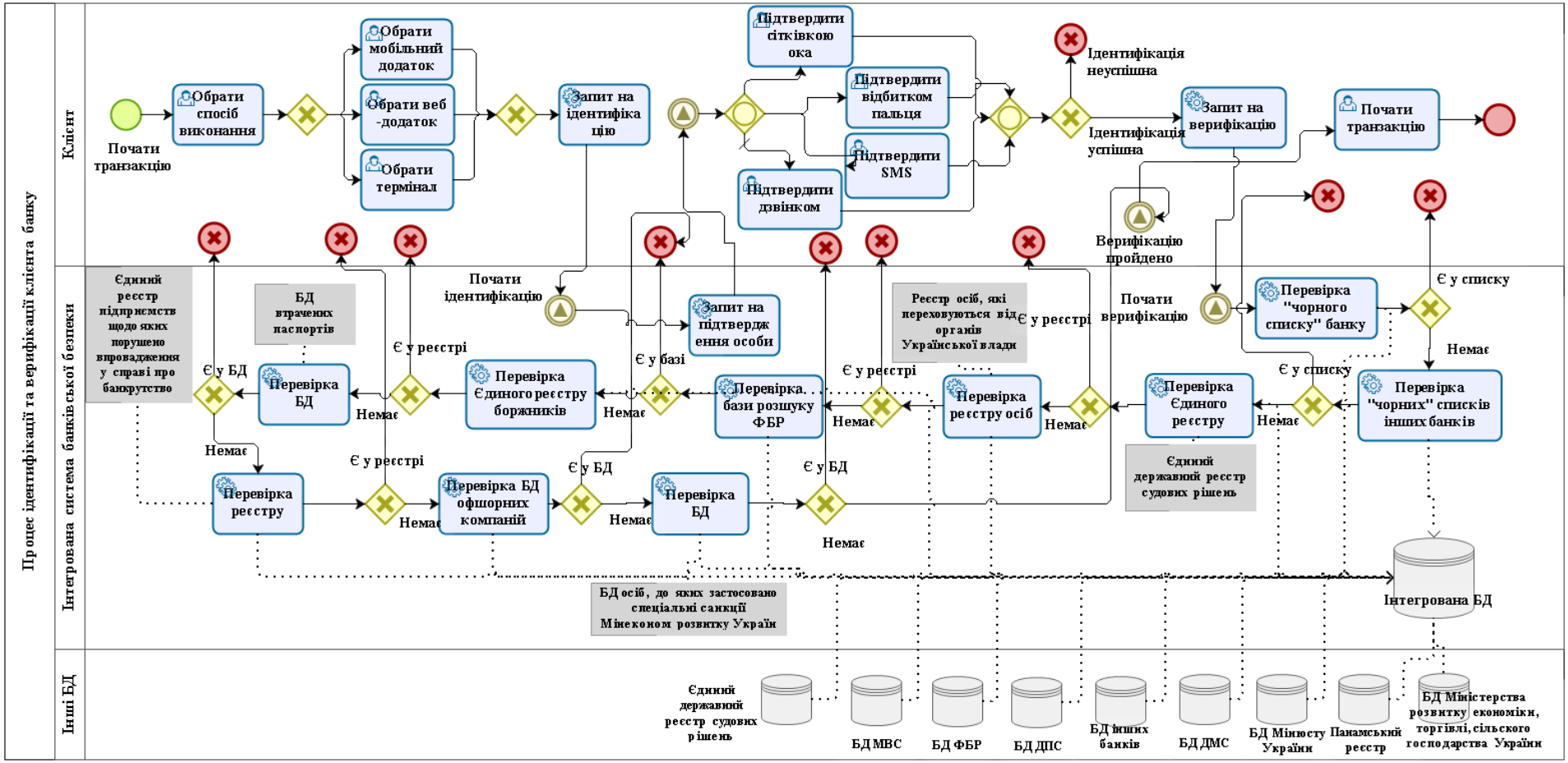


Рисунок 1.22 – Бізнес-модель процесу ідентифікації та верифікації клієнта в інтегрованій системі банківської безпеки

[Ошибка! Источник ссылки не найден.]

Якщо клієнт успішно проходить верифікацію, то система надає йому дозвіл на здійснення операції, в протилежному випадку система його блокує та повідомляє відповідні органи безпеки.

Результати проведеної симуляції для даного бізнес-процесу представлені на рисунку А.1 у додатку А. В якості умов симуляції було задано: кількість операцій – 1000; ймовірність відхилення операції у випадку не проходження перевірки – 1% по кожному вузлу (бажаний показник); час на виконання операцій в інтегрованій системі банківської безпеки – 1 с. (це максимальний час на виконання 1 запиту у будь-якій системі) [**Ошибка! Источник ссылки не найден.**]; для операцій ідентифікації час виставлявся, виходячи із власних замірів максимального часу в процесі користування мобільним банкінгом. В результаті отримано, що середній час на ідентифікацію та верифікацію клієнта за умови впровадження даної схеми на практиці буде дорівнювати 69,75 с. Кількість операцій після перевірки – 903 (формула 2.1). Коефіцієнт результативності – 0,903 (формула 2.2). Тобто за умови визначення 1% операцій такими, які носять ознаки шахрайських по кожному з критеріїв перевірки, система буде ідентифікувати після верифікації та ідентифікації 90,3% операцій як таких, що пройшли моніторинг.

Після проходження ідентифікації та верифікації пропонується перевірка операцій у відповідності із їх сумами. Якщо сума транзакції перевищує 400 000 грн., то банк зобов’язаний здійснити її моніторинг за критеріями на предмет легалізації кримінальних доходів [**Ошибка! Источник ссылки не найден.**]. В протилежному випадку, рекомендується перевірити на наявність ознак шахрайства. Це актуально в умовах зростання кількості постраждалих від соціальної інженерії. Так, на 4 квартал 2019 року цей вид злочинності у поєднанні із шкідливим програмним забезпеченням використовувався у 54%. Для приватних осіб соціальна інженерія склала 67%, для приватних – 62%. При чому для різних компаній його доля є значною: для державних компаній – 66%, промислових – 88%, фінансових організацій – 94%, ІТ-компаній – 50%, торгівля – 36% [**Ошибка! Источник ссылки не найден.**].

На практиці установи зобов’язані здійснювати моніторинг, але процес перевірки організується банками самостійно. Тому більшість з них його проводить

вручну. Згідно із Постановою НБУ №65 від 19.05.2020 «Про затвердження Положення про здійснення банками фінансового моніторингу» налаштування та автоматизацію відповідних процесів банки повинні організувати до 30.06.2021 року [**Ошибка! Источник ссылки не найден.**].

Науковці різних країн світу пропонують власні підходи до організації автоматизованого моніторингу. Так, авторським колективом Чен З., Ван Хоа Л.Д., Тео Е.Н. , Назір А., Каруппія Е.К., Лам К.С. досліджено техніки машинного навчання, як засіб протидії відмивання коштів [**Ошибка! Источник ссылки не найден.**]. Авторами Гао С., Сю Д., Ванг Х., Грін П. розроблено мультиагентну систему з використанням технології інтелектуальних агентів, яка може бути інтегрована в бізнес-процеси банку для виявлення операцій, пов'язаних з відмиванням грошей [**Ошибка! Источник ссылки не найден.**]. Робота Дівії Е. та Умадеві П. присвячена розробці інформаційної моделі, яка базується на аналізі потоку транзакцій, що дозволяє здійснювати кластеризацію банківських операцій з точки зору ймовірності відмивання грошей [**Ошибка! Источник ссылки не найден.**].

Цікавий підхід представили у своїй роботі Калдера Х., Хейн Д. та Шерлок К., які запропонували платіжну систему з доповненим автоматизованим функціоналом протидії відмиванню незаконно отриманих коштів, яку було ними запатентовано [**Ошибка! Источник ссылки не найден.**]. Колхаткар Д., Фатнані С., Яо Ю. та Мацумото К. представили та запатентували багатоканальну систему протидії легалізації коштів для платіжних карт, яка здійснює моніторинг операцій у режимі реального часу [**Ошибка! Источник ссылки не найден.**]. В роботі Діонісія С. Деметиса розглянуто сучасний напрямок реалізації сучасних систем протидії відмиванню коштів (Anti-money laundering), які базуються на підходах визначення ризиків [**Ошибка! Источник ссылки не найден.**]. У дослідженні Коельо Р., Де Сімоні М. та Преніо Дж. представлений новий напрямок “Suptech”, який є передовим інструментом збору даних та їх аналізу на основі штучного інтелекту та машинного навчання, який застосовується у боротьбі з легалізацією кримінальних доходів [**Ошибка! Источник ссылки не найден.**]. У праці Йонг Лі висвітлені аспекти

технічної реалізації AML-інформаційних систем, особливо планування їх впровадження, проектування, аналізу поточного та майбутнього стану, деяких технічних рішень та практичних підходів [**Ошибка! Источник ссылки не найден.**].

Не дивлячись на значний вклад закордонних вчених у вирішення проблеми протидії відмивання коштів, вітчизняна наука відстає в питанні створення, розвитку, удосконалення інформаційних систем та технологій моніторингу, які використовуються для виявлення кримінальних доходів в процесі їх легалізації. Тому вирішення даного питання є досить актуальним для економіки та наукової спільноти України. Практичного досвіду вітчизняних банків пропонується бізнес-модель процесу первинного фінансового моніторингу банку, який здійснюється в умовах автоматизованої обробки інформації (рисунок 1.23). Запропонована модель (рисунок 1.23) демонструє здійснення автоматизованого моніторингу за 13-ма показниками. Якщо операція не проходить хоча б одну із запрограмованих перевірок, система її блокує та вводить до бази даних запис про ризик, пов'язаний із здійсненням даної транзакції, після чого дані надсилаються до Держфінмоніторингу. У разі проходження транзакцією всіх етапів перевірки, приймається рішення щодо обслуговування клієнта та ухвалення даної операції.

Впровадження запропонованої автоматизованої системи моніторингу дозволить розвантажити працівників фронт-офісу щодо перевірки потенційних операцій, пов'язаних з відмиванням грошей. Також її функціонування сприятиме підвищенню ефективності роботи персоналу банку під час проведення фінансового моніторингу. По-перше, це дозволить здійснювати онлайн-перевірку транзакцій на постійній основі. По-друге, вплив працівника на процес перевірки та приховування чи спотворення його результатів більше не буде можливим. Це відбудеться тому, що система передбачає застосування логіки бізнес-правил, яка сприятиме автоматичному вибору тих операцій, які не відповідають заданим умовам. Адміністратор системи несе відповідальність за їх налаштування, а інші банківські працівники не матимуть достатніх прав для цілеспрямованого впливу на процес верифікації. По-третє, запропонована система дозволяє перевіряти більші обсяги операцій щодо їх участі у відмиванні

грошей та фінансуванні тероризму. Наприклад, оскільки обов’язковий моніторинг застосовується до операцій, сума яких перевищує 400 000 гривень, то операції з меншими сумами, які можуть мати кримінальні джерела походження та приймати участь у схемах з відмиванням, залишаються поза увагою.

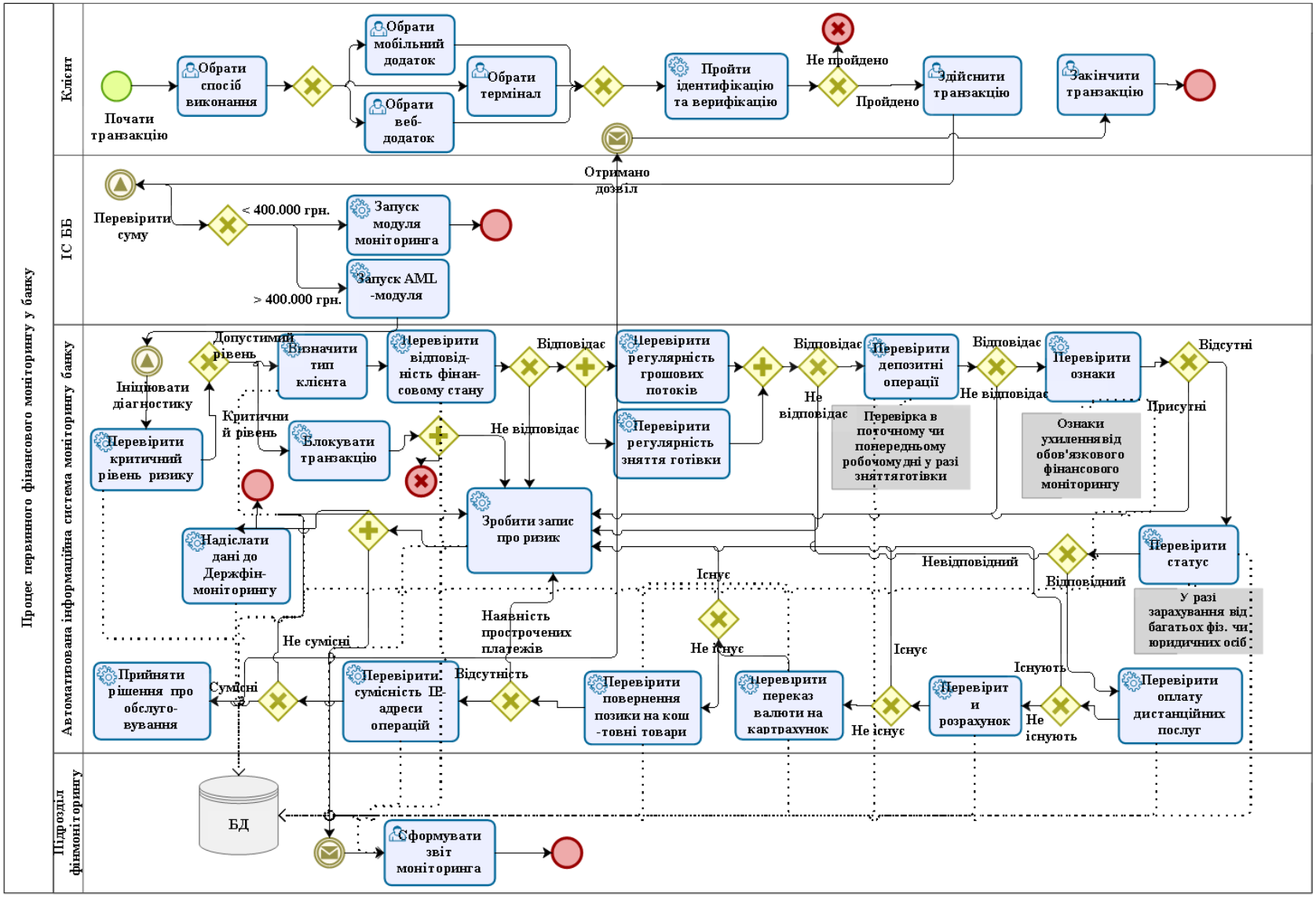


Рисунок 1.23– Бізнес-модель процесу автоматизованого фінансового моніторингу банку [49]

Використання автоматизованої системи полегшить перевірку всього обсягу транзакцій, незалежно від їх суми. По-четверте, перевагою запропонованого рішення є гнучкість налагодження системи у разі зміни законодавства, положень НБУ, інструкцій банків щодо перевірки таких операцій.

При здійсненні симуляцій враховуються два важливих твердження:

– враховуємо, що час на виконання операцій автоматизованою системою та фахівцем є однаковим, що відповідає принципу співставності витрат, якого потрібно дотримуватися у разі визначення ефективності та порівняння витрат;

– симуляції результатів здійснюємо, виходячи з автоматизованої та ручної обробки даних, оскільки запропоновані бізнес-процеси мають вже елементи оптимізації, тобто процеси, реалізовані на практиці є вже застарілими та прогнозуються удосконалюватися, виходячи із дотримання норм законодавства.

Результати проведеної симуляції по даному процесу представлені на рисунку А.2 у додатку А. Умовами симуляції були наступні: кількість операцій – 1000; ймовірність відхилення операції у випадку не проходження перевірки – 1% по кожному вузлу (бажаний показник); час на виконання 1 запиту в автоматизованій системі фінансового моніторингу – 1 с.; час було задано тільки для операцій перевірки, щоб виявити тільки той його обсяг, який буде витрачено на моніторинг. В результаті отримано, що середній час на перевірку 1 транзакції на предмет наявності ознак фінансового моніторингу 42,35 с., тобто на 1000 операцій буде витрачено 11,77 год. Оскільки симуляція не враховує потужність серверів, то даний показник в дійсності може бути завищеним. На практиці подібна перевірка досвідченим фахівцем займає 20 хвилин. Тобто людині буде потрібно витратити на перевірку 1000 операцій 333,33 годин: ((20 хв. * 1000 оп.) / 60 хв.). Тільки по показнику часу ефективність впровадженого запропонованого процесу буде наступною: автоматизована система у 28,33 рази швидше здійснюватиме перевірку операцій у розрахунку на 1000 транзакцій.

Кількість операцій після перевірки – 877, розрахованих за формулою (2.1), коефіцієнт результативності – 0,877 (за формулою (2.2)). Тобто за умови 1% операцій, які носять ознаки відмивання кримінальних доходів, по кожному з

критеріїв перевірки, система буде позитивно ідентифікувати 87,7% операцій, що є високим результатом.

Проведемо симуляцію по ресурсам. Для цього задаємо фахівця, який здійснює моніторинг, та автоматизовану інформаційну систему фінансового моніторингу (AML-модуль). Визначимо їх вартісні оцінки, а саме вартість людино-години та машино-години. Для розрахунків використаємо дані, які відображають фактичні витрати азіатських банків, понесені на AML-систему (AML – Anti-Money Laundering – протидія відмиванню коштів), які за принципами роботи у даному напрямку схожі з українськими. Інформація міститься у звіті компанії LexisNexis та охоплює період 09.2015 – 01.2016 [**Ошибка! Источник ссылки не найден.**]. Розрахунки наведені у таблиці 1.7:

Таблиця 1.7 – Розрахунки вартості людино-години та машино-години [49]

Назва показника	Фактичне значення, узяте із звіту [**Ошибка! Источник ссылки не найден.**]	Розраховане значення
Кількість опитаних компаній	210	X
Кількість опитаних банків	50%	105
Витрати на AML по всім банкам, дол. США	1500000000	X
Середні витрати на 1 банк, дол. США	X	14285714,29
Витрати на програмне та технічне забезпечення (зовнішні та внутрішні), дол. США	19%	2714285,71
Витрати на персонал, задіяний в AML, дол. США	81%	11571428,57
Час функціонування AML-системи за рік за умови 24-годинної роботи, год.	X	8760
Вартість машино-години, дол. США	X	309,85
Вартість людино-години, дол. США	X	1320,94

Наведені у таблиці 1.7 розрахунки показують вартість машино-години, якщо задіяно увесь комплекс програмно-технічних засобів, та вартість людино-години, якщо задіяно увесь штат працівників. Оскільки значення вартісних

показників є комерційною таємницею для банків, то можна скористатися тільки умовним визначенням витрат. Але й ці розрахунки можуть дати уявлення про ефективність. Використовуючи отримані значення вартості машино-години та людино-години, проведемо симуляцію «Аналіз ресурсів», результат якої представлений на рисунку 1.24.

Результати, представлені на рисунку 1.24, показують, що витрати на 1000 транзакцій, перевірених фахівцями фінансового моніторингу, у 4,26 разів вище, ніж витрати на 1000 транзакцій, перевірених AML-модулем.

Scenario information

Название	Scenario 1
Time unit	Seconds
Продолжительность	030,00:00:00

Resource	Utilization	Total fixed cost	Total unit cost	Total cost
Фахівець з фінансового моніторингу	100,00 %	0	15 223,83	15 223,83
AML-модуль	0,00 %	0	0	0
	Total	0	15 223,83	15 223,83

Resource	Utilization	Total fixed cost	Total unit cost	Total cost
Фахівець з фінансового моніторингу	0,00 %	0	0	0
AML-модуль	100,00 %	0	3 571,02	3 571,02
	Total	0	3 571,02	3 571,02

Рисунок 1.24 – Результати симуляції за ресурсами для бізнес-процесу автоматизованого фінансового моніторингу банку [49]

Можна зробити висновок, що при реалізації запропонованого бізнес-процесу фінансового моніторингу, його ефективність буде вищою для автоматизованого варіанту, ніж для ручного. Для остаточних розрахунків важливо мати інформацію щодо витрат на придбання та впровадження такої системи, а також мати інформацію щодо її результативності.

Перед розробкою моделі бізнес-процесу для перевірки транзакцій на

ознаки кібершахрайства, необхідно створити інформаційну модель виявлення шахрайств в операціях, ініційованих зовнішнім середовищем. Ця модель відображає функціонування інформаційних потоків у автоматизованому середовищі і побудована з використанням нотації DFD (data flow diagrams), що є одним з інструментів структурного моделювання та проектування інформаційних систем. Для цього використовується програмне забезпечення "All Fusion Process Modeller" [51].

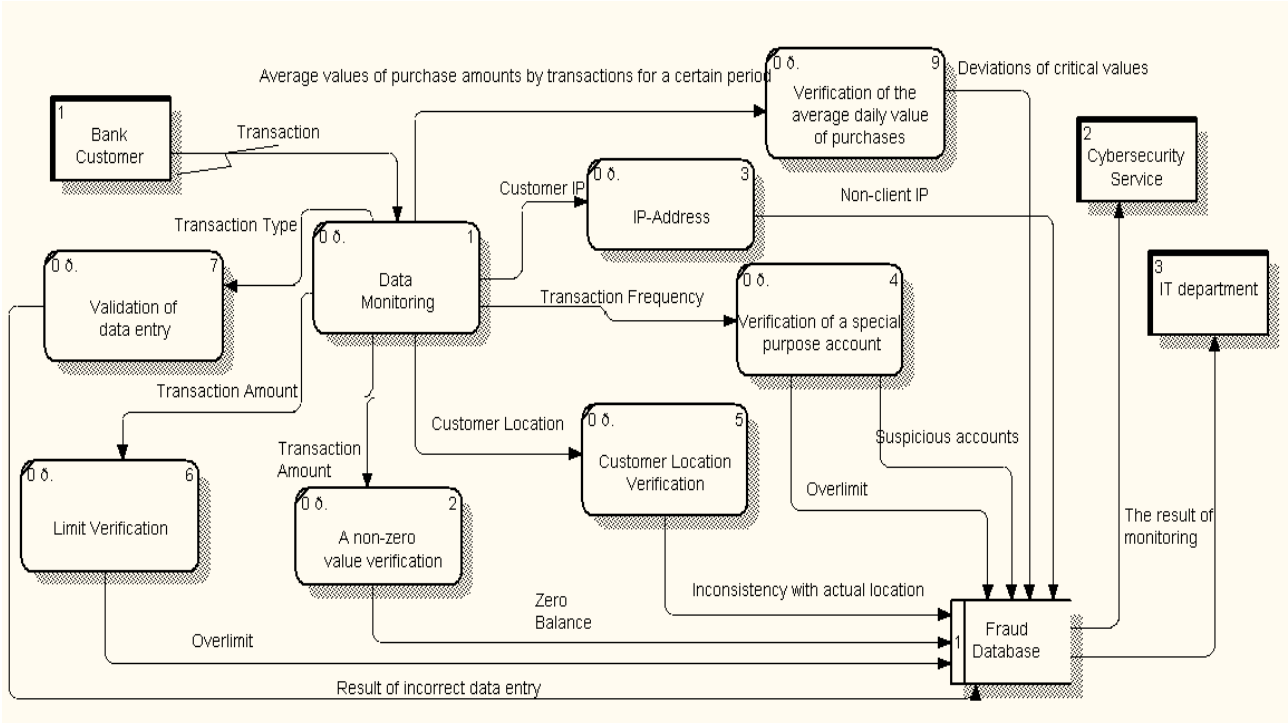


Рисунок 1.25– Інформаційна модель виявлення властивостей кібершахрайств клієнтів [49, 51, 52]

Модель 1.25 показує рух інформації, що здійснюється на основі моніторингу транзакцій клієнта із залученням відповідних функцій перевірки даних. Система перевіряє:

– суми операції на предмет вичерпання коштів з рахунку. Як правило, кіберзлочинець знімає усі кошти з рахунку, що не може бути характерним для власника, який залишає певний ліміт на рахунку. Система подає інформацію про нульовий баланс;

– суми операцій на предмет перевищення лімітів, відповідних характеру та типу рахунку. Оскільки кіберзлочинець не знає суми ліміту, то в процесі злочину

він може перевищити необхідну суму, в результаті чого система буде повідомляти про спробу зняття недозволеної суми коштів;

– геолокація клієнта, оскільки транзакція може ініціюватися з будь-якої точки, яка може не відповідати фактичному місцезнаходженню клієнта;

– мету рахунку. Наприклад, його може бути занесено до “чорного списку” клієнтів банку, або його цільовим призначенням виступатиме інший вид діяльності, не пов'язаний з метою транзакції, або його було відкрито в іншому банку;

– IP-адресу клієнта. Перевіряється ситуація, якщо транзакція здійснюється з адреси, яка не зареєстрована в базі даних клієнта;

– коректності введеної інформації, яка залежить від типу банківської операції. Якщо система фіксує спроби неправильного введення даних, то це може свідчити про потенційне зламування клієнтського акаунту;

– перевищення середньоденної суми покупок. На вході аналізуються середньоденні значення витрачених коштів та у випадку критичного їх перевищення система може сигналізувати про можливості шахрайства.

Результати перевірки щодо потенційного кібершахрайства автоматично надсилаються до бази даних кіберзлочинів і потім обробляються вручну. Висновки перевірки передаються департаментам банку – інформаційних технологій та кібербезпеки.

Базуючись на поданій вище інформаційній моделі (рисунок 1.25) розроблено бізнес-модель процесу перевірки транзакцій на наявність ознак кібершахрайств, виконану у нотації BPMN 2.0 (рисунок 1.26).

Рисунок 1.26 показує, що процес починається з того, що потенційний кіберзлочинець активізує роботу із системою або на основі веб-сайту банку, або мобільного додатку, або терміналу. Після вдалої ідентифікації та верифікації система моніторингу буде перевіряти транзакцію в залежності від її суми або на властивості легалізації незаконних коштів, або на кіберзлочин. Процес перевірки на можливість здійснення кібершахрайства зображено на рисунку 1.26.

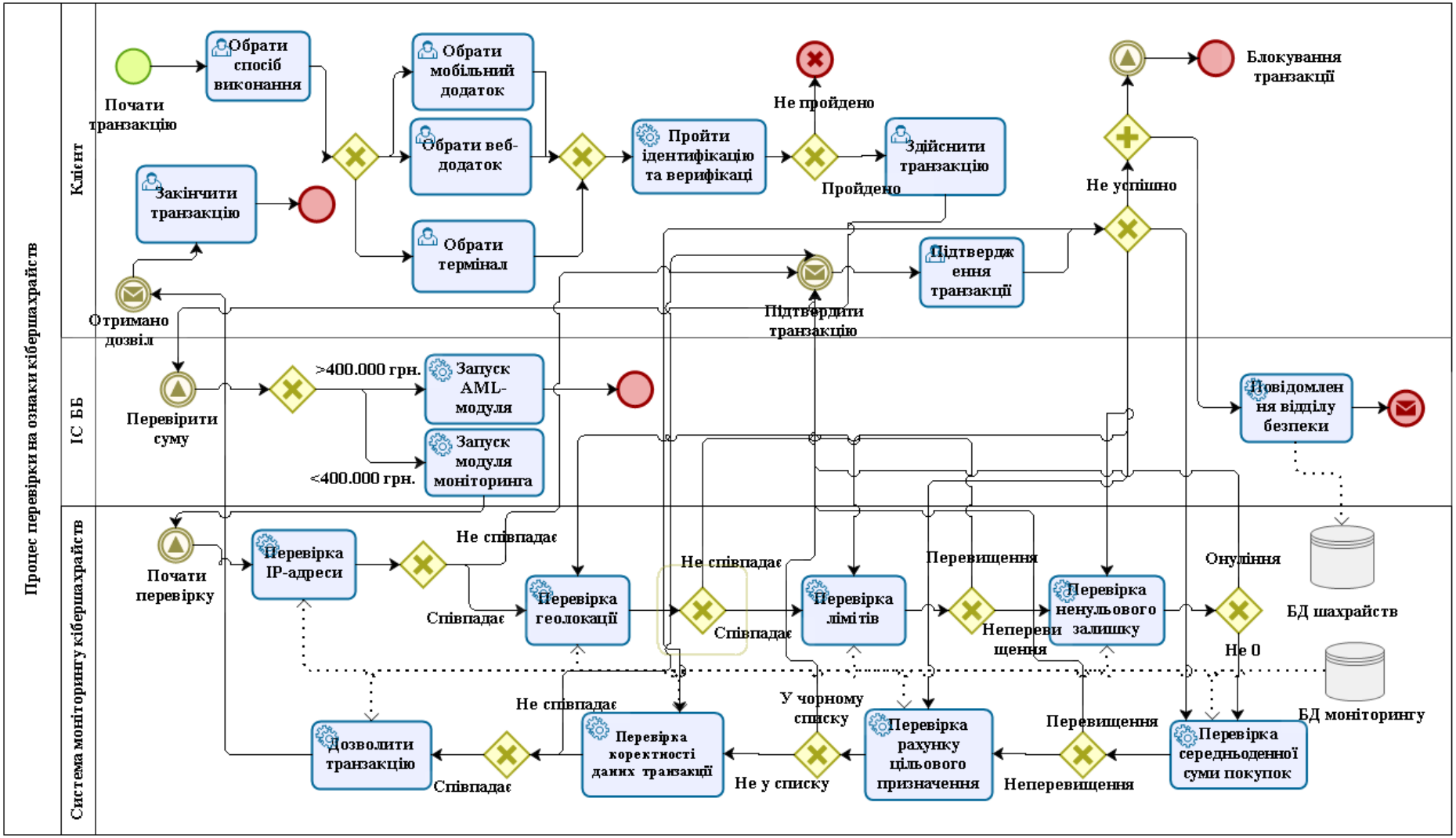


Рисунок 1.26– Бізнес-модель процесу перевірки транзакцій на наявність ознак кібершахрайств [49]

У випадку не виявлення властивостей ймовірного кіберзлочину модуль моніторингу дає дозвіл здійснити транзакцію, яка може бути завершена. У результаті ідентифікації кібершахрайства система надсилає запит клієнту щодо підтвердження ним шляхом дзвінка, текстового повідомлення, повідомлення в додатку. У цьому випадку клієнт повинен виконати додаткову аутентифікацію. Якщо транзакція була ним ініційована, то система дає дозвіл на її завершення. Якщо клієнт є кіберзлочинцем, то ця процедура є для нього недоступною і система його заблокує. Результат автоматично буде згенерованим для служби кібербезпеки банку.

По даному процесу було проведено симуляції по витратам часу та вартісним витратам ресурсів (рисунок А.3 у додатку А). Умови симуляції: кількість операцій – 1000; ймовірність відхилення операції у випадку не проходження перевірки – 1% по кожному вузлу (бажаний показник); для вузла, який відповідає додатковій автентифікації після того, як система виявила потенційну загрозу, ймовірність була розподілена пропорційно; час на виконання 1 запиту в автоматизованій системі – 1 с. Виявлено, що середній час на перевірку 1 транзакції на предмет наявності ознак кібершахрайств дорівнює 9,86 с., тобто на 1000 операцій буде витрачено 2,71 год.

Виявилося, що 7 операцій не пройшли перевірок та повторної ідентифікації. Оскільки тільки 976 операцій із 1000 підлягали перевірці на ознаки кібершахрайств, то показник результативності склав 99,28%. Це значення може свідчити про високу ефективність системи. На практиці такий результат можливо досягти за рахунок ефективного налаштування параметрів моніторингу, що потребує постійної перевірки з боку відділу внутрішнього аудиту банку.

Проведемо симуляцію процесу по ресурсах. Для цього визначимо собівартість людино-години та машино-години. У звіті компанії Deloitte зазначається, що у 2020 році банки здійснювали витрати на інформаційну безпеку в розмірі від 0,6% всіх витрат, що склало приблизно 9,4% від IT-бюджету або $2688 на 1 людину на рік [**Ошибка! Источник ссылки не найден.**].

Виходячи із того, що у 2020 році було 251 робочий день, та беручи до уваги 8-годинний робочий день, визначаємо, що вартість 1 машино-години буде дорівнювати $1,34: $2688 / (251 днів * 8 годин). Для порівняння даного процесу із ручною обробкою визначаємо, що заробітна плата банківського аналітика в Україні дорівнює 17500 грн. на місяць [**Ошибка! Источник ссылки не найден.**]. Виходячи із того, що у 2020 році було 251 робочий день, та беручи до уваги 8-годинний робочий день, визначаємо, що вартість 1 машино-людини буде дорівнювати $1,34: $2688 / (251 днів * 8 годин).

Результати проведеної симуляції по ресурсах представлено на рисунку 1.27.

Scenario information

Название	Scenario
Time unit	Seconds
Продолжительность	030,00:00:00

Resource	Utilization	Total fixed cost	Total unit cost	Total cost
Аналітик	99.99 %	0	9.12	9.12
Система моніторингу	0.00 %	0	0	0
	Total	0	9.12	9.12

Scenario information

Название	Scenario
Time unit	Seconds
Продолжительность	030,00:00:00

Resource	Utilization	Total fixed cost	Total unit cost	Total cost
Аналітик	0.00 %	0	0	0
Система моніторингу	99.99 %	0	3.27	3.27
	Total	0	3.27	3.27

Рисунок 1.27 – Результати симуляції за ресурсами для бізнес-процесу перевірки транзакцій на наявність ознак кібершахрайств

На рисунку 1.27. можна побачити, що у разі забезпечення практично 100% виконання транзакцій автоматизованою системою та аналітиком, витрати

ресурсів для першого варіанту є меншими у 2,79 разів. Тобто економічно доцільним є здійснення перевірки із використанням автоматизованого модулю (3,27 дол. витрат на 1000 операцій) у порівнянні із здійсненням перевірки фахівцем (9,12 дол. витрат на 1000 операцій).

Інформаційна модель виявлення кіберзлочинів, які здійснюються інсайдерами банку, представлена на рисунку 1.28.

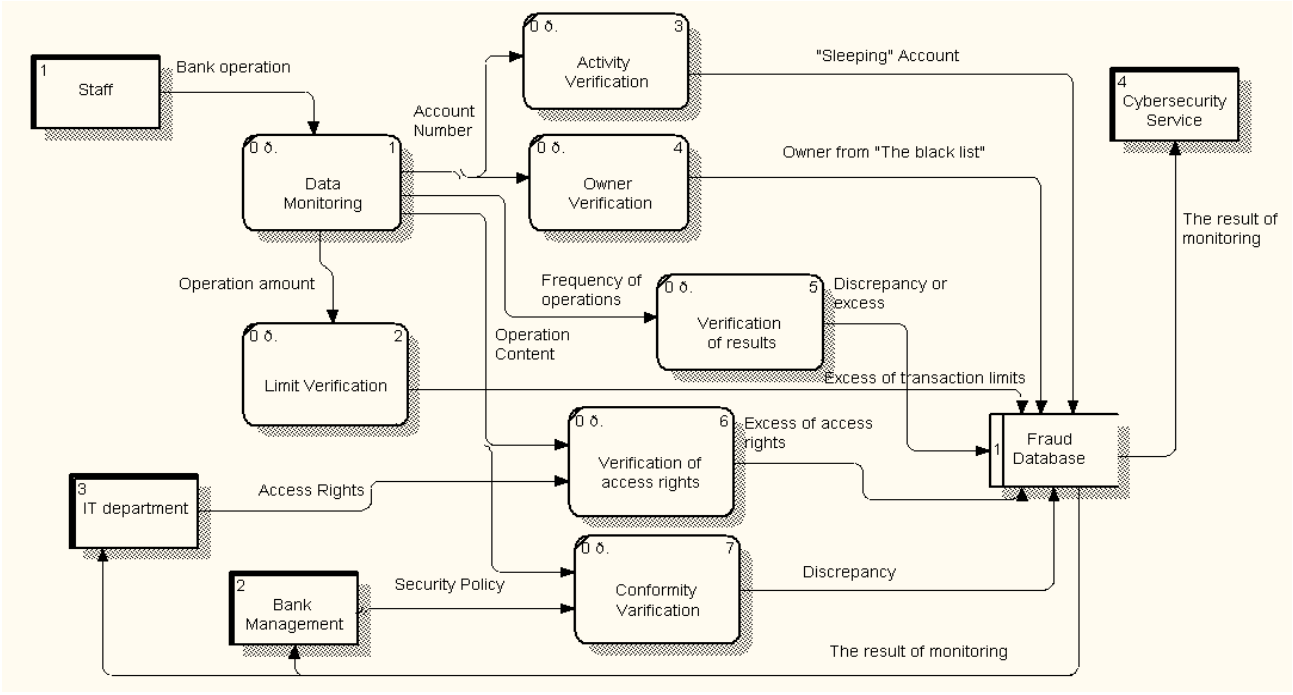


Рисунок 1.28– Інформаційна модель виявлення властивостей кіберзлочинів інсайдерів [49, 51, 52]

Модель 1.28 ідентифікує інформаційні потоки в процесі моніторингу банківських транзакцій, ініціаторами яких є працівники банку. Підлягають моніторингу:

– дії на рахунку. Це відбувається тоді, коли працівники банку використовують “сплячі рахунки” для задоволення власних потреб;

– власники рахунку. Персонал банку може використовувати рахунки, які належать померлим, іноземцям, особам із “чорного списку”;

– ліміти транзакцій. Інсайдери можуть відслідковувати рахунки, по яким встановлено ліміти із порушенням вимог НБУ та політикою банку;

– дії персоналу на предмет дотримання ними банківських нормативів та

посадових інструкцій;

– дії персоналу на відповідність правам доступу, встановлених у відповідності до посадових інструкцій;

– дії працівників на відповідність політиці безпеці банку.

Результати накопичуються у базі даних кіберзлочинів, обробляються та надсилаються відділу кібербезпеки банку, IT-відділу та менеджменту банку.

У відповідністю із запропонованою інформаційною моделлю (рисунок 1.28) розроблено схему процесу здійснення операції персоналом з урахуванням її перевірки на ознаки шахрайства у нотації BPMN 2.0 (рисунок 1.29).

Процес виглядатиме наступним чином:

– банківський співробітник, який може бути потенційним шахраєм, авторизується в банківській системі та здійснює банківську операцію;

– система моніторингу кібершахрайств перевіряє операцію на предмет кіберзлочину із використанням зазначених критеріїв перевірки, а саме: прав доступу, операцій на відповідність політики безпеки, особи працівника, дотримання банківських нормативів, сплячих рахунків, активностей рахунків та лімітів по операціях ;

– у випадку, коли транзакція відповідає всім параметрам та не містить ознаки кіберзлочину з боку інсайдерів, то система дозволяє її здійснення та працівник може її завершити;

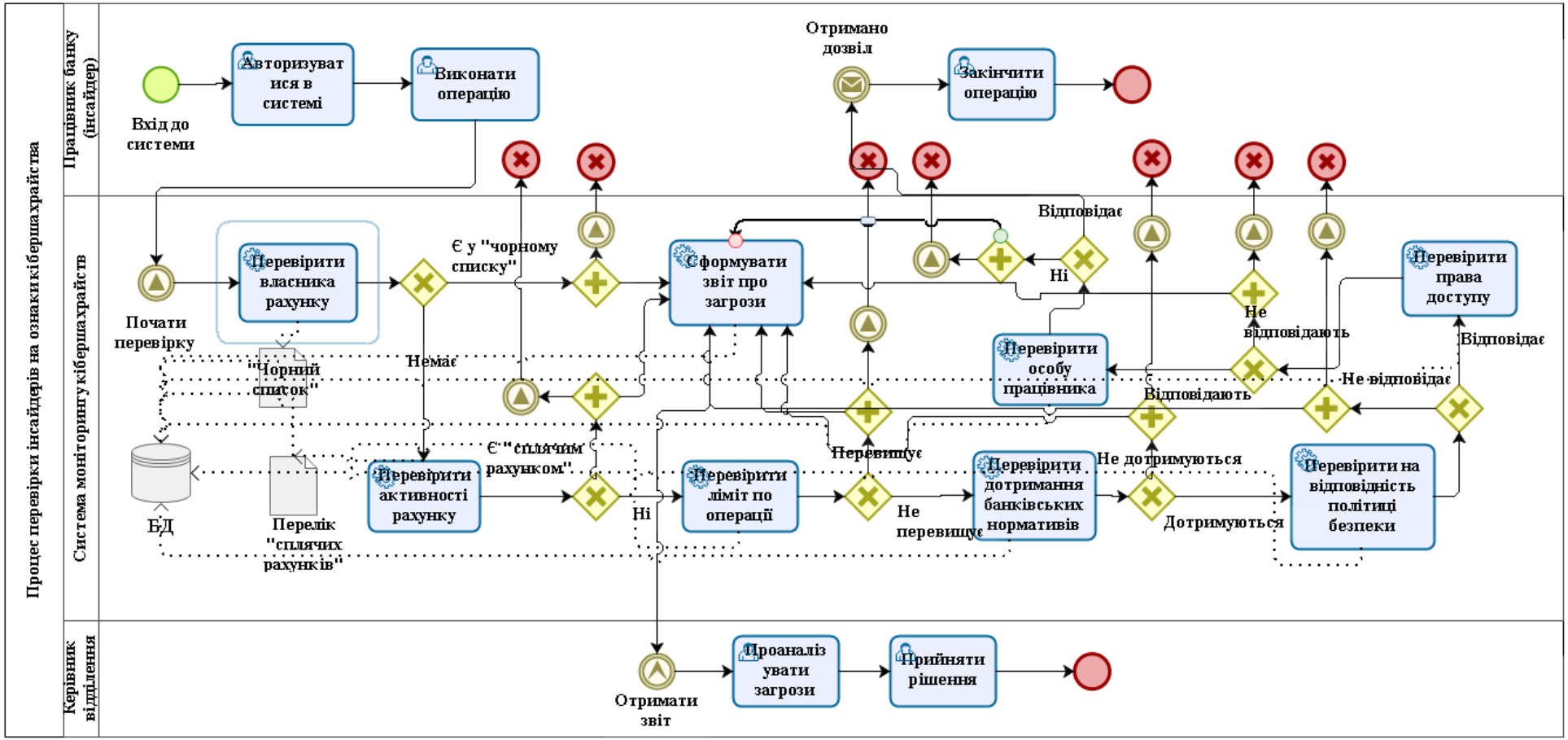


Рисунок 1.29– Бізнес-модель процесу перевірки дій інсайдерів на ознаки кібершахрайств [49]

– якщо система знаходить властивості кіберзлочину, то вона автоматично генерує повідомлення менеджеру відповідного відділу банку, де було ініційовано транзакцію, який проводить аналіз даних та приймає рішення щодо ймовірного кібершахрайства.

По даному процесу було проведено симуляції по витратам часу та ресурсів (рисунок А.4 у додатку А). За умовами: кількість операцій – 1000; ймовірність відхилення операції у випадку не проходження перевірки – 1% по кожному вузлу (бажаний показник); час на виконання 1 запиту в автоматизованій системі – 1 с. Виявлено, що середній час на перевірку 1 транзакції на предмет наявності ознак кібершахрайств з боку інсайдерів дорівнює 6,86 с., тобто на 1000 операцій буде витрачено 1,90 год. Було виявлено 65 операцій з ознаками шахрайств, відповідно показник результативності системи складає 93,5%. Результати симуляції по ресурсам (рисунок 1.30) показують, що ефективність автоматизованого виявлення ознак кіберзагроз є менш витратним в 2,79 разів.

Scenario information

Название	Scenario 1
Time unit	Seconds
Продолжительность	030,00:00:00

Resource	Utilization	Total fixed cost	Total unit cost	Total cost
Аналітик	100.00 %	0	7.12	7.12
Система моніторингу	0.00 %	0	0	0
	Total	0	7.12	7.12

Scenario information

Название	Scenario 1
Time unit	Seconds
Продолжительность	030,00:00:00

Resource	Utilization	Total fixed cost	Total unit cost	Total cost
Аналітик	0.00 %	0	0	0
Система моніторингу	100.00 %	0	2.55	2.55
	Total	0	2.55	2.55

Рисунок 1.30 – Результати симуляції за ресурсами для бізнес-процесу перевірки дій інсайдерів на ознаки кібершахрайств

Запропонована методика оптимізації бізнес-процесів представляє собою організаційний рівень ключових алгоритмів інтеграції систем фінансового моніторингу і кібербезпеки. Її реалізація дозволить формувати передумови виявлення транзакцій, наслідком яких може бути здійснення шахрайства з боку зовнішнього злочинця чи інсайдера, а також відмивання кримінальних доходів. Впровадження в практичну діяльність розроблених моделей дозволить охопити широке коло операцій незалежно від їх належності до зовнішнього чи внутрішнього середовища. Запропоновані алгоритми дозволять не тільки виявити слабкі місця в захисті інформації, але також вони слугують передумовою конвергенції систем кібербезпеки та фінансового моніторингу в рамках єдиної інтегрованої банківської автоматизованої системи. Це сприятиме здійсненню системного моніторингу для перевірки операцій клієнтів банку на предмет наявності властивостей кібер- і фінансових злочинів. Врешті-решт впровадження запропонованого підходу до оптимізації бізнес-процесів сприятиме підвищенню ефективності системи управління банком в цілому за рахунок прийняття більш дієвих рішень.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**, **Ошибка! Источник ссылки не найден.**].

1.2.2 Математичне забезпечення алгоритмів виявлення кібершахрайських операцій з кредитними картками

Економічна криза, низький рівень доходів населення, зростання кількості комерційних банків, спеціалізованих фінансових компаній та стрімкий розвиток інформаційних технологій сформували умови для появи банківських шахрайств. Їх мета полягає у незаконному привласненні коштів однією особою або групою осіб. Як правило, процес скоєння банківських шахрайств є прихованим. Масовість їх здійснення може вразити фінансову безпеку банку, призвести до

фінансових збитків та, як наслідок, викликати втрату довіри та репутації серед клієнтів, стати однією з причин банкрутства банківської установи.

Найбільш поширеним видом банківського шахрайства є ті, що пов’язані із кредитними операціями, тобто процесом кредитування клієнтів, а також кредитними картками клієнтів. Це відбувається завдяки спрощення процедури надання кредитів, а також створення гнучких умов для клієнтів щодо використання ними кредитних коштів та засобів платежу. Також судово-бухгалтерською експертизою фінансово-кредитних установ дедалі частіше фіксуються махінації, що стосуються незаконних кредитних операцій, до яких вдаються не тільки позичальники (юридичні та фізичні особи), але й кредитори (банки, фонди, асоціації). Останнім часом відсоток таких шахрайств зростає у порівнянні із іншими видами. Так, у 2020 році шахрайство з кредитними картками зайняло друге місце серед п’ятірки найбільш розповсюджених фінансових злочинів та становило 29,7% (шахрайства із державними пільгами, на які подано заявку, або отримано – 32,0%; різні крадіжки особистих даних – 22,9%; шахрайства із позиками для бізнесу/особистісного користування – 8,1%; податкове шахрайство – 7,3%) [**Ошибка! Источник ссылки не найден.**].

За часту шахрая виявляють вже після того, як було скоєно злочин, тому існує потреба саме у передбаченні потенційних шахрайств. Це можливо тільки в процесі оцінювання ймовірності їх виникнення в ході кредитування клієнтів банку. У контексті даної проблеми є потреба у створенні комплексу заходів для попередження кредитних шахрайств. З цією метою доцільно застосовувати математичні методи, за допомогою яких, можна створювати математичні моделі для проведення ідентифікації банківських транзакцій на предмет шахрайства, або ідентифікації потенційного клієнта банку, який може його скоїти. Використання новітніх інформаційних технологій та мов програмування дозволяє будувати моделі будь-якого рівня складності та спрощувати їх розрахунки.

Проблема виявлення та попередження шахрайств у банківській сфері є досить актуальним напрямом дослідження. Для виявлення його тенденцій

проведемо бібліометричний аналіз наукових досліджень, результати яких опубліковані в міжнародних виданнях, що були проіндексовані у базі даних Scopus. З цією метою було зроблено вибірку таких джерел та шляхом застосування аналітичного програмного забезпечення VOSviewer побудовано карту наукометричної бібліографії досліджень, присвячених проблемі шахрайств щодо кредитування клієнтів банків (див. рис. 1.31).

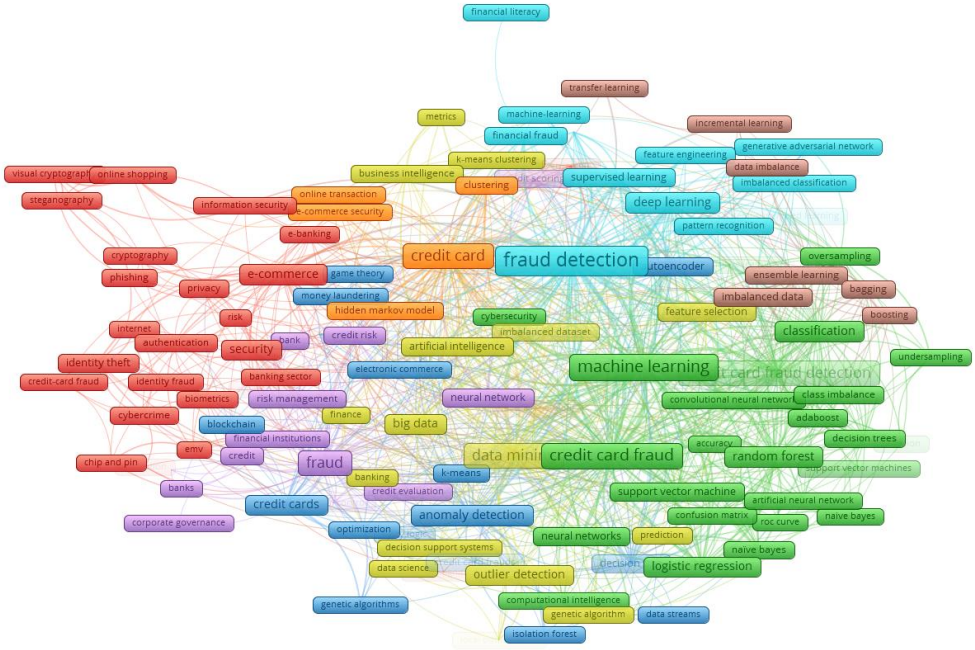


Рисунок 1.31 – Карта наукометричної бібліографії досліджень, присвячених проблемі шахрайств щодо кредитування клієнтів банків

Джерело дослідження: побудовано авторами на основі бази даних Scopus

Червоний кластер досліджень (див. рис. 1.31) охоплює публікації, які стосуються процесів захисту транзакцій електронного банкінгу, інтернет-банкінгу, електронної комерції, онлайн-магазинів, що відбуваються за допомогою засобів інформаційної безпеки – аутентифікації, ідентифікації, криптографії, стеганографії, біометрії, тощо. Тут можна виділити роботи Чжен Л., Лю Г., Ян Ч., Цзян Ч. [**Ошибка! Источник ссылки не найден.**], Пріша П., Нео Х.-Ф., Онг Т.-С., Тео Ч.-С. [**Ошибка! Источник ссылки не найден.**] та інші. Публікації зеленого кластеру (див. рис. 1.31) розкривають напрямки застосування машинного навчання для вирішення проблем кредитних шахрайств

у банках. Так, застосовують такі інструменти, як кластерний аналіз, випадковий ліс, логістична регресія, нейронні мережі, методи опорних векторів, тощо. В даному напрямку працювали Діліп М.Р., Наванет А.В., Абхішек М. [**Ошибка! Источник ссылки не найден.**], Цуй Ю., Сон З., Ху Дж. [**Ошибка! Источник ссылки не найден.**] та інших. Дослідження коричневого кластеру (див. рис. 1.31) перетинаються із попереднім кластером, оскільки вивчаються питання ансамблевого та інкрементного навчання, бегінг та бустінг, що застосовується до незбалансованих даних. Ці аспекти досліджували Ван Р., Лю Г. [**Ошибка! Источник ссылки не найден.**], Собанадеві В., Раві Г. [**Ошибка! Источник ссылки не найден.**] та інші. Роботи помаранчевого кластеру (див. рис. 1.31) стосуються проблематики операцій з кредитними картками, в яких вирішуються питання за допомогою байесівського підходу та моделі Марковіца. Тут можна виділити публікації Чжоу Ю., Сон Х., Чжоу М. [**Ошибка! Источник ссылки не найден.**], Мішра С.П., Кумарі П. [**Ошибка! Источник ссылки не найден.**], тощо. Дослідження синього кластеру стосуються процесів виявлення аномалій у операціях, пов’язаних із відмиванням незаконних коштів у банках та кредитними картками, для чого дана проблема вирішується за допомогою нечіткої логіки, теорії ігор, оптимізації, великих даних та блокчейнів. Ця проблема була розкрита такими науковцями, як Рачавеліас М.Г. [**Ошибка! Источник ссылки не найден.**], Нана З., Сюцзянь В., Чжунцю З. [**Ошибка! Источник ссылки не найден.**] та інші. Публікації блакитного кластеру (див. рис. 1.31) відображають напрям процесу знаходження шахрайств у фінансовій сфері, для чого застосовуються розпізнавання патернів, методи глибокого навчання, навчання з вчителем та без вчителя, feature engineering, тощо. В цій сфері працювали Зоу Х. [**Ошибка! Источник ссылки не найден.**], Мектерович І., Каран М., Пінтар Д., Бркіч Л. [**Ошибка! Источник ссылки не найден.**], тощо. Напрям бузкового кольору (див. рис. 1.31) розкриває дослідження процесу виявлення шахрайств у банках та інших фінансових інститутах за допомогою ризиків – кредитного та операційного. Тут можна виділити роботи Джанотті Е., Даміан да Сілва Е. [**Ошибка! Источник ссылки не найден.**], Цзоу В., Страуб Д., Венс А., Ян Дж.

[**Ошибка! Источник ссылки не найден.**] та інших. Ключовою проблемою жовтого кластеру (див. рис. 1.31) є Data Mining та питання, які з ним пов’язані, а саме штучний інтелект, генетичні алгоритми, великі дані, системи прийняття рішень, бізнес-аналітика. Цим напрямом займалися Цзін Р., Тянь Х., Чжоу Г., Чжан Х., Чжен Х., Цзен Д.Д. [**Ошибка! Источник ссылки не найден.**], Уедраго А.-Ф., Геученн Ц., Нгуєн Ж.-Т., Тран Г. [**Ошибка! Источник ссылки не найден.**], тощо.

Не дивлячись на велику кількість досліджень щодо вирішення проблеми шахрайств, публікації в окреслених напрямах розкривають їх різні аспекти. Стосовно питання оцінювання ймовірності виникнення шахрайства щодо кредитування клієнтів банків, то воно потребує досліджень, особливо для реалій вітчизняних економічних відносин.

Метою дослідження є побудова математичних моделей для оцінювання ймовірності виникнення шахрайства щодо кредитування клієнтів банків, їх реалізація і візуалізація за допомогою мови програмування Python.

Для проведення дослідження окресленої проблеми було узято базу даних з відкритих джерел даних, яка відображає операції кредитування клієнтів та містить основні їх характеристики. Так, набір даних сформували 122 змінні та 307511 спостережень. До незалежних змінних ввійшли такі показники, як: вік, стать, наявність нерухомості, тип нерухомості, наявність рухомого майна, сімейний статус, кількість дітей, тип зайнятості, рівень освіти та ін. Цільовою виступає бінарна змінна, яка є індикатором ймовірного шахрайства в процесі кредитування, а саме: 0 – виявлено клієнта – ймовірного шахрая; 1 – не виявлено клієнта – ймовірного шахрая.

З метою подальшої побудови математичних моделей розроблено концептуальну модель оцінювання ймовірності виявлення ознак шахрайства в процесі кредитування клієнтів банку, яка відображає основні етапи роботи з масивом вхідних даних та представлена на рисунку 1.32.

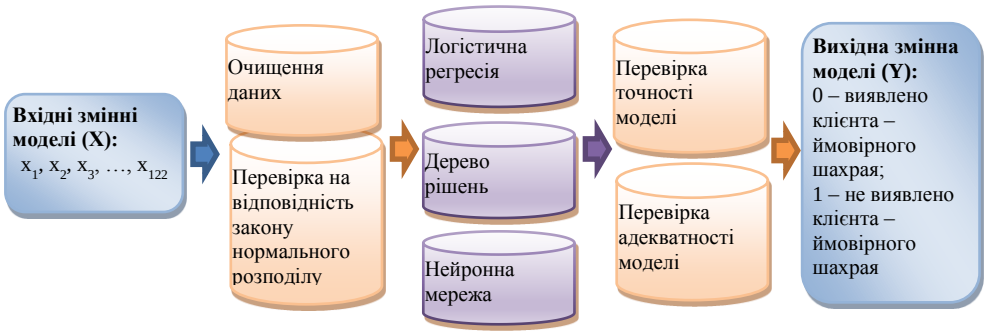


Рисунок 1.32 – Концептуальна модель оцінювання ймовірності виявлення ознак шахрайства в процесі кредитування клієнтів банку

Виходячи з інформації концептуальної моделі (рис. 1.32), процес моделювання передбачає:

– попередню обробку набору даних, а саме їх очищення від пропусків та перевірка на відповідність закону нормального розподілу;

– побудову математичних моделей для оцінки ймовірності виникнення шахрайства у процесі кредитування клієнтів банку: логістичну регресію; дерево рішень; нейронну мережу;

– верифікацію побудованих моделей, тобто перевірка їх точності та адекватності.

Оскільки дослідження стосується оцінки ймовірності виникнення шахрайства, то для вирішення даної проблеми найбільш ефективними є методи інтелектуального аналізу даних Їх переваги та недоліки представлені в таблиці 1.8.

Таблиця 1.8 – Переваги та недоліки математичних моделей

Назва моделі	Переваги	Недоліки
Логістична регресія	має один з найпростіших алгоритмів; є легкою у виконанні та інтерпретації; є простою у оновленні нових даних; є ефективнішою за лінійну регресію	алгоритм чутливий до викидів; необхідна мінімальне значення або відсутність мультиколінеарності між незалежними змінами
Дерево рішень	вимагає менше зусиль та часу на підготовку даних; є дуже легким у поясненні; не вимагає нормалізації даних;	обчислення можуть бути набагато складнішими за інші алгоритми; передбачає багато часу для навчання моделі;

	дозволяє мати пропущені дані	є недостатнім для прогнозування безперервних значень
Нейронна мережа	досить стійка до шуму в навчальних даних; помилки в навчальному наборі не впливають на результат; використовується для швидкої оцінки функції	вимагає паралельної обробки даних; складнощі з відображенням; результативні значення не є оптимальними

Логістична регресія – це статистичний регресійний метод, що застосовується у випадку, коли залежна зміна являється бінарною, тобто може набувати значення 0 або 1. Логістична регресія є прогностичним аналізом та використовується для опису даних та пояснення взаємозв'язку між одним залежним фактором (змінною) та однією або декількома незалежними. Її математичний вираз можна представити формулою (1.16):

$$\mathrm{P}(\hat{y} = 1) = \frac{1}{1 + exp^{-\hat{y}}} = \frac{1}{1 + exp^{-(\beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n)}}, \quad (1.16)$$

де $x_1, x_2, \dots, x_n$ – множина незалежних змінних-факторів, які впливають на результатний показник;

$\beta_0, \beta_1, \dots, \beta_n$ – множина параметрів регресії, які необхідно оцінити в процесі побудови моделі логістичної регресії;

$\hat{y}$ – залежна змінна-фактор, значення якої прогнозується в процесі моделювання;

$\mathrm{P}(\hat{y} = 1)$ – ймовірність виникнення випадку, при якому значення результативної змінної дорівнює 1.

Для побудови логістичної регресії використовувалася мова програмування Python. Для її реалізації дані були очищені від пропущених значень та перевірені на відповідність закону нормального розподілу. Далі набір вхідних даних було розділено на дві вибірки – тестову та тренувальну. Після здійснення даної

процедури було оцінено точність опису залежної змінної незалежними, результат чого представлений на рисунку 1.33.

```
print("Training set score: %f" % LogReg.score(X_train, Y_train))
print("Test set score: %f" % LogReg.score(X_train, Y_train))
```

```
Training set score: 0.930932
Test set score: 0.936000
```

Рисунок 1.33 – Точність опису залежної змінної

Результати показують, що частка правильних прогнозів у тренувальній вибірці становить 93,09%, а у тестовій – 93,60%, що свідчить про адекватність даних обох вибірок та високу точність прогнозування.

Після отримання придатного для моделювання набору, було побудовано модель логістичної регресії. Оскільки для отримання адекватної моделі необхідно, щоб її параметри були статистично значущими, то було проведено їх оцінку із використанням значення p-value та довірчих інтервалів. Статистично незначущі фактори було усунуто з моделі. Процедура повторювалася доти, доки було отримано модель із усіма статистично значущими параметрами. Так, було проведено 9 ітерацій. Результати логістичної регресії представлені на рисунку 1.34.

```
Optimization terminated successfully.
         Current function value: 0.627703
         Iterations 9
                           Logit Regression Results
==============================================================================
Dep. Variable:                 TARGET   No. Observations:                40416
Model:                          Logit   Df Residuals:                    40409
Method:                           MLE   Df Model:                            6
Date:                Fri, 11 Jun 2021   Pseudo R-squ.:                 -1.539
Time:                        11:24:35   Log-Likelihood:                -25369.
converged:                       True   LL-Null:                       -9992.8
Covariance Type:            nonrobust   LLR p-value:                    1.000
================================================================================================
                                     coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------------------------
NAME_TYPE_SUITE_Other_A             -4.1769      1.008     -4.145      0.000      -6.152      -2.202
NAME_FAMILY_STATUS_Widow            -3.4096      0.293    -11.622      0.000      -3.985      -2.835
Occupation type_Accountants         -2.8988      0.090    -32.302      0.000      -3.075      -2.723
Organization Type_Electricity       -5.2689      1.003     -5.255      0.000      -7.234      -3.304
Organization Type_Industry: type 3  -3.3706      0.322    -10.482      0.000      -4.001      -2.740
Organization Type_Military          -3.0647      0.184    -16.679      0.000      -3.425      -2.705
Organization Type_School            -2.6126      0.128    -20.489      0.000      -2.862      -2.363
================================================================================================
```

Рисунок 1.34 – Результати оцінки та відбору параметрів логістичної регресії

На рисунку 1.34 можна побачити, що найбільш значущими виявилися 7 параметрів логістичної регресії. Також було отримано від’ємне значення коефіцієнту детермінації. Оскільки в даному випадку розраховувався псевдо-R2, то він не має корисності, тому що не обмежений знизу. Тому для перевірки моделі на адекватність доцільно використати ROC-криву. Результати її побудови представлені на рисунку 1.35.

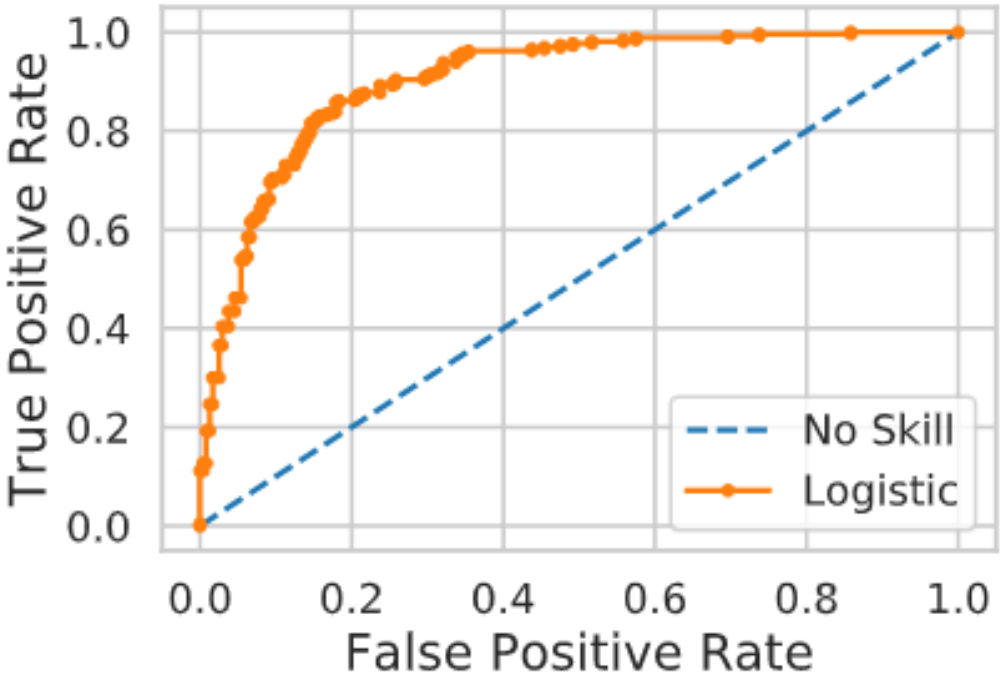


Рисунок 1.35 – ROC-крива для логістичної регресії

На рисунку 1.35 показано, що логістична модель має істинне позитивне значення, яке наближається до верхнього лівого кута (до 1), тобто ROC-крива показує високу залежність кількості правильно класифікованих позитивних прикладів від кількості неправильно класифікованих негативних прикладів. Результати логістичної регресії є придатними для оцінки ймовірності виникнення шахрайств в процесі кредитування. Отриманні значення запишемо у вигляді математичної моделі логістичної регресії – формули (1.17):

$$P = \frac{1}{1 + E^{-2.579-4.1769x_1-3.4096x_2-2.8988x_3-5.2689x_4-3.3706x_5-3.0647x_6-2.6126x_7}} \qquad (1.17)$$

Прогнозні оцінки експоненти цієї моделі вказують на те, що коли буде змінюватися значення незалежних змінних $x_1, x_2, \dots, x_7$ на 1, ймовірність шахрайської операції буде зростати або зменшуватися у кількість разів, що відповідає визначеному значенню параметра. Наприклад, при зміні значення сімейного статусу (x_2), ймовірність шахрайства знизиться у 3,4096 разів.

Наступний метод математичного моделювання – дерево рішень. Спочатку для його побудови береться весь набір даних, що представляється кореневою вершиною. Потім визначаються варіанти розбивки даних на гілки, що відповідають кореневому вузлу. Дані гілки утворюють дерево, повернене корою вниз. Способи розбивки множини даних називають вирішальним правилом, яке відбувається за формулою (1.18):

$$a_{ik} = \begin{cases} 1, \ s_i = r_k; \\ 0, \ s_i \neq r_k, \end{cases} \tag{1.18}$$

де $a_{ik} = 1$, якщо умова s_i для правила r_k виконується;

$a_{ik} = 0$, якщо умова s_i для правила r_k не виконується;

$S\{s_i\}, i = \overline{1, l}$ – множина умов, що описують параметри обраної предметної області.

Дане правило фактично являє собою алгоритм «якщо, …то..» та ділить множину записів на дві частини [**Ошибка! Источник ссылки не найден.**].

Перевірка точності наборів даних виявила, що точність тестового набору дорівнює 0,9915, а тренувального – 1,0. Це означає, що точність прогнозування майже дорівнює 100%. Дерево рішень показало кращий результат ніж логістична регресія у значеннях точності опису залежної змінної незалежними змінними. Після цього проведемо побудову дерева рішень, а також здійснимо його навчання з метою отримання найкращої комбінації даних. Його результати представлені на рисунку 1.36.

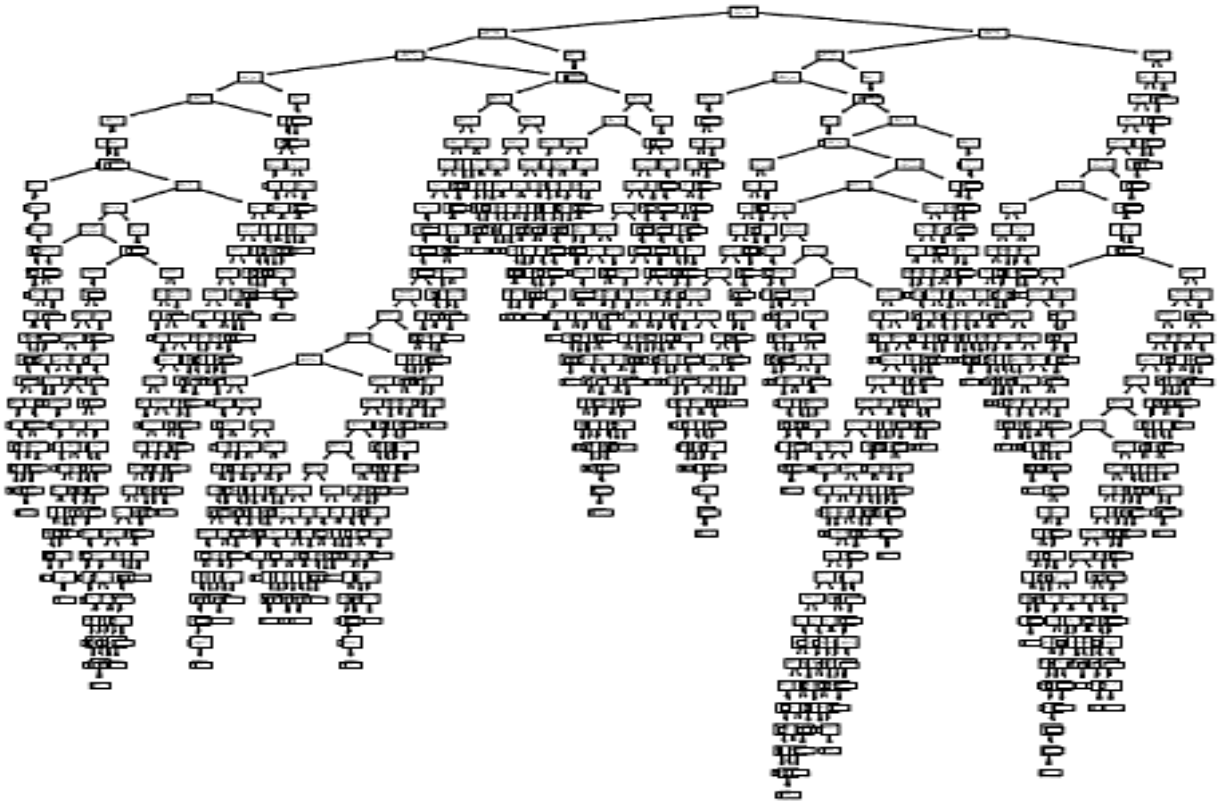


Рисунок 1.36 – Дерево рішень

Побудувавши дерево рішень можна зробити висновок, що воно має досить складну для інтерпретації структуру, хоча має високу точність прогнозування цільової змінної, ніж логістична регресія.

Третьою моделлю є нейронна мережа, яка представляє собою алгоритм, що поєднує в собі біологічні принципи та вдосконалену статистику для вирішення задач у різних сферах. Нейронна мережа приймає базову модель нейронних аналогів, пов'язаних між собою різними способами.

Проведена перевірка точності вибірок виявила, що нейронна модель забезпечує гарну точність, що є вищою за базову (66%): для тренувального набору 100,00 %, для тестового – 86,67%. Дана модель може бути удосконалена за допомогою зміни аргументів в процесі здійснення оцінювання параметрів нейронної мережі або шляхом перехресної перевірки.

В результаті побудови та проведення навчання нейронної мережі було отримано модель, графічна інтерпретація якої представлена на рисунку 1.37.

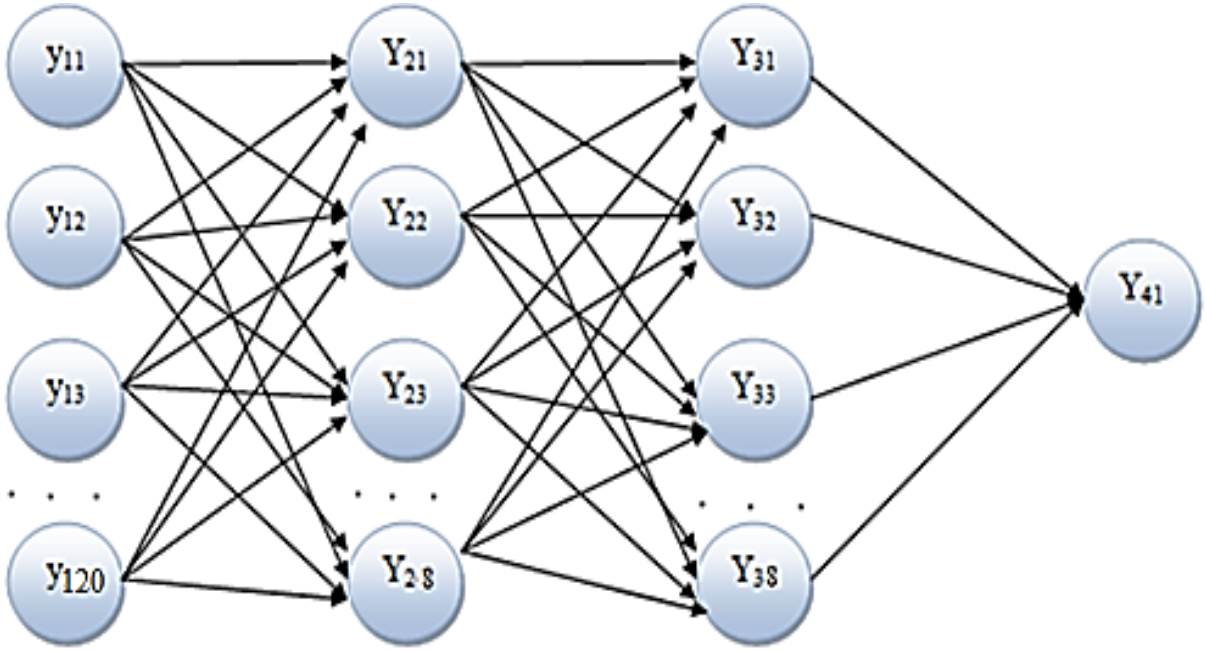


Рисунок 1.37 – Графічна інтерпретація отриманої нейронної моделі

На рисунку 1.37 можна побачити, що мережа має три приховані шари. В першому буде 120 вихідних змінних, у другому – 28, у третьому – 38. На виході отримаємо фінальне рівняння. Фрагмент математичної інтерпретації нейронної мережі представлений формулами (1.19)-(1.26):

$$y_{11} = -0.1161x_1 - 0.5501x_2 + 0.3569x_3 + \cdots + 0.2065x_8 \quad (1.19)$$

$$y_{12} = -0.6962x_1 + 0.3309x_2 + 0.2792x_3 + \cdots + 0.2065x_8 \quad (1.20)$$

…

$$y_{120} = 0.3499x_1 - 0.4575x_2 + 0.0761x_3 + \cdots + 0.3483x_8 \quad (1.21)$$

$$y_{21} = -0.3926x_1 - 0.6920x_2 + \cdots - 0.1789x_8 \quad (1.22)$$

…

$$y_{28} = 0.5144x_1 - 0.0822x_2 + \cdots - 0.2014x_8 \quad (1.23)$$

$$y_{31} = -4.4977x_1 + 9.19e^{-01}x_2 + 0.27 + \cdots - 3.5448e^{-01}x_8 \quad (1.24)$$

…

$$y_{38} = 4.44e^{-02}x_1 - 3.4965e^{-02} + \cdots + 3.4229e^{01}x_8 \quad (1.25)$$

$$y_{41} = 0.2948x_1 - 0.92x_2 + \cdots - 0.6388x_8 \quad (1.26)$$

Після побудови трьох моделей проведемо їх оцінку за точністю та якістю опису моделей тренувальним та тестовим набором даних (див. табл. 1.9).

Таблиця 1.9 – Порівняння точності та якості побудованих моделей

Назва моделі	Точність, %		MSE	
	Тестові дані	Тренувальні дані	Тестові дані	Тренувальні дані
Дерево рішень	99,15	100,00	0,008	0,000
Логістична регресія	96,60	93,00	0,064	0,069
Нейронна мережа	86,70	100,00	0,089	0,004

Отримані результати точності моделей (див. табл. 1.9), дозволяють зробити висновок, що дерево рішень найкраще моделює ймовірність шахрайства в процесі кредитування банківських клієнтів, оскільки її точність і для тестового, і для тренувального практичного дорівнює 100%. Відповідно, значення середньоквадратичної похибки (MSE) є дуже малим та наближається до 0. Логістична регресія та нейронна мережа є майже рівноцінними, оскільки: логістична регресія має вищу точність для тестових даних, ніж нейронна мережа, а нейронна мережа, навпаки, має вищу точність для тренувального набору даних. Що стосується середньоквадратичної похибки, то її значення для обох моделей є також малим та наближається до 0. В цілому, усі три моделі дають гарні результати, тому їх можна застосовувати для оцінювання ймовірності виявлення шахрайства у процесі кредитування клієнтів банків.

Проблема шахрайств у фінансовому секторі на сьогоднішній день є досить актуальною, що пояснюється впливом різних факторів. Тому на практиці існує потреба не тільки у виявленні таких випадків, але й у попередженні їх настання. Це можливо здійснити тільки із використанням сучасних інформаційних технологій та математичних методів. У даному дослідженні ця проблема вирішувалася за допомогою побудови трьох математичних моделей, що належать до класу інтелектуального аналізу даних, а саме логістичної регресії, дерева рішень та нейронної мережі. Відповідні розрахунки було проведено із використанням сучасної мови програмування Python. Дослідження передбачало формування такого набору даних, який включав не тільки кількісні

характеристики клієнтів банку (дохід, депозити), але й якісні параметри – рівень освіти, тип зайнятості, сімейний стан, тип житла та ін. В результаті було отримано математичну модель логістичної регресії, яка показала досить високі значення частки правильних прогнозів у тренувальній вибірці (93,09%) та тестовій (93,60%). Тобто обидві вибірки є адекватними та демонструють високу точність прогнозування. В результаті проведеного відбору найбільш значущих параметрів було побудовано логістичну регресію із використанням семи змінних. Її гарну якість підтвердила ROC-крива. Також було побудовано дерево рішень, модель якого продемонструвала точність наборів даних вищу, ніж для логістичної моделі (для тестового набору – 0,9915, а для тренувального – 1,0). Не дивлячись на її кращі результати, модель виявилася дуже складною для інтерпретації. Нейронна мережа показала гарні показники точності вибірок: для тренувального набору 100,00 %, для тестового – 86,67%. На останньому етапі було проведено розрахунок точності та якості моделей, в результаті чого найкращі результати продемонструвала модель дерева рішень, а нейронна мережа та логістична регресії також показали гарні результати, хоча й дещо нижчі, ніж для дерева рішень. Щоб мати постійне уявлення про ймовірні шахрайства результати повинні регулярно доповнюватися, оновлюватися для використання їх у фінансових установах, що дозволить вчасно реагувати на злочинні дії та попереджати їх виникнення у процесі надання кредиту.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден., Ошибка! Источник ссылки не найден.**].

1.3 Аналіз можливих сценаріїв взаємодії систем кібербезпеки та протидії фінансовим злочинам

1.3.1 Збалансованість детермінант розвитку країн: барицентрична модель

Динамічні процеси, які відбуваються зараз у суспільстві, призводять до того, що деякі сфери його життєдіяльності розвиваються не досить рівномірно. Це можна спостерігати на прикладі стрімкого розвитку інформаційних технологій, який за останнє десятиліття призвів до трансформації багатьох процесів економічної, політичної та соціальної сфер. В даному контексті його наслідки є в більшій мірі позитивними для суспільства та країни, оскільки призводять до побудови нових компаній сфери ІТ, створення нових робочих місць, розширення можливостей людини за рахунок використання нею останніх розробок. Що стосується інших аспектів, які також впливають на розвиток країни, то їх вплив може бути як позитивним, так й негативним. Відповідно ця проблема потребує вивчення та удосконалення в частині визначення тих детермінант, які впливають на забезпечення сталого розвитку країни за умови балансування між потребами суспільства та захистом інтересів майбутніх поколінь, а також рівномірного розвитку усіх сфер життєдіяльності суспільства. Так, це повинно узгоджуватися із цілями сталого розвитку Організації Об'єднаних Націй, які було оголошено 25 вересня 2015 року на Саміті ООН, а також із стратегіями розвитку країн, що розробляються відповідно до їх пріоритетів. Саме тому метою даного дослідження було обрано визначення рівня збалансованості соціальних, економічних, політичних детермінант та детермінант цифрової спроможності і кібербезпеки, як композитних таргетів, характерних для будь-якої країни світу. Побудова відповідної моделі дозволить визначити ті таргети, що впливають на незбалансованість розвитку країни, а також окреслити відповідні напрямки реалізації державної політики уряду країни щодо розробки ефективних стратегій, які зазначають пріоритети для підвищення рівня добробуту населення, якості соціальних стандартів та рівня його життя, подолання політичних та військових конфліктів, вирішення екологічних проблем в умовах тих викликів перед суспільством, які генерують глобальні проблеми.

Збалансований розвиток країни передбачає, що зміни, які відбуваються, носять системний характер та мають рівномірний вплив на всі сфери. Це може

забезпечуватися рядом детермінант, серед яких найбільший вплив мають економічні [**Ошибка! Источник ссылки не найден.**]. Суттєвий дисбаланс в економіці країн викликають недосконаліть законодавчої бази та наявність корупційної складової, вплив яких порушує макроекономічну стабільність [**Ошибка! Источник ссылки не найден.**]. Автори [**Ошибка! Источник ссылки не найден.**] емпірично довели, що криза довіри у фінансовому секторі також дестабілізує економіку. Автори [**Ошибка! Источник ссылки не найден.**] досліджували залежність макроекономічної стабільності від фіскальної децентралізації та зробили акцент на децентралізації витрат, децентралізації доходів та децентралізації витрат одночасно. Формування системи фінансування сталого розвитку є одним з головних стратегічних пріоритетів, що повинно відбуватися з урахуванням особливостей функціонування корпоративного сектору [**Ошибка! Источник ссылки не найден., Ошибка! Источник ссылки не найден.**]. Автори [**Ошибка! Источник ссылки не найден.**] доводять, що грошові кошти є не тільки платіжним засобом, але й виступають інструментом пропаганди та відмивання нелегальних доходів, що виступає фактором розвитку тіньового сектору. Це в свою чергу впливає на інноваційний потенціал [**Ошибка! Источник ссылки не найден.**]. Формування сприятливого інвестиційного клімату в країні є одним з напрямів підвищення рівня добробуту населення країни, що математично доведено в роботі [**Ошибка! Источник ссылки не найден.**]. Оптимальний розподіл частки приватних та державних інвестицій було змодельовано в контексті економічного розвитку [**Ошибка! Источник ссылки не найден.**].

Рівень розвитку країни суттєво впливає на умови забезпечення суспільного добробуту, формування стійких парадигм його покращення, а також призводить до якісних змін у суспільних відношеннях. З іншого боку, соціальні детермінанти формують відповідну модель суспільного життя, наслідки якої є драйвером сталого розвитку [**Ошибка! Источник ссылки не найден.**]. Для даного аспекту є важливим забезпечення саме соціальної безпеки, коли існують мінімальні ризики життєдіяльності населення в країні [**Ошибка! Источник ссылки не**

найден.]. Формування кращого соціального клімату в країні є одним з головних джерел залучення інвестицій [**Ошибка! Источник ссылки не найден.**]. Також забезпечення якісної освіти [**Ошибка! Источник ссылки не найден.**] (Lyeonov & Liuta, 2016) та системи охорони здоров'я [**Ошибка! Источник ссылки не найден.**] формують модель успішного суспільства. Тріада впливу економічних, соціальних та політичних детермінант в контексті забезпечення зростання економічної безпеки країн була проаналізована та із використанням методу експоненційного згладжування було зпрогнозовано рівень інноваційних змін [**Ошибка! Источник ссылки не найден.**].

Окрім економічних, політичних та соціальних детермінант на розвиток країни можуть впливати й екологічні фактори. Дане питання було досліджено [**Ошибка! Источник ссылки не найден.**] та визначено, що існує зв'язок між валовим внутрішнім продуктом на душу населення, викидами парникових газів, відновлюваними джерелами енергії у загальному кінцевому споживанні енергії та екологічні інвестиції, а також їх взаємодія позитивно впливає на розвиток окремих сфер життєдіяльності. Автори [**Ошибка! Источник ссылки не найден.**] досліджували взаємозв'язки між економічними, соціальними та екологічними аспектами розвитку та побудували для України та ЄС екологічну криву Кузнеця. Синергетичний ефект від взаємодії зелених інвестицій та інституційними детермінантами проявляється у національній економіці та призводить до зниження її енергоефективності [**Ошибка! Источник ссылки не найден.**]. Існування конвергенції між податковою та екологічною системами було доведено на основі бета- та сігма-конвергенцій [**Ошибка! Источник ссылки не найден.**].

Наслідки четвертої промислової революції сприяли цифровізації багатьох процесів, що впливає, в першу чергу, на динамічність розвитку економіки країни та підвищує рівень її національної безпеки. Новіковим В. [**Ошибка! Источник ссылки не найден.**] на основі бібліометричного аналізу досліджень довів, що збалансованість розвитку країни в більшій мірі залежить від її соціальної, економічної та інформаційної безпеки. Процеси цифровізації є актуальними

також й для фінансового сектору економіки, де існують найбільші потреби у цифровізації фінансових послуг. В умовах зростання інформаційних потоків повинна забезпечуватися конфіденційність великих даних. Паралельно також зростають ризики фінансових втрат завдяки здійснення масових кібератак, що призводить до дестабілізації процесів та систем, а також гальмування їх розвитку [**Ошибка! Источник ссылки не найден.**]. Хоча більшість країн намагаються вирішувати цю проблему шляхом застосування технологій штучного інтелекту, але нажаль попередження кібератак залишається важливою задачею забезпечення національної безпеки країн. Тому при визначенні детермінант слід враховувати не тільки ті фактори, що характеризують розвиток ІТ галузі, але й напрям інформаційної безпеки [**Ошибка! Источник ссылки не найден.**].

Для проведення дослідження збалансованості розвитку країн обираємо тріаду економічних, політичних та соціальних детермінант, а також детермінант, що характеризують розвиток інформаційних технологій та кібербезпеки, групу яких буде названо як цифрова спроможність а кібербезпека. Оскільки екологічні фактори мають більш вузько направлений вплив на розвиток країни, то в даній роботі їх не буде враховано для побудови моделі.

Для моделювання в сфері економічного, політичного, соціального, інформаційного розвитку застосовується широкий спектр математичних методів. Ці проблеми вирішувались науковцями шляхом побудови оптимізаційних моделей [**Ошибка! Источник ссылки не найден.**], структурного моделювання [**Ошибка! Источник ссылки не найден.**], методів інтелектуального аналізу даних [**Ошибка! Источник ссылки не найден.**], гравітаційного моделювання [**Ошибка! Источник ссылки не найден.**], нечітких множин [**Ошибка! Источник ссылки не найден.**], ймовірнісних методів [**Ошибка! Источник ссылки не найден.**], економетричного інструментарію [**Ошибка! Источник ссылки не найден.**], статистичного аналізу [**Ошибка! Источник ссылки не найден.**].

На збалансованість розвитку країн можуть впливати різні детермінанти, які або призводять до підвищення її рівня, або знижують його. Для

обґрунтування їх вибору було застосовано методи наукового пізнання, які дозволили визначити найбільш релевантні показники для кожного композитного таргету. Так, вимір цифрової спроможності і кібербезпеки формується під впливом тенденцій розвитку ІТ-галузі та її складових, рівня цифрового розвитку та безпекової складової. Оскільки не існує єдиних підходів до визначення даного виміру, то в дану групу було віднесено 5 ключових індикаторів, які характеризують: слабкі сторони країн щодо кібербезпеки та покращення можливостей для них шляхом розробки стратегії кібербезпеки та відповідних стандартів (The Global Cybersecurity Index – GCI), рівень готовності країн протидіяти кіберзагрозам та керувати кіберінцидентами (The National Cyber Security Index – NCSI), рівень розвитку інформаційних та комунікаційних технологій в країні (ICT Development Index – ICTDI), ступінь технологічної готовності країни для застосування новітніх інформаційно-комунікаційних технологій в різних сферах життєдіяльності (Networked Readiness Index – NRI), ступінь відповідності цифровізації країни рівню її кібербезпеки для формування рекомендацій щодо коректування програм кібербезпеки (Digital Development Level – DDL). Оскільки значення цих індикаторів позитивно впливає на інтегральне значення виміру цифрової спроможності і кібербезпеки, тобто із зростанням їх значення підвищується його рівень, то враховуємо їх як показники-стимулятори. Вважається, що країна, для якої характерне високе значення композитного таргету цифрової спроможності і кібербезпеки, має потужний розвиток інформаційних технологій та вважається країною із найвисоким рівнем інформаційної безпеки.

Фактори економічного розвитку країни є ключовим компонентом у здійсненні його збалансованості, оскільки дозволяють оцінити рівень добробуту громадян (The Global Competitiveness Index), умови ведення бізнесу у країні та захисту прав власності (Ease of Doing Business), вплив розвитку фінансових систем на зростання, стабільність та нерівність різних економік (Financial Development Index), етнічну, расову, регіональну, освітню, тощо нерівність, яка формує економічну різницю між даними групами, що врешті решт впливає на

економічний розвиток країни (Uneven Economic Development Index), людські можливості контролювати власну працю та майно, рівень власного споживання та інвестування (Economic Freedom Index). Чим вище рівень економічного розвитку країни, тим більше можливостей для неї бути лідером на світових ринках та забезпечувати високий рівень життя її населення. Серед обраних індикаторів тільки Uneven Economic Development Index є показником дестимулятором, із збільшенням значення якого зростає інтегральний рівень таргету економічного розвитку. Інші індикатори є стимуляторами за своєю сутністю.

Соціальний вимір направлений на визначення спроможності країни забезпечити населенню високий рівень життя, що полягає у створенні сприятливих умов для отримання населенням таких соціальних благ, як освіта, якісні медичні послуги, рівень «екологічного сліду», забезпечення та підтримання миру в середині країни, тощо. Для аналізу даного композитного таргету було обрано індикатори, що вимірюють якість поточного життя населення (Happiness Index), рівень забезпечення країнами основних потреб людини, їх добробуту та можливостей для прогресу (Social Progress Index) та таких основних характеристик людського потенціалу, як рівень життя, грамотності, освіченості і довголіття (Human Development Index). Обрані детермінанти є показниками-стимуляторами, тому високе значення інтегрального рівня соціального виміру свідчитиме про високий рівень забезпечення соціальних стандартів для життя населення в даній країні.

Політичний розвиток будь-якої країни є неодмінна частина її загального розвитку, оскільки характеризує динаміку політичного життя країни, її можливості взаємодіяти із іншими країнами у зовнішньому політичному просторі, налаштовувати діалог між державою та населенням. Існування політичних коливань може призводити до дестабілізації соціальних настроїв населення та гальмування розвитку економіки, тому його вимірювання є вкрай важливим для визначення рівня збалансованого розвитку країни. Тому для визначення його інтегрального рівня було обрано індикатори, які вимірюють:

рівень ймовірності того, що уряд країни може бути дестабілізований або зруйнований засобами, які носять неконституційний та насильницький характер (Political Stability Index), якість демократії у країні з урахуванням оцінок виборчого процесу, громадянської свободи, функціонування уряду, політичної культури (Democracy Index), якість діяльності уряду країни на основі оцінки якості державних послуг та органів, якості формування та реалізації політичних заходів, ступеня незалежності від політичного тиску, тощо (Government Effectiveness Index), рівень корупційної складової державного сектору (Corruption Perceptions Index). Зростання значень обраних детермінант впливає на зростання рівня композитного таргету, що свідчить про політичну стабільність в країні та високий рівень її політичного розвитку.

Збалансованість будь-якої системи – це її стан, при якому забезпечується оптимальне співвідношення її компонентів, яке дозволяє їй знаходитися у рівновазі та бути стійкою у випадку впливу зовнішніх факторів. Відповідно збалансованість розвитку країни показує рівномірний або зважений розвиток її компонентів, що забезпечує його стійкість протягом тривалого часу. Для його моделювання найбільш оптимальними є моделі, які базуються на визначені центру мас. Тобто в залежності від кількості компонент, які приймають учать у вимірі збалансованості розвитку, будується геометрична фігура, вершинами якої виступають композитні їх значення, які в свою чергу формуються під впливом різних детермінант. В даному дослідженні було обрано чотири основні сфери – економічна, політична, соціальна та сфера цифрової спроможності і кібербезпеки, які представляють собою найбільш впливові на розвиток будь-якої країни компоненти або таргети. При цьому їх формування здійснюється на основі обраних детермінант, що найбільше характеризують їх розвиток. Відповідно модель, що будується у даному дослідженні, є барицентричною.

Популяризацією підходу визначення центру мас для економічних наук займалися автори праці [**Ошибка! Источник ссылки не найден.**]. Вони розробили модель трикутника для визначення стійкості ринку страхування та перестрахування, де зробили акцент на розрахунку та аналізі радіусу описаного

кола. Методологію побудови барицентричної моделі для аналізу ділової активності компаній запропонували автори [**Ошибка! Источник ссылки не найден.**], де не передбачалася графічна інтерпретація моделі та були відсутні практичні розрахунки. Її розвиток було продовжено Яровенко Г.М. [**Ошибка! Источник ссылки не найден.**] для визначення рівня збалансованості розвитку національної економіки.

Методика базується на визначенні та проведенні аналізу трьох компонентів барицентричної моделі: значень композитних вимірів, збалансованості пар таргетів та збалансованості всіх чотирьох таргетів.

Для визначення значень композитних таргетів необхідно провести нормалізацію значень тих детермінант, які здійснюють вплив. Ця процедура необхідна, оскільки обрані фактори мають різну природу та відрізняються значеннями їх абсолютних величин. Проведення нормалізації дозволить звести значення всіх факторів в діапазоні від 0 до 1. Відповідно це спростить процес згортки даних для визначення композитного значення економічного, соціального, політичного вимірів та виміру цифрової спроможності та кібербезпеки, яке також буде знаходитися в діапазоні від 0 до 1. Якщо значення композитного таргету буде наближатися до 1, то це свідчитиме про потужний розвиток відповідної сфери життєдіяльності країни. В протилежному випадку, якщо воно буде ближчим до 0, то це показник гальмування розвитку.

Існують різні види нормалізації даних, але в даній роботі буде використано лінійну нормалізацію для показників-стимуляторів (1.27) та нормалізацію Севіджа для дестимуляторів (1.28), оскільки дані дослідження є просторовими:

$$\widetilde{x_{ik}} = \frac{x_{ik} - x_{min_i}}{x_{max_i} - x_{min_i}}, \tag{1.27}$$

$$\widetilde{x_{ik}} = \frac{x_{max_i} - x_{ik}}{x_{max_i} - x_{min_i}}. \tag{1.28}$$

де $\widetilde{x_{ik}}$ – нормалізоване значення *i*-го фактору економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки для *k*-ої країни;

x_{ik} – вхідне значення *i*-го детермінанту економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки для *k*-ої країни;

x_{min_i} та x_{max_i} – відповідно мінімальне та максимальне значення *i*-го детермінанту економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки серед сукупності спостережень, тобто країн.

Розрахунок композитного таргету для економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки відбувається на основі середньогеометричної функції (1.29). Її вибір обумовлений тим, що в результаті отримуємо середнє пропорційне значення таргету для кожної країни:

$$G(\widetilde{x_1}, \widetilde{x_2}, \dots, \widetilde{x_n}) = \left(\prod_{i=1}^{n} \widetilde{x_i} \right)^{1/n}, \tag{1.29}$$

де $G(\widetilde{x_1}, \widetilde{x_2}, \dots, \widetilde{x_n})$ – середньогеометричне значення сукупності нормалізованих значень детермінант економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки, визначене для *k*-ої країни;

n – кількість детермінант, які формують відповідний композитний таргет.

При визначенні середньогеометричного значення ті фактори, нормалізоване значення яких дорівнює 0, відбувається значне зміщення значення

композитного таргету у бік 0. То для усунення цього фактору для таких значень можна скористатися формулою Мінковського (1.30):

$$R(x_i) = 1 - \sqrt{\sum_{j=1}^{k} \omega_j \left|1 - \frac{x_{ij}}{x_{max_j}}\right|^2 + \sum_{j=k+1}^{n} \omega_j \left|1 - \frac{x_{min_j}}{x_{ij}}\right|^2}, \quad (1.30)$$

де *R(xi)* – значення композитного таргету економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки;

ω_j – вага кожного детермінанту при формуванні композитного таргету при чому $\sum_{j=1}^{n} \omega_j = 1$. Для їх визначення можна провести канонічний аналіз, побудувати стандартизоване рівняння регресії, або врахувати їх рівномірний вплив на формування таргету, що було використано в даній роботі.

Для визначення збалансованості пар таргетів та всіх чотирьох таргетів необхідно побудувати чотириполюсну барицентричну модель, що здійснюється як побудова геометричної фігури - чотирикутника та визначення його основних характеристик. Даний процес передбачає відкладення чотирьох точок на координатній площі, координати яких відповідають значенням композитних таргетів. Але щоб розуміти наскільки збалансованим є розвиток країни доцільно поряд із моделлю фактичних даних будувати й еталонну, в якості якої виступає квадрат, координати вершин якого дорівнюють максимальному значенню таргету, тобто 1. Для виміру цифрової спроможності і кібербезпеки це буде точка із координатами (1; 1), для соціального виміру – (1; -1), економічного – (-1; -1), політичного – (-1; 1). Точки з'єднуються прямими, які утворюють боки квадрату. Його центроїд або центр мас знаходиться у точці перетину його діагоналей ("Center of Mass"), яка співпадає із початком координатних вісей і має координатами (0; 0). Еталонна модель представлена на рисунку 1.38. Її було побудовано із використанням програмного забезпечення GeoGebra.

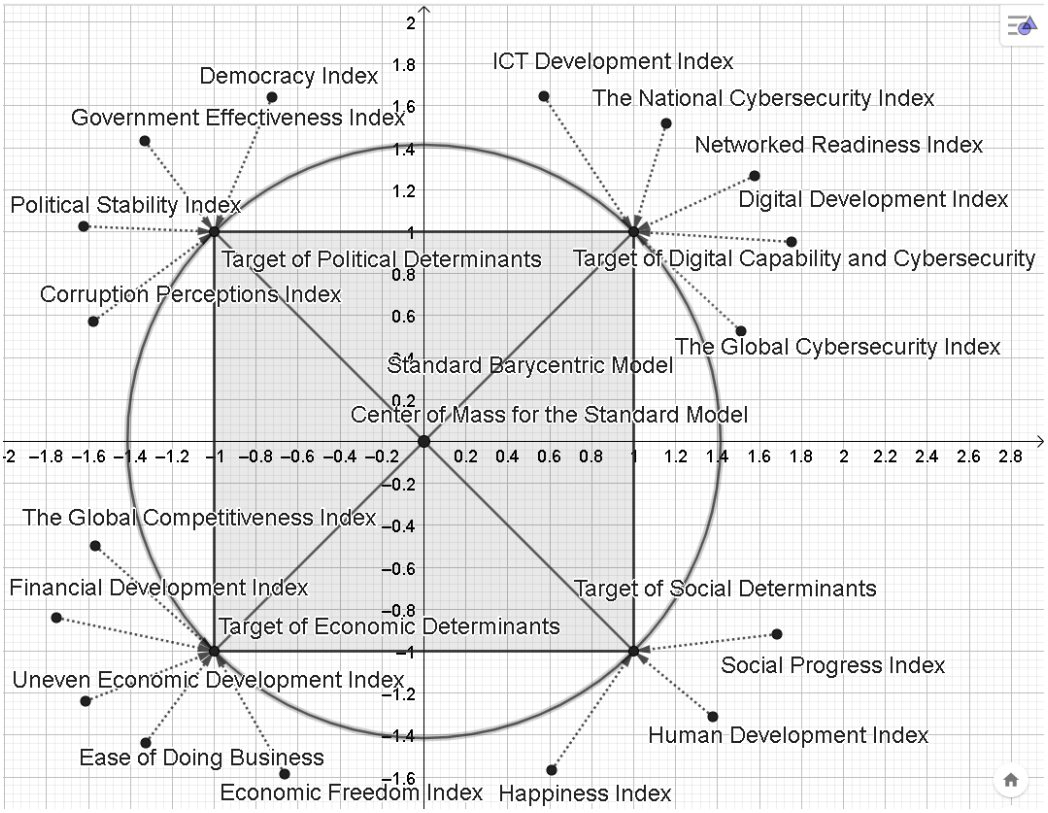


Рисунок 1.38 – Еталонна чотириполюсна барицентрична модель збалансованості розвитку країни

Для емпіричних даних побудувати барицентричну модель у вигляді квадрату досить складно, оскільки за таких умов країна має однаковий рівень розвитку економічної, соціальної, політичної сфери та сфери інформаційної безпеки. На практиці для різних країн можна отримати різні форми чотирикутників, з різними довжинами боків та різними кутами. Тому також необхідно накреслити коло навколо чотирикутника. Це може бути можливим за умови, якщо сума його протилежних кутів дорівнює 180^{o}. В протилежному випадку даний факт свідчитиме про існування дисбалансу між парами вимірів.

Для визначення градусної міри кутів чотирикутника і перевірки можливості побудувати коло навколо нього необхідно чотирикутник поділити на два трикутники та розрахувати їх довжину сторін як довжину відрізків за формулою (1.31):

$$AB = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2}, \quad (1.31)$$

де AB – це довжина відрізку між двома точками А та В, які є вершинами одного з двох трикутників АВС;

$(x_a; y_a)$ – координати точки А, що є значеннями відповідних композитних таргетів;

$(x_b; y_b)$ – координати точки В, що є значеннями відповідних композитних таргетів.

Аналогічно знаходяться інші сторони трикутника АБС та сторони другого трикутника, що разом формують чотирикутник.

Розраховуємо косинуси кожного кута для кожного з двох трикутників за формулою (1.32):

$$\cos\alpha = \frac{b^2 + c^2 - a^2}{2 \cdot b \cdot c}, \quad (1.32)$$

де a, b, c – це значення довжин трьох сторін трикутника.

Отримані значення переводимо у градусну міру із використанням спеціальних таблиць або калькуляторів. У даному дослідженні розрахунки відбувалися із використанням програмного забезпечення MS Excel, де застосовуються відповідні функції.

Сумуємо значення градусної міри двох кутів, які знаходяться біля основи одного трикутника, із значеннями градусної міри кутів іншого, щоб отримати значення двох протилежних кутів чотирикутника. Спочатку проводимо перевірку, чи дорівнює сума чотирьох кутів 360 градусів, а потім перевіряємо умову збалансованості двох пар вимірів шляхом визначення суми пар протилежних кутів чотирикутника. У випадку, якщо їх суми дорівнюють 180

градусів, робимо висновок, що навколо даного чотирикутника можна описати коло, тобто пари вимірів є збалансованими.

Для визначення третьої компоненти барицентричної моделі (збалансованості чотирьох таргетів) необхідно визначити центр мас чотирикутника, що передбачає розрахунок значень його координат за формулами (1.33)-(1.34):

$$O_x = \frac{1}{6A}\sum_{i=0}^{n-1}((x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i)), \tag{1.33}$$

$$O_y = \frac{1}{6A}\sum_{i=0}^{n-1}((y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i)), \tag{1.34}$$

де *Ox* та *Oy* – координати точки O, яка є центром мас чотирикутника;

(x_i; y_i), (x_{i+1}; y_{i+1}) – координати вершин чотирикутника, де вершина з координатами (x_n; y_n) буде співпадати з вершиною з координатами (x_0; y_0);

A – площа чотирикутника, яка визначається за формулою (1.35):

$$\mathrm{A} = \frac{1}{2}\sum_{i=0}^{n-1}(x_i y_{i+1} - x_{i+1} y_i). \tag{1.35}$$

Визначення рівня збалансованості чотирьох таргетів здійснюється шляхом отримання різниці між центром мас, який відповідає даним певної країни, та центром мас еталонної моделі. Для цього розраховуємо дану відстань як довжину відрізку за формулою (1.35). Чим ближче отримане значення до 0, тим ближче центр мас барицентричної моделі країни до еталонного значення, що говорить про збалансований розвиток країни на основі її чотирьох композитних таргетів.

Для проведення дослідження та розрахунків було узято дані обраних детермінант для 127 країн світу за 2018 рік з джерел таких організацій, як Світовий банк, Міжнародний валютний фонд, Всесвітній економічний форум, незалежний шведський фонд «Gapminder», глобальна некомерційна організація «Імператив соціального прогресу». Даний період було обрано тому, що більшість детермінант не мають фактичних значень після нього, особливо індикатори виміру цифрової спроможності і кібербезпеки. Для аналізу отриманих результатів проведемо групування країн за рівнем їх економічного розвитку відповідно до класифікації Міжнародного валютного фонду, згідно із якою вони поділяються на розвинені, ті, що розвиваються, та найменш розвинені. Також виділимо серед них групу країн, які вважаються новими індустріальними завдяки їх високих темпів технологічного розвитку, який виступає драйвером розвитку їх економіки. Сюди відносяться Аргентина, Бразилія, Мексика, Індія, Малайзія, Тайланд, Чилі, Індонезія, Туреччина, Китай, Іран, Філіппіни [**Ошибка! Источник ссылки не найден.**], та перспективні індустріальні країни з Групи одинадцяти (Нігерія, Єгипет, Пакистан, Бангладеш, В'єтнам) [**Ошибка! Источник ссылки не найден.**].

На рисунку 1.39 представлено результати значень розрахованих композитних таргетів економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки для двадцяти розвинених країн або країн із високим рівнем економіки, перелік яких було означено у відповідністю із Міжнародним валютним фондом [**Ошибка! Источник ссылки не найден.,** с. 132]. Перші десять країн мають найвище сумарне значення таргетів, друга десятка – найнижче з групи розвинених країн.

Порівняння значень композитних таргетів із еталонним рівнем (рисунок 1.39) показує, що для більшості розвинених країн їх значення прямують до 1, але не досягають його. На практиці цього досягти не можливо для будь-якої країни, тому чим ближче розраховані значення прямують до нього, тим вищий рівень розвитку даного виміру країни. Слід відмітити, що найкращий результат демонструє Швейцарія, яка має найвище сумарне значення таргетів

економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки (3,735).

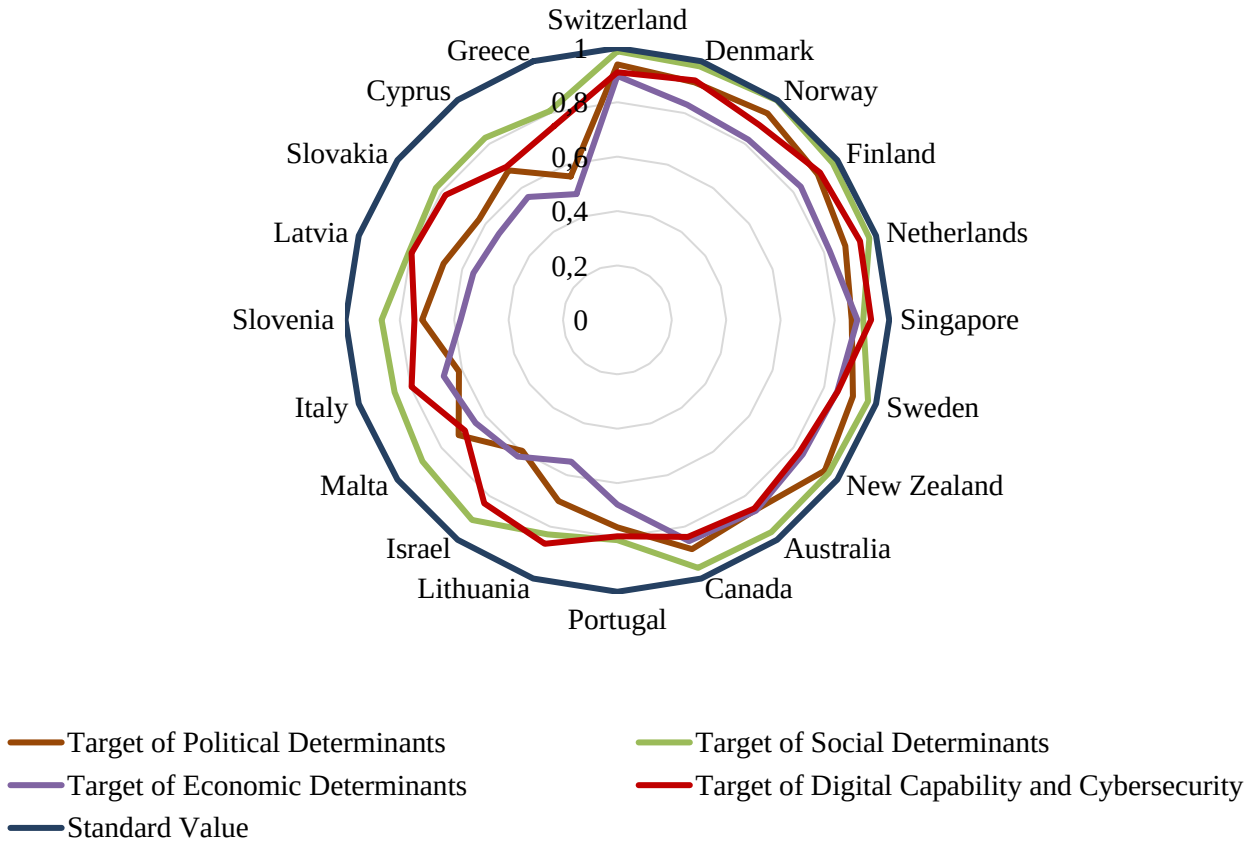


Рисунок 1.39 – Значення композитних таргетів економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки для розвинутих країн

Такі країни, як Данія, Фінляндія, Норвегія, Нідерланди, Сінгапур, Швеція, Нова Зеландія, Австралія, Канада, входять у топ-десять країн із найвищими значеннями композитних таргетів, що свідчить про досить високий рівень розвитку кожного з них. Найнижчі значення серед аналізованої групи розвинених країн демонструють Португалія, Литва, Ізраїль, Мальта, Італія, Словенія, Латвія, Словакія, Кіпр та Греція. Найвищий рівень розвитку демонструє соціальний вимір, що говорить про ефективну соціальну політику уряду цих країн по відношенню до їх населення. Політичний вимір та вимір цифрового розвитку і кібербезпеки для переважної кількості країн переважають над значеннями економічного таргету. Це свідчить про те, що економічний

потенціал цих країн надав поштовх для прискорення розвитку інших таргетів, що в подальшому сприятиме більш потужному економічному розвитку цих країн.

На рисунку 1.40 представлені результати розрахованих значень композитних таргетів економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки для країн, що розвиваються згідно із переліком Міжнародного валютного фонду. Було винесено десять країн із найвищими значеннями та десять – з найнижчими.

Провідними в даній групі країн є Польща, Арабські Емірати, Маврикій, Уругвай, Хорватія, Угорщина, Катар, Болгарія, Коста Ріка та Румунія. Найменш розвиненими є таргети таких країн, як Киргизстан, Болівія, Кенія, Алжир, Нікарагуа, Гондурас, Сурінам, Кот-Д'Івуар, Таджикистан та Камерун. У порівнянні із даними, отриманими для розвинених країн, спостерігається значний дисбаланс, що виникає між соціальним виміром і виміром кібербезпеки та економіко-політичним. При чому різниці є досить суттєвими. Наприклад, для Польщі таргет соціального виміру дорівнює 0,8357, цифрової спроможності і кібербезпеки – 0,7773, економічний – 0,6533, політичний – 0,6410, а для Алжиру – відповідно 0,6542, 0,3486, 0,1365, 0,3002. Тобто, розвиток економічної та політичної сфери є досить критичним для країн, що розвиваються. Це обумовлене або нестійкою політичною ситуацією, що склалася в них (наприклад, Україна, Гондурас, Гватемала), або неефективністю дій уряду та прийнятих ним політичних законів та рішень.

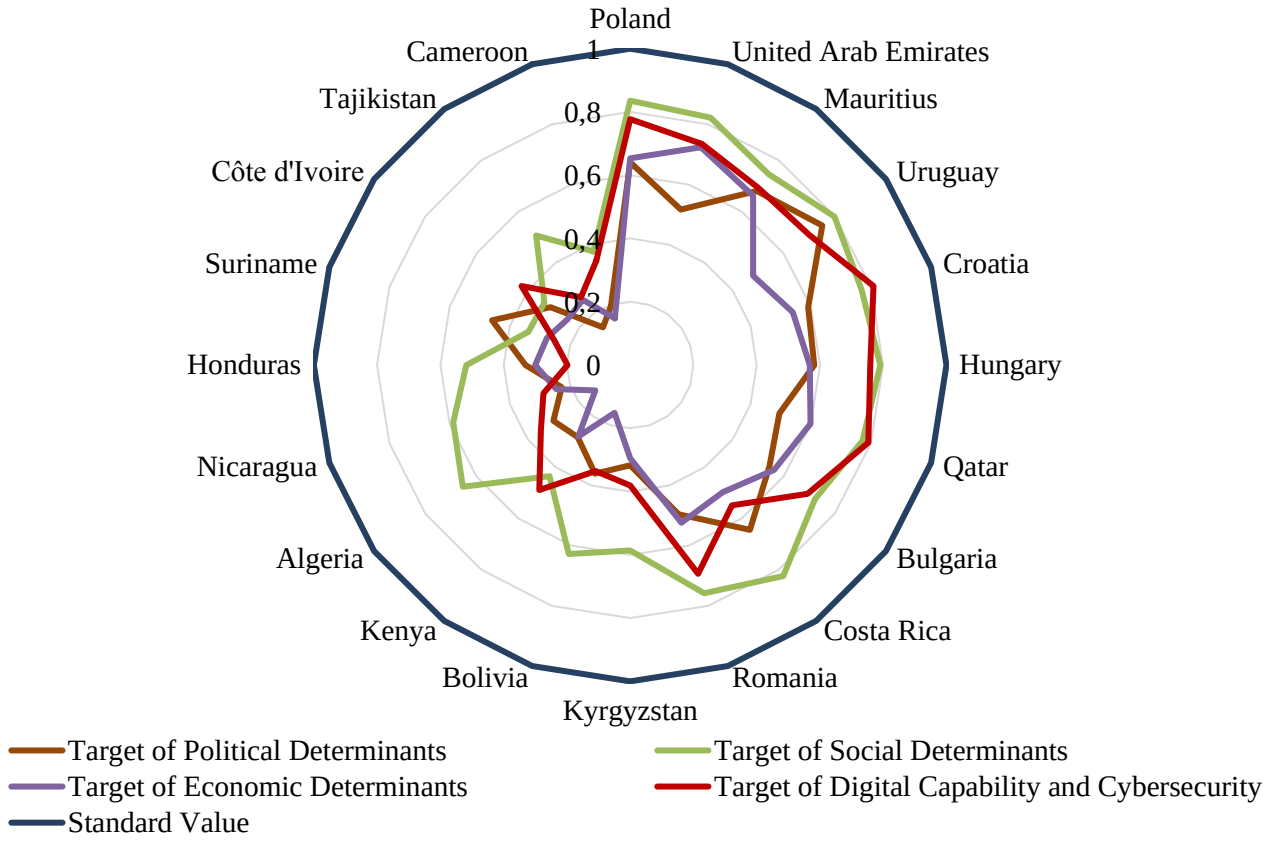


Рисунок 1.40 – Значення композитних таргетів економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки для країн, що розвиваються

В свою, чергу нестабільність економічного розвитку таких країн є прямим наслідком кризи їх політичної сфери, що призводить до гальмування їх розвитку в цілому. Тобто для країн, що розвиваються, є важливим, в першу чергу, посилення політичного виміру шляхом трансформації законодавства, боротьби з корупційною складовою, прийняття урядом більш ефективних рішень, спрямованих на розвиток економіки та проведення реформ, тощо.

Результати розрахованих значень композитних таргетів для нових індустріальних країн представлені на рисунку 1.41. У розвитку нових індустріальних країн також присутній дисбаланс, при цьому явно спостерігається однаковий напрям розвитку у соціальному вимірі та вимірі цифрової спроможності і кібербезпеки, а також економіко-політичному. Хоча на відмінність від значень композитних таргетів, представлених на рисунку 1.41,

для більшості даних країн, окрім Чилі, Аргентини та Бразилії, характерний більш рівномірний розвиток обраних сфер, який не містить аномальних перепадів. Найвищі значення таргетів мають Малайзія, Чилі, Тайланд, Туреччина та Аргентина, найгірший результат характерний для Ірану, Бангладеш, Пакистану та Нігерії. Оскільки представлені країни вважаються такими, що вже пройшли певні етапи соціо-економічного розвитку та досягли успіхів, або мають всі шанси на індустріальний стрибок, то можна сказати, що для більшості із них, а саме Туреччині, Тайланду, Аргентині, Нігерії, Пакистану, Чилі, Бразилії, Бангладеш, Мексиці та Ірану, слід звернути увагу на політичний та економічний виміри для забезпечення розвитку соціальної сфери та сфери цифрової спроможності.

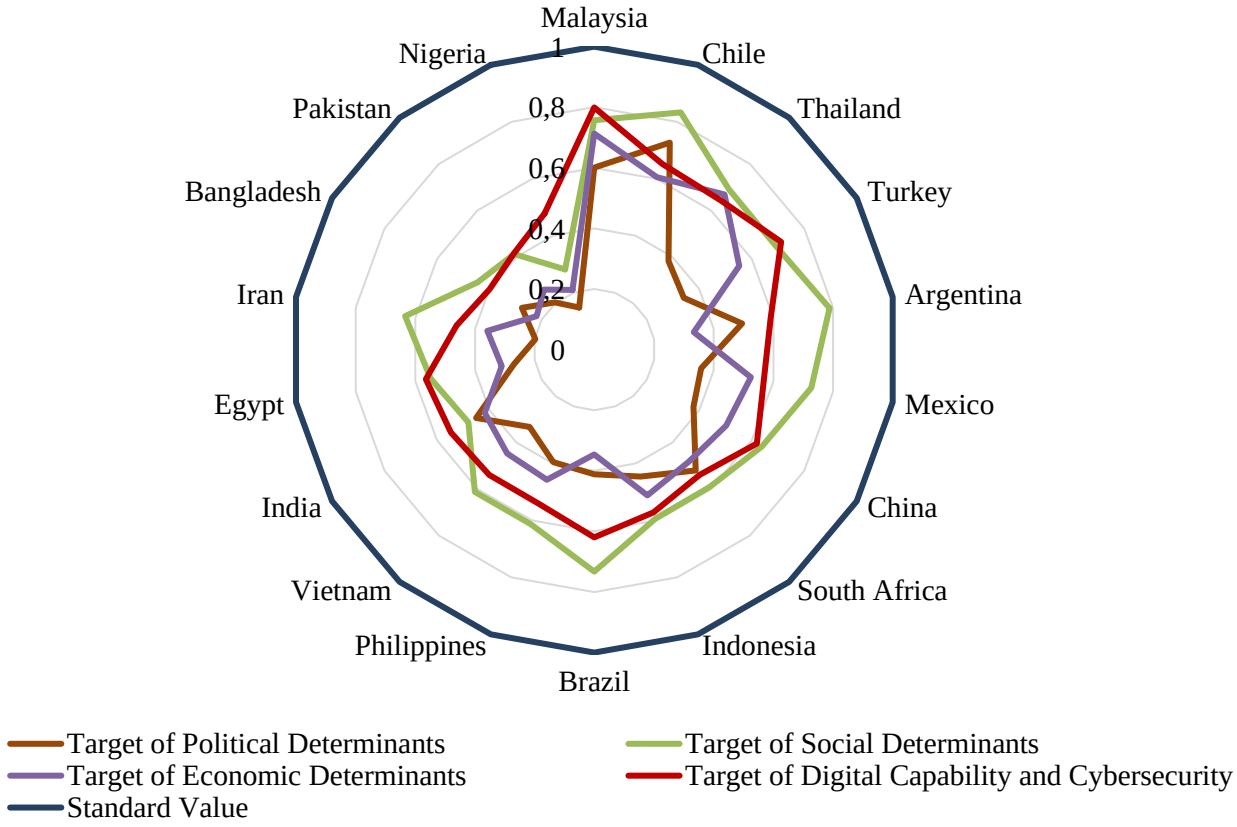


Рисунок 1.41 – Значення композитних таргетів економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки для нових індустріальних країн

На рисунку 1.42 представлені побудовані композитні таргети економічного, соціального, політичного вимірів та виміру цифрової

спроможності і кібербезпеки для найменш розвинутих країн, перелік яких визначено Організацією Об'єднаних Націй (далі ООН) [**Ошибка! Источник ссылки не найден.**, с. 1]. Було виділено 10 країн із найвищими та найменшими значеннями показників серед своєї групи країн.

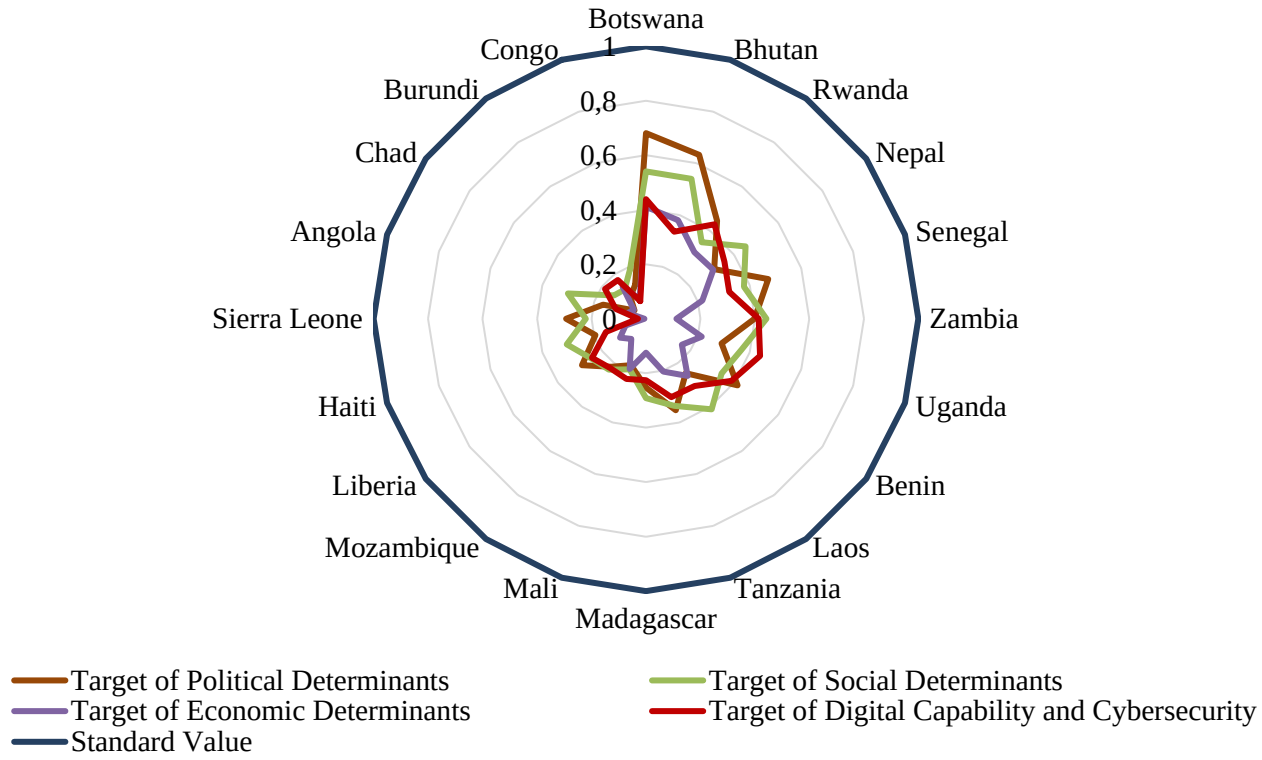


Рисунок 1.42 – Значення композитних таргетів економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки для найменш розвинених країн

Практично усі країни, окрім Ботсвани та Бутану, мають дуже низькі значення чотирьох таргетів (рисунок 1.42). При цьому можна побачити нерівномірний розвиток усіх вимірів, особливо економічного. Отримані результати свідчать про існування реальних проблем економічного, соціального, політичного характеру та недостатній рівень розвитку інформаційних технологій, що потребує допомоги з боку всесвітніх та міжнародних організацій.

Для проведення аналізу збалансованості пар таргетів було визначено градусну міру кутів чотирикутників для 127 країн світу та сформовано їх пари, які було обрано, виходячи із наступних міркувань. Найбільший поштовх

сьогодні в економіці забезпечується за рахунок розвитку саме інформаційних технологій, що призводить до поступової її трансформації у цифрову. З іншого боку, саме економічний розвиток країни стимулює науково-технічний прогрес, наслідком якого є розвиток ІТ-сфери в країні. Також дані міркування були підкріплені шляхом розрахунку лінійного коефіцієнту кореляції між значеннями чотирьох композитних таргетів. Виявилося, що між інтегральними значеннями економічного виміру та виміру цифрової спроможності і кібербезпеки існує найтісніший кореляційний зв'язок (0.9144), між парою соціального та політичного вимірів цей зв'язок є також тісним (0.8343). Візуалізуємо дані отриманих розрахунків, які у відсотках показують співвідношення сум протилежних кутів чотирикутника, що дозволяє зробити висновок про збалансованість або незбалансованість розвитку пар вимірів – соціо-політичного та економіко-цифрового (для скорочення назви виміру цифрової спроможності та кібербезпеки застосовуємо «інформаційний»). Тобто, якщо значення буде прямувати до 50% (для пари соціального та політичного вимірів) та 100% (для пари економічного виміру та виміру цифрової спроможності і кібербезпеки), то це говорить, що сума пари кутів є рівною 180 градусів, в протилежному випадку, вона буде або більшою, або меншою ніж 180. Так, на рисунку 1.43 представлені результати розрахунків для розвинених країн.

Аналізуючи дані рисунку 1.43, можна зробити висновок про те, що такі країни як Італія, Японія, Франція та Ізраїль мають найбільш збалансовані пари таргетів, оскільки значення сум пар протилежних кутів прямують до 180 градусів. Тобто при побудові їх барицентричної моделі можна накреслити коло навколо їх чотирикутника. Для Іспанії, Сінгапуру, Естонії, Великобританії, Німеччини та США сума кутів має незначне відхилення від 180 градусів, але для інших країн розбіжність зростає.

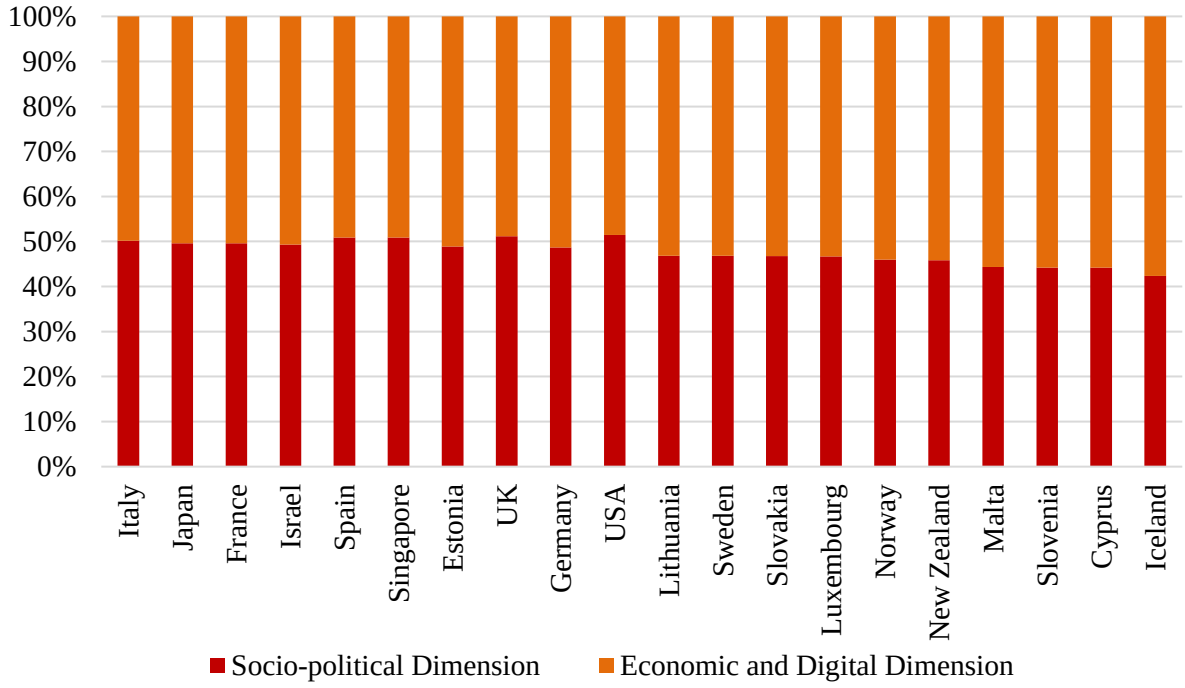


Рисунок 1.43 – Збалансованість пар композитних таргетів економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки для розвинених країн

При цьому можна побачити, що значення для соціально-політичного виміру є меншим для переважної кількості країн, що свідчить про більшу значущість цієї пари для збалансованого розвитку економічно розвинених країн, а також їх стрімкий розвиток в порівнянні з парою економічного виміру та виміру цифрової спроможності.

Результати розрахунків сум протилежних кутів для побудови барицентричної моделі країн, що розвиваються, представлені на рисунку 1.44, де можна побачити, що Ко-д'Івуар, Болгарія, Оман та Румунія мають значення, близькі до 180 градусів. Для інших країн розбіжність зростає, що свідчить про не можливість описати коло навколо чотирикутника моделі. Отримані результати показують, що для частини країн є превалюючою парою соціально-політичний вимір, а для таких, як Молдова, Грузія, Кенія, Вірменія, Катар, Арабські Емірати, Україна, Бахрейн, Казахстан, Саудівська Арабія, Азербайджан та Російська Федерація (частина з них не представлена на рисунку 1.44), превалює вимір

економічного розвитку та цифрової спроможності і кібербезпеки. Аналіз цієї пари вимірів для цих країн показав, що найбільш потужний розвиток вони мають в сфері ІТ, а економічний розвиток значно відстає. Тому у випадку цих країн економічний прорив можливий за рахунок потужного потенціалу ІТ сфери.

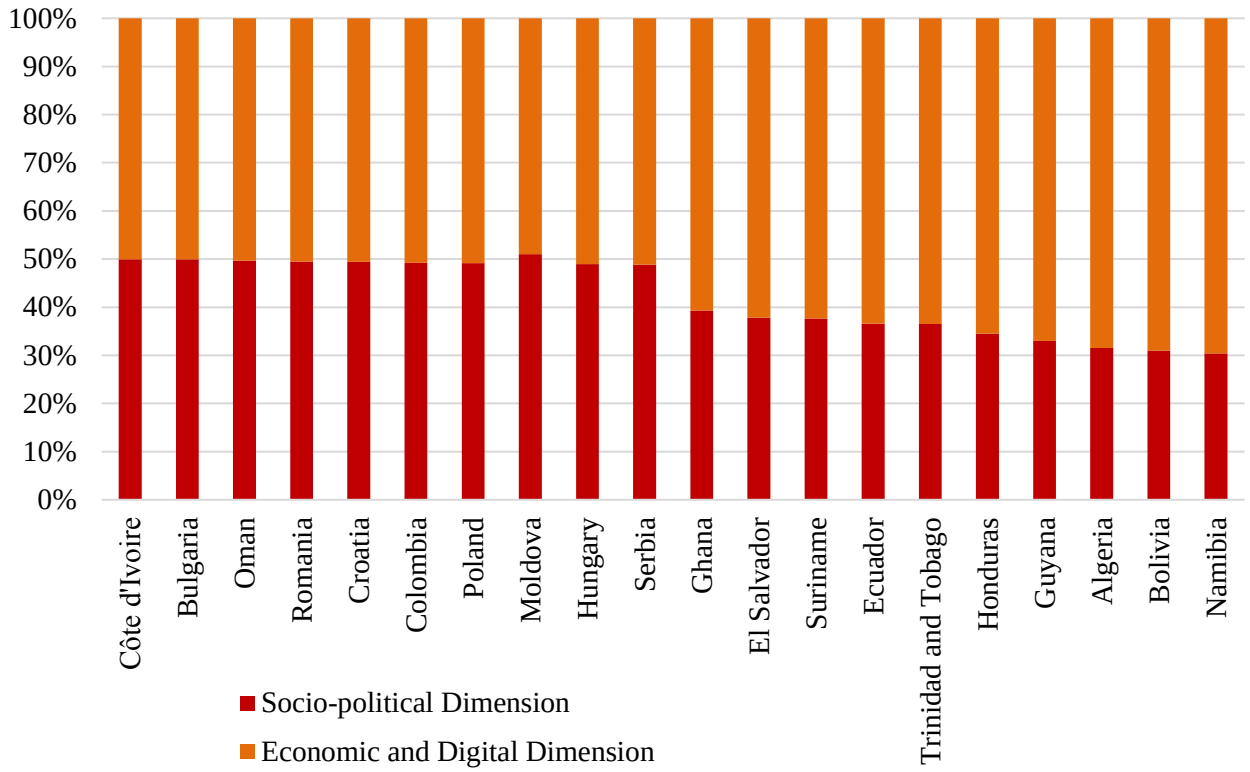


Рисунок 1.44 – Збалансованість пар композитних таргетів економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки для країн, що розвиваються

Розраховані значення сум протилежних кутів для нових індустріальних країн представлено на рисунку 1.45, де можна побачити, що тільки модель Філіппін та Індії мають значення сум протилежних кутів чотирикутника, які приблизно дорівнюють 180 градусів. Інші країни мають незбалансовані пари вимірів, причому для одних є характерним превалювання збалансованості економіко-цифрового виміру (Індія, Індонезія, Мексика, Єгипет, В'єтнам, Малайзія, Пакистан, Іран, Китай, Тайланд, Туреччина, Нігерія), для інших – соціально-політичного (Південна Африка, Аргентина, Бангладеш, Бразилія, Чилі). Аналіз окремих індикаторів показав, що такі країни, наприклад, як Китай,

Індія, Єгипет та інші, мають потужний розвиток ІТ-галузі та кібербезпеки. Такі країни, як Мексика та Малайзія мають розвиток економічної та сери ІТ приблизно на одному рівні. Для Бразилії, Чилі та Аргентини дестабілізуючим таргетом є політичний, що є наслідком політичної нестабільності цих країн. Тобто, група нових індустріальних країн має різні напрямки розвитку країн, що потрібно враховувати їх урядом для розробки стратегії розвитку.

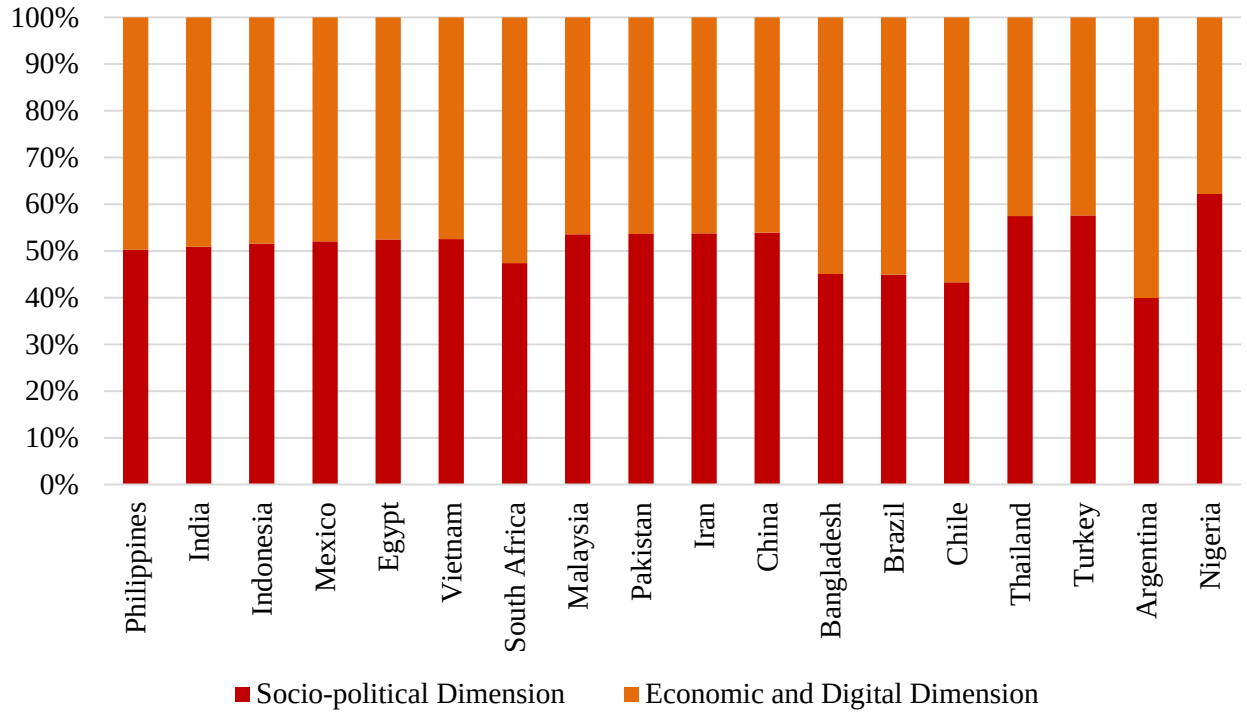


Рисунок 1.45 – Збалансованість пар композитних таргетів економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки для нових індустріальних країн

Значення сум протилежних кутів для найменш розвинених країн представлено на рисунку 1.46. Тільки для Камбоджі значення сум протилежних кутів чотирикутника дорівнюють 180 градусів, а для Ефіопії ці значення є близькими. Інші країни мають ярко виражену незбалансованість пар вимірів, причому для переважної більшості із них є характерним превалювання соціально-політичного виміру та незбалансованість економіко-цифрового. Для Чаду, Малі та Бурунді ситуація є зворотною. Незважаючи на те, що значення їх таргетів є низькими, можна спостерігати, що деякі країни мають урівноважений

розвиток таргетів для пар вимірів. Наприклад, Бутан (0.5400 – соціальний таргет, 0.6322 – політичний, 0.3800 – економічний, 0.3356 – цифрова спроможність і кібербезпека), який має соціально-політичний та економіко- цифровий розвиток на приблизно однаковому рівні, але між даними парами є суттєва різниця, що свідчить про недостатній потенціал економіко-цифрової сфери. Інші країни цієї групи можуть мати інші сценарії розвитку, де домінуватиме тільки один з таргетів, що пов’язано з їх історичними, культурними, політичними та іншими особливостями існування та розвитку.

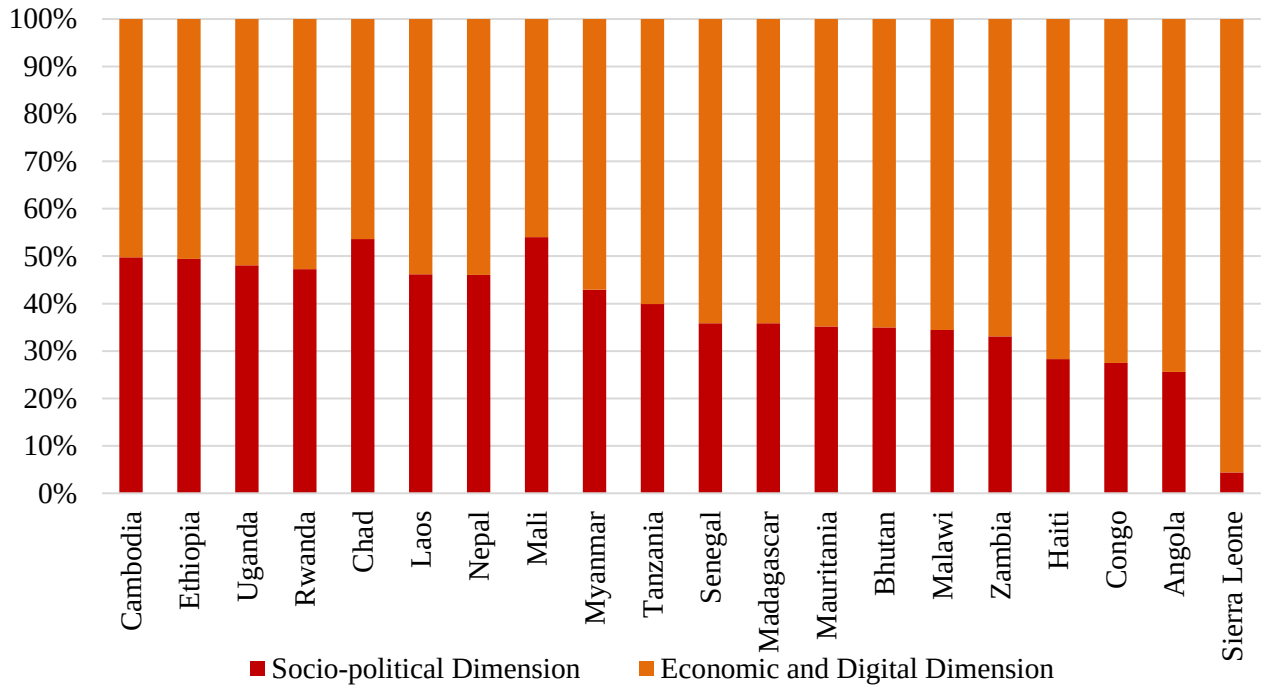


Рисунок 1.46 – Збалансованість пар композитних таргетів економічного, соціального, політичного вимірів та виміру цифрової спроможності і кібербезпеки для топ-двадцяти найменш розвинених країн

Розраховані значення відстаней між центрами мас для всіх країн, які представляють собою відхилення фактичних значень їх центрів мас від еталонного, представлені на рисунку 1.47.

Так, найбільш збалансованими є Нова Зеландія (0.0106), Малі (0.0196), Бурунді (0.0214), Швейцарія (0.0236), Швеція (0.0272), Сінгапур (0.0312),

Маврикій (0.0340), Канада (0.0350), Мавританія (0.0357). Тобто, найбільш збалансованим є країни, як розвинуті, так й ті, що розвиваються, та найменш розвинені. Даний фактор свідчить про те, що не залежно від значень таргетів, рівня збалансованості їх пар, будь-яка країна не залежно від рівня її економічного розвитку може мати ефективне поєднання чотирьох таргетів.

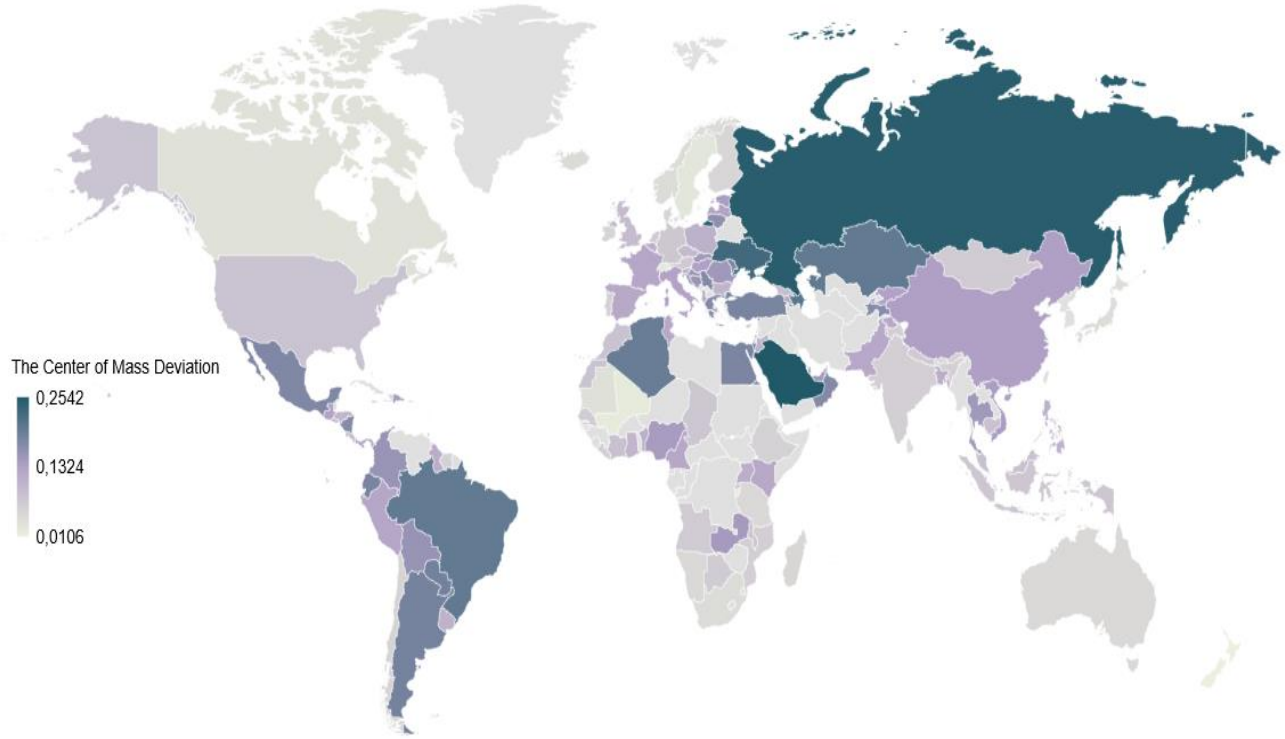


Рисунок 1.47 – Рівень збалансованості розвитку країн на основі відхилення центрів мас їх барицентричних моделей

Наприклад, Малі має низькі значення таргетів, але їх комбінація є збалансованою, що в подальшому може виступати драйвером для їх більш стрімкому та динамічному розвитку. Найменш збалансованими виявилися Парагвай (0,1860), Алжир (0,1946), Бразилія (0,1990), Казахстан (0.1998), Азербайджан (0,2063), Бахрейн (0,2093), Іран (0,2113), Україна (0,2269), Російська Федерація (0,2467), Саудівська Аравія (0,2542). Цей результат говорить про те, що дані країни мають дисбаланс за рахунок превалювання переважно одного (наприклад, як в Україні таргет цифрової спроможності і кібербезпеки) або двох таргетів над іншими, що свідчить про несистемність їх

розвитку та необхідність трансформації їх стратегій з урахуванням отриманих даних.

Побудуємо чотириполюсні барицентричні моделі збалансованості розвитку країн з кожної чотирьох груп, які мають найменшу відстань розрахованого центру їх мас від еталонного значення. Для розвинених країн таким представником є Нова Зеландія (0.0106), для країн, що розвиваються – Маврикій (0.0340), нових індустріальних – Південна Африканська Республіка (0.0428), для найменш розвинених – Малі (0.0196). На рисунку 1.48 представлена барицентрична модель Нової Зеландії.

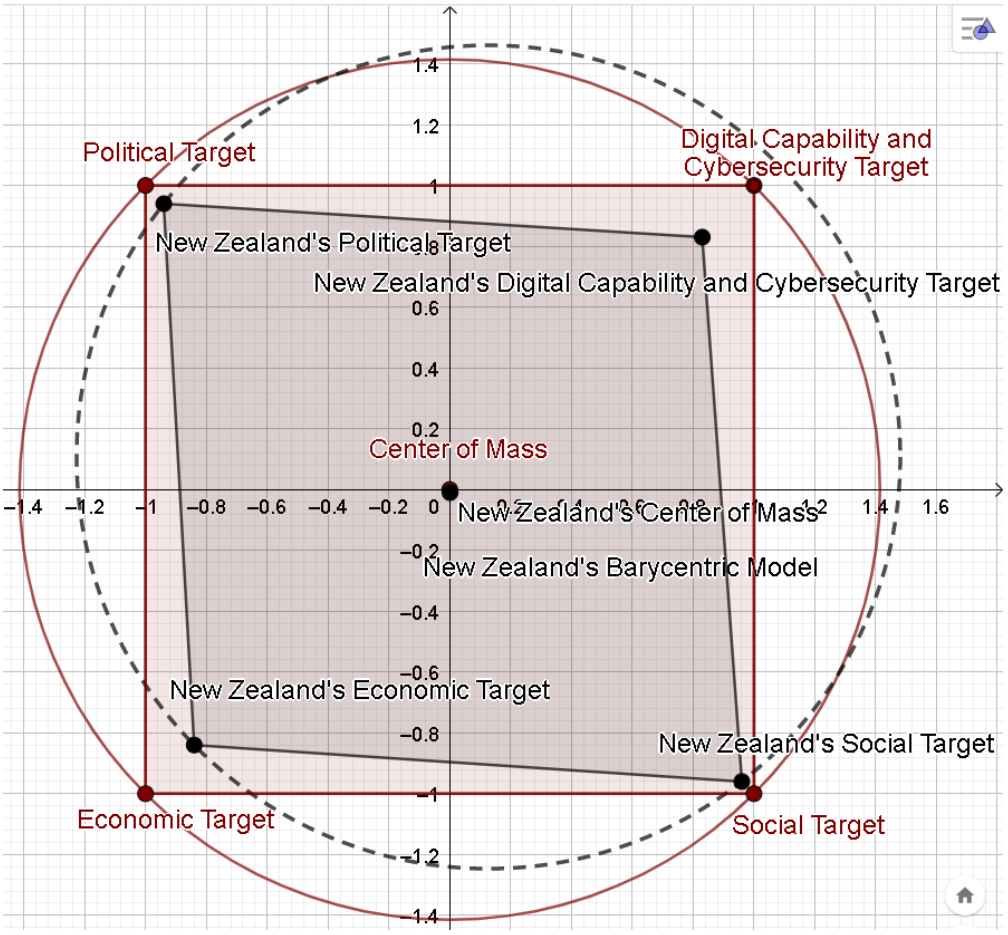


Рисунок 1.48 – Чотириполюсна барицентрична модель збалансованості розвитку Нової Зеландії

Барицентрична модель (рисунок 1.48) показує розвиток Нової Зеландії з урахуванням збалансованості чотирьох композитних таргетів – економічного, соціального, політичного та цифрової спроможності і кібербезпеки. Їх значення

є досить високими та наближаються до 1, що свідчить про високий рівень розвитку економіки, соціальних стандартів, політичної стабільності, а також значний потенціал ІТ сфери та кібербезпеки. На рисунку 1.48 відмічено центр мас чотирикутника, координати якого практично дорівнюють координатам еталонного центру мас, що свідчить про повну збалансованість чотирьох таргетів. Але коло описати навколо даного чотирикутника не можливо, оскільки суми пар протилежних кутів не дорівнюють 180 градусів. Це відбувається за рахунок того, що вимір цифрової спроможності і кібербезпеки, а також економічний вимір мають значення значно нижче ніж соціальний та політичний. За даною моделлю можна зробити наступний висновок: розвиток країни є стійким, оскільки відстань між центрами мас є незначною. Співвідношення між парами вимірів (економіко-цифровим та соціо-політичним) є незбалансованим, але розвиток однієї сфери можна компенсувати за рахунок іншої. Значення композитних таргетів економічного та цифрового вимірів є слабкішими, тому країні треба змістити акцент у даний напрямок розвитку, особливо в частині цифровізації та автоматизації різних сфер діяльності економічних агентів. При цьому драйвером розвитку можуть виступати політичний та соціальний виміри.

Чотириполюсна барицентрична модель збалансованості розвитку Маврикія представлена на рисунку 1.49. Отримана модель дозволяє зробити наступні висновки: розвиток країни є досить стійким, оскільки відстань між центрами мас дорівнює 0,0340, що практично наближає його до мінімального значення серед відстаней. Співвідношення між парами вимірів (економіко-цифровим та соціо-політичним) не є збалансованим, оскільки суми протилежних кутів не дорівнюють 180 градусів, причому дисбаланс є найбільшим для економіко-цифрового вимірів. Найменш ефективним в цій пар є таргет цифрової спроможності і кібербезпеки, що свідчить про відставання рівня розвитку інформаційних технологій та заходів кібербезпеки від інших. Це пояснюється тим, що Маврикій є острівною державою, економіка якого орієнтується на розвиток туристичної галузі. Для даного таргету є важливим підвищення рівня національної кібербезпеки, оскільки в порівнянні з іншими даними його

значення є низьким, що свідчить про можливі проблеми в системі державного кіберзахисту. Значення таргетів є вище середнього, серед яких найбільш ефективним є таргет соціального розвитку, що може бути відповідним драйвером для розвитку економіки та її цифровізації.

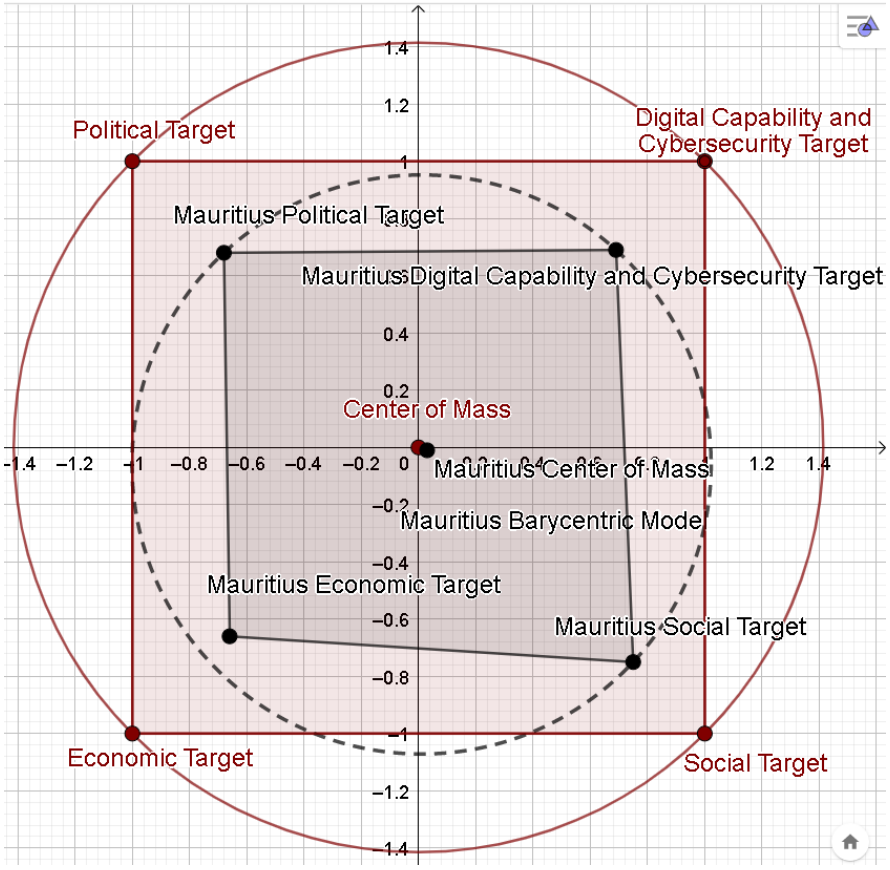


Рисунок 1.49 – Чотириполюсна барицентрична модель збалансованості розвитку Маврикія

Побудуємо чотириполюсну барицентричну модель збалансованості розвитку однієї із нових індустріальних країн, а саме Південно Африканської Республіки, результат якої представлений на рисунку 1.50. За результатами моделі (рисунок 1.50) можна зробити наступні висновки: розвиток країни є стійким, оскільки відстань між центрами мас наближається до 0 і дорівнює 0,0428. Співвідношення між парами вимірів (економіко-цифровим та соціально-політичним) – незбалансоване, оскільки суми протилежних кутів не дорівнюють 180 градусів, причому дисбаланс є найбільшим для економіко-цифрового вимірів, ніж для соціально-політичного. Значення таргетів коливаються біля

середнього рівня, але найбільш неефективним є таргет економічної сфери, що стримує розвиток країни та унеможливлює розвиватися комплексно. Чотириполюсна барицентрична модель збалансованості розвитку Малі представлена на рисунку 1.51.

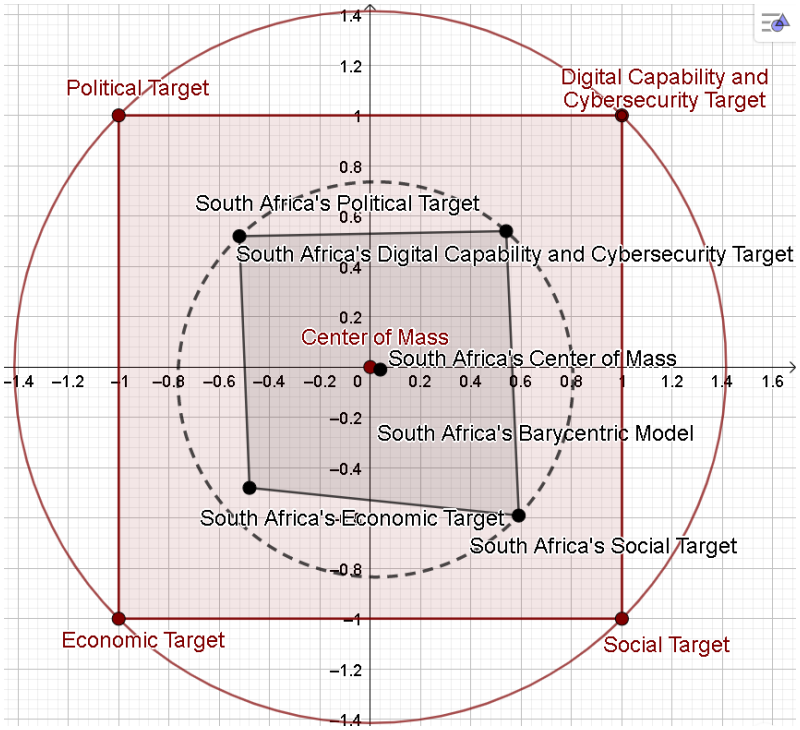


Рисунок 1.50 – Чотириполюсна барицентрична модель збалансованості розвитку Південно Африканської Республіки

За результатами моделі (рисунок 1.51) можна зробити наступні висновки: розвиток країни є збалансованим, оскільки відстань між центрами мас дорівнює 0.0196, але оскільки значення таргетів є досить низькими (0,2323 – для виміру цифрової спроможності та кібербезпеки; 0,1940 – для економічного; 0,1954 – для соціального; 0,1791 – для політичного) та наближаються до 0, то це говорить про досить слабкі темпи розвитку всіх сфер Малі. Співвідношення між парами вимірів незбалансоване, оскільки суми протилежних кутів не дорівнюють 180 градусів. За умови ефективних політичних рішень, фінансової допомоги міжнародних організації, трансформації стратегій умови збалансованості можуть сприяти більш динамічному подальшому розвитку країни.

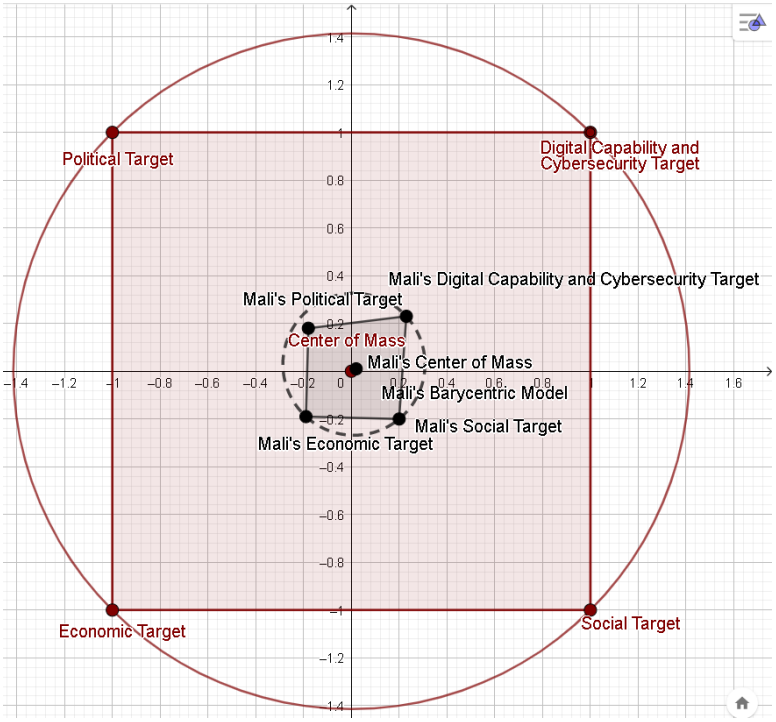


Рисунок 1.51 – Чотириполюсна барицентрична модель збалансованості розвитку Малі

В умовах динамічних змін, що відбуваються в різних сферах життєдіяльності суспільства, є важливим визначення ключових детермінант, які забезпечують збалансований розвиток економічної, політичної, соціальної сфер, а також інформаційних технологій та кібербезпеки. Відповідно слід розуміти, що викликає дисбаланс та на скільки це може бути критичним. З цією метою автори даного дослідження використали підхід визначення центру мас геометричної фігури для моделювання рівня збалансованості розвитку країн, що відбувається на основі факторів економічного, соціального, політичного розвитку та цифрової спроможності і кібербезпеки. Модель представляє собою чотириполюсну барицентричну модель, сформовану на основі композитних таргетів, що формуються під впливом окреслених детермінант. Перелік факторів та таргети було обрано на основі проведеного аналізу літератури та методів наукового пізнання. Розрахунки проводилися для даних 127 країн світу, обраних за 2018 рік. Побудова барицентричної моделі відбувалася з урахуванням її трьох компонентів: значень композитних таргетів, рівня збалансованості пар таргетів

та збалансованості всіх чотирьох таргетів, тобто визначення центру мас чотирикутника. В процесі аналізу отриманих значень було виявлено, що розвинуті країни мають інтегральні значення чотирьох композитних таргетів вищі, ніж для груп країн, що розвиваються, нових індустріальних та найменш розвинених. Це свідчить про високий рівень добробуту цих країн, соціального захисту їх населення, політичної стабільності, розвитку ІТ-сфери. Розраховані значення сум протилежних пар кутів для більшості країн не дорівнюють значення 180 градусів, що засвідчило про незбалансований рівень їх розвитку. Для розвинутих країн найбільш ефективною є пара соціально-політичного розвитку, що є наслідком високих темпів розвитку економіки. Але ця пара вимірів може слугувати також й драйвером для прискорення розвитку пари економічного таргету та таргету цифрової спроможності і кібербезпеки. Для країн, що розвиваються, та нових індустріальних незбалансованість може бути викликана різними детермінантами. Так, для більшості з них таким таргетом виступає цифрова спроможність та кібербезпека. Сюди відносяться Кот-д'Івуар, Хорватія, Грузія, Кенія, Молдова, Катар, Російська Федерація, Саудівська Аравія, Сербія, Україна, Єгипет, Індія, Малайзія, Пакистан, Туреччина. Цей факт може сприяти розвитку четвертинного сектору економіки цих країн та призвести до поступової її трансформації у цифрову площину. Для більшості найменш розвинених країн незбалансованою є пара економіко-цифрового таргетів, в якій саме економічний є критичним для подальшого розвитку країни в цілому. Отримані висновки аналізу відстані розрахованого центру мас від еталонного значення показали, що є країни, для яких характерний збалансований розвиток на основі всіх чотирьох композитних таргетів. При цьому виявилося, що збалансованими можуть бути не тільки розвинені країни, але й ті, що розвиваються, та найменш розвинуті, такі як Малі й Бурунді. Але не дивлячись на цей факт, рівень розвитку їх таргетів відповідає їх класифікаційній групі, що дозволило зробити висновок про можливості подальшого динамічного розвитку таких країн за умови підтримки ефективності соціальної, політичної, економічної сфер, а також сфери цифрової спроможності і кібербезпеки. Тим

країнам, розвиток яких є найбільш незбалансованим, наприклад, Саудівська Аравія, Україна, побудова барицентричної моделі дозволяє виявити той напрям або напрями, які призводять до цього. Це може бути викликано політичними коливаннями, військовими конфліктами, низькою якістю соціальної сфери, тощо. Результати даного дослідження слід прийняти до уваги відповідним державним органам, що відповідають за певну сферу розвитку країни, для розробки більш ефективних стратегій.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден., Ошибка! Источник ссылки не найден.**].

1.3.2 Сценарії взаємодії систем кібербезпеки та протидії фінансовим злочинам для країн з різним рівнем економічного розвитку

На сьогодні проблема боротьби із відмиванням кримінальних доходів та фінансування тероризму є вкрай актуальною для країн світу. Це пов’язано із тим, що за рахунок процесу легалізації коштів, джерела походження яких мають незаконний характер, значні грошові суми уникають оподаткування, сприяють розвитку тіньового сектору, стимулюють підвищення рівня злочинності та, врешті-решт, можуть вплинути на дестабілізацію економіки країни, створення конфліктів у суспільстві, зниження довіри до країни з боку міжнародних партнерів. За результатами опитування, проведеного консалтинговою компанією “PwC” за 2018 рік, обсяг операцій з відмивання кримінальних доходів та фінансування тероризму становив 1 трлн. дол., що склало приблизно від 2% до 5% світового ВВП [**Ошибка! Источник ссылки не найден.**]. Саме тому світова спільнота схвильована існуванням даної проблеми, оскільки з’являються загрози міжнародній фінансовій системі. Профільна міжнародна організація FATF пропонує необхідні заходи щодо здійснення боротьби та протидії легалізації (відмивання) доходів, одержаних злочинним шляхом, фінансуванню тероризму та фінансуванню розповсюдження зброї масового знищення, в рамках яких

розроблено спеціальні стандарти та інструменти, які періодично оновлюються у відповідності до реалій функціонування фінансової системи.

З іншого боку, наслідки промислової революції 4.0 викликали стрімкий розвиток інформаційних технологій та впровадження їх в усі сфери життєдіяльності людини. Процеси автоматизації та діджиталізації призвели до зростання рівня кіберзлочинів, особливо у фінансовій сфері, яка входить у п’ятірку найбільш атакованих сфер світу [**Ошибка! Источник ссылки не найден.**]. Також рівень збитків від кіберзлочинності зростає у геометричній прогресії та за прогнозованими оцінками експертів він дорівнюватиме за 2021 рік 6 трлн. дол. [**Ошибка! Источник ссылки не найден.**]. Тому проблема забезпечення відповідного рівня кіберзахисту фінансової системи країни та інших її систем є критично важливою та практично значущою.

Вирішення окреслених проблем є можливим за рахунок конвергенції системи кібербезпеки та протидії легалізації кримінальних доходів, оскільки синергетичний ефект від їх взаємодії буде значно більшим ніж від їх окремого функціонування. Це можливо за рахунок їх системного поєднання на технологічному, програмному, інформаційному, правовому та організаційному рівнях. Процес інтеграції є досить складним і потребує застосування зважених рішень, оскільки наслідки від неправильних заходів можуть бути катастрофічними. Тому попередньо необхідно здійснити оцінку фактичного стану системи кібербезпеки та протидії фінансовим шахрайствам для визначення потенційного рівня їх конвергенції для різних країн.

Питання конвергенції системи кібербезпеки та протидії легалізації кримінальних доходів є досить новим для сучасної наукової спільноти. Тому можна виділити тільки ряд тих наукових досліджень, які проводилися у близькому до даного питання напрямку.

Найбільш актуальним серед практиків та науковців є напрям дослідження проблеми протидії шахрайству із кредитними картками. Це відбувається завдяки зростання обсягу шахрайств по відношенню до фізичних осіб – клієнтів банків за рахунок появи низки методів, таких як соціальна інженерія. Дану тематику

досліджували Діліп М.Р., Наванет А.В., Абхішек М. [**Ошибка! Источник ссылки не найден.**], Ванг Р., Лью Дж. [**Ошибка! Источник ссылки не найден.**], Мішра С.П., Кумарі П. [**Ошибка! Источник ссылки не найден.**], Мектерович І., Каран М., Пінтар Д., Брккіч Л. [**Ошибка! Источник ссылки не найден.**], та інші.

Також вивчаються інструменти протидії фінансовим та кібершахрайствам. Особливо популярними є засоби машинного навчання та штучного інтелекту, які використовуються в процесі виявлення операцій, що носять ознаки шахрайських. Так, Чен З., Ван Хоа Л.Д., Тео Е.Н. , Назір А., Каруппія Е.К., Лам К.С. дослідили можливості застосування засобів машинного навчання для виявлення операцій з легалізації кримінальних доходів [**Ошибка! Источник ссылки не найден.**]. Чжоу Ю., Сонг Кс., Чжоу М. запропонували метод бустінгу для прогнозування шахрайських операцій [**Ошибка! Источник ссылки не найден.**]. Мультиагентна система для виявлення операцій з відмивання коштів, отриманих злочинним шляхом, яку можна інтегрувати в банківську інформаційну систему, була розроблена Гао С., Сю Д., Ванг Х., Грін П. [**Ошибка! Источник ссылки не найден.**]. Карпуніна Є.К., Михайлов А.М., Бондарева Н.А., Любименко О.А., Федотова Є.В. досліджували можливості застосування блокчейн-технологій для протидії фінансовим та кібершахрайствам [**Ошибка! Источник ссылки не найден.**].

Важливими є питання організаційного, технологічного, правового та інформаційного забезпечення системи кібербезпеки та протидії легалізації кримінальних доходів. Так, в напрямку інтеграції політичної, освітньої та технологічної сфери для забезпечення ефективності функціонування системи кібербезпеки та протидії фінансовим шахрайствам проведено дослідження М. Доусоном [**Ошибка! Источник ссылки не найден.**]. Діонісій С. Деметіс розглядав технології виявлення операцій з легалізації незаконних коштів, серед яких виділяв ризикологію та методи оцінювання ризиків [**Ошибка! Источник ссылки не найден.**]. Гальяні Г. досліджував поняття «технологічної нейтральності» по відношенню до кібербезпеки у контексті формування та

забезпечення міжнародного правового поля з даного питання [**Ошибка! Источник ссылки не найден.**].

Не зважаючи на широке коло наукових публікацій, які охоплюють напрямок дослідження проблеми боротьби і протидії фінансовим та кібершахрайствам, досить багато питань є мало вивченими і потребують уточнення, удосконалення та подальшого дослідження. Особливо це стосується можливості конвергенції системи кібербезпеки та протидії фінансовим шахрайствам й легалізації кримінальним доходам.

Метою дослідження є здійснення оцінки рівня потенційної конвергенції системи кібербезпеки та протидії легалізації кримінальних доходів і фінансуванню тероризму на основі визначення їх інтегральних показників та застосування функції Харрінгтона – Менчера, що буде виконано в рамках формування сценаріїв конвергенції.

Для здійснення оцінки рівня конвергенції скористаємося підходом, запропонованим Яровенко Г.М. у роботі [**Ошибка! Источник ссылки не найден., Ошибка! Источник ссылки не найден.**] для оцінювання рівня загрози інформаційної безпеки, суть якого полягає у визначенні інтегрального показника. Але для нашого дослідження необхідно розрахувати два композитних індикатори, один з яких характеризуватиме рівень кібербезпеки в країні, а інший – рівень протидії легалізації кримінальних доходів.

На першому етапі оберемо вхідні дані, які будуть використовуватися для здійснення розрахунків. Першу групу сформували світові індекси, що застосовуються для вимірювання окремих сфер кібербезпеки країни, узяті з офіційного сайту організації «e-Governance Academy Foundation» за 2018 рік: Глобальний індекс кібербезпеки (Global Cybersecurity Index) оцінює можливості країн світу протидіяти кіберзагрозам у світі, а також визначає їх слабкі сторони та потенційні можливості; Національний індекс кібербезпеки (National Cyber Security Index) визначає стан готовності окремої країни протидіяти кіберзагрозам та керувати кіберінцидентами; Індекс мережевої готовності (Networked Readiness Index) дозволяє оцінити рівень технологічної готовності країни для

впровадження сучасних інформаційних систем та технологій для автоматизації різних процесів життєдіяльності суспільства; Рівень цифрового розвитку (Digital Development Level) показує ступінь цифровізації країни. Кожен з цих обраних показників характеризує стан кібербезпеки країни з огляду на різні її аспекти, тому їх аналіз у сукупності дозволить сформувати комплексне бачення на її розвиток та можливості інтеграції.

Другу групу індикаторів сформували індекси, які дозволяють оцінити стан системи протидії легалізації кримінальних доходів та фінансування тероризму. Сюди увійшли: Індекс політичної стабільності (Political Stability Index), який дозволяє оцінити ймовірність дестабілізації уряду країни із використанням неконституційних та насильницьких заходів, що є сприятливим або несприятливим в залежності від значення фактором для процесів легалізації незаконних коштів; Індекс ефективності уряду (Government Effectiveness Index), який вимірює його якість, що полягає у його незалежності від політичного тиску, ефективності роботи державних служб, рівня довіри до його діяльності; Легкість ведення бізнесу (Ease of Doing Business) характеризує умови для ведення бізнесу в країні, що впливає на ризики зростання тіньового сектору та відмивання коштів; Індекс злочинності (Crime Index) характеризує рівень злочинності в країні, який впливає на нестабільність соціальної, політичної та економічної сфер; Глобальний індекс тероризму (Global Terrorism Index) свідчить про рівень терористичної активності, що впливає на ризики легалізації кримінальних доходів та фінансування тероризму; Індекс фінансової таємниці (Financial Secrece Index) свідчить про ступінь захисту фінансових операцій, що багатьма країнами використовується для формування сприятливих умов для приховування незаконних доходів та здійснення фінансових операцій, джерела коштів яких є кримінальними. Дані обраних показників було узято з офіційного джерела Світового банку. Емпіричні дані обох груп відповідають 76 країнам світу за 2018 рік, оскільки саме цей період характеризується найбільш повним набором значень.

В роботі [**Ошибка! Источник ссылки не найден.**] авторами Кузьменко О.В., Яровенко Г.М., Радько В.В. проведено попередній аналіз процесу конвергенції систем кібербезпеки та фінансового моніторингу, що дозволило довести релевантність саме цих показників для подальшого дослідження.

На другому етапі проведемо нормалізацію вхідних даних для їх приведення до співставного вигляду. Для цього використаємо нелінійну нормалізацію, яка згладжує різні за знаками та значеннями дані більш ефективно, ніж інші методи (формула (1.36)):

$$Z_{ij} = \left(1 + e^{\frac{\overline{y_j} - y_{ij}}{\sigma(y)}}\right)^{-1}, \tag{1.36}$$

де Z_{ij}– нормалізоване значення j-го показника, обраного для здійснення оцінки рівня конвергенції системи кібербезпеки та протидії легалізації кримінальних доходів, в розрізі і-ої країни;

$\overline{y_j}$– середнє значення j-го показника в межах досліджуваного переліку країн;

y_{ij} – фактичне значення j-го показника в розрізі і-ої країни;

$\sigma(y_j)$ – середнє квадратичне відхилення j-го показника в межах досліджуваного переліку країн.

Всі обрані показники за своїм впливом на стан системи є стимуляторами, окрім двох – індексу злочинності та фінансової таємниці, які є дестимуляторами. Тому для того, щоб правильно врахувати їх значення при формуванні інтегрального індексу, необхідно їх розраховане нормалізоване значення відняти від одиниці.

На третьому етапі проведемо трансформацію нормалізованих значень обраних показників бази дослідження до безрозмірної шкали бажаності Харрінгтона за допомогою формули (1.37):

$$d_{ij} = \exp(-\exp(-Z_{ij})), \quad (1.37)$$

де d_{ij} - проміжне значення j-го показника, обраного для здійснення оцінки рівня конвергенції системи кібербезпеки та протидії легалізації кримінальних доходів, в розрізі і-ої країни, приведене до безрозмірної шкали бажаності Харрінгтона;

Z_{ij} – нормалізоване значення j-го показника, в розрізі і-ої країни.

Для подальшої побудови інтегрального показника оцінки рівня конвергенції системи кібербезпеки та протидії фінансовим шахрайствам необхідно дослідити характер поведінки кривої перетворення Харрінгтона-Менчера, яка характеризує залежність d_{ij} від фактичних значень кожного вхідного показника. З цією метою проведемо візуалізацію залежностей на четвертому етапі. В результаті було виявлено, що для більшості показників є характерним перший тип кривої – S-подібна, зростаюча, симетрична. Індексу злочинності та фінансової таємниці відповідає четвертий тип – S-подібна, спадаюча, симетрична крива. Приклади отриманих графіків кривої першого та другого типів представлені на рисунках 1.52 та 1.53.

Рисунок 1.52 – Графік кривої першого типу для «Національного індексу кібербезпеки»

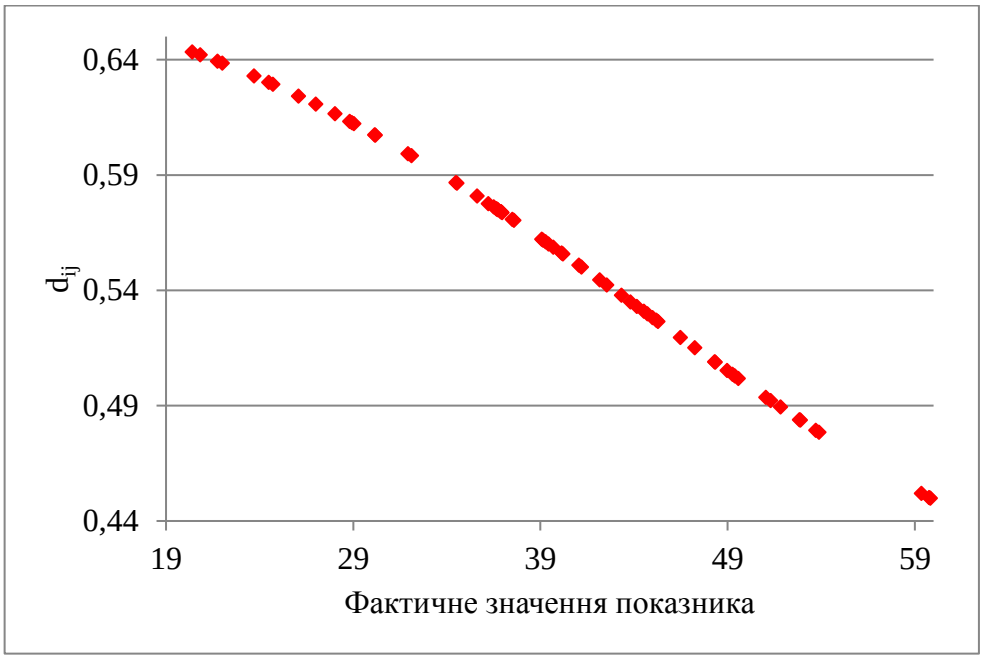


Рисунок 1.53 – Графік кривої четвертого типу для «Індексу злочинності»

На п’ятому етапі проведемо формалізацію перетворення Харрінгтона-Менчера в межах обраної на попередньому кроці залежності d_{ij} від фактичних значень в розрізі кожного вхідного показника. Тобто розрахуємо проміжні значення показників для оцінки рівня конвергенції системи кібербезпеки та протидії фінансовим шахрайствам з урахуванням їх приведення до безрозмірної шкали бажаності Харрінгтона-Менчера у відповідності із визначеним типом кривої.

Для показників, залежності для яких описуються кривою першого типу, використаємо формулу (1.38):

$$d_{ij}^{*} = \exp\left(-\exp\left(-\left(9\left(\frac{z_{ij}-\min\limits_{i} z_{ij}}{\max\limits_{i} z_{ij}-\min\limits_{i} z_{ij}}\right)^{1.927}-2\right)\right)\right), \qquad (1.38)$$

де d_{ij}^{*} - проміжне значення j-го показника, обраного для здійснення оцінки рівня конвергенції системи кібербезпеки та протидії легалізації кримінальних доходів, в розрізі і-ої країни, приведене до безрозмірної шкали бажаності Харрінгтона-Менчера;

$\min_i Z_{ij}$ – мінімальне значення нормалізованого j-го показника в розрізі і-ої країни;

$\max_i Z_{ij}$ – максимальне значення нормалізованого j-го показника в розрізі і-ої країни.

Для показників, залежності для яких описуються кривою четвертого типу, використаємо формулу (1.39):

$$d_{ij}^{*} = \exp\left(-\exp\left(-\left(9\left(\frac{\max_i Z_{ij} - Z_{ij}}{\max_i Z_{ij} - \min_i Z_{ij}}\right)^{1.927} - 2\right)\right)\right). \qquad (1.39)$$

На шостому етапі необхідно визначити ваги показників для того, щоб розрахувати узагальнену функцію. З цією метою проведемо канонічний аналіз, який дозволить визначити ступінь залежності між двома множинами показників, а також розрахувати їх канонічні ваги, які буде використано для інтегральної оцінки. Аналіз виконано із використанням модуля канонічного аналізу аналітичного пакету “STATISTICA”, результати якого представлені на рисунку 1.54.

N=76	Canonical Analysis Summary (Konvergentcia2.sta) Canonical R: .93762 Chi²(24)=200.41 p=0.0000 **Left Set**	Right Set
No. of variables	4	6
Variance extracted	100.000%	83.8201%
Total redundancy	70.3694%	47.9580%
Variables: 1	Global Cybersecurity Index	Political stability index
2	Networked Readiness Index	Government effectiveness index
3	National Cyber Security Index	Ease of doing business
4	Digital Development Level	Crime Index
5		Global Terrorism Index
6		Financial Secrece Index

Рисунок 1.54 – Підсумки канонічного аналізу

З рисунку 1.54 можна побачити, що значення канонічної кореляції R = 0,93762, що свідчить про наявність дуже сильного кореляційного зв’язку між множиною факторів, які характеризують рівень розвитку системи кібербезпеки та протидії фінансовим шахрайствам.

Статистичну значимість коефіцієнта кореляції підтверджує високе значення критерію Пірсона ($\chi^2 = 200{,}00$), рівень значущості якого не перевищує 0,05 (p = 0,0000). Значення надмірності для лівої множини, яку сформували індекси кібербезпеки, дорівнює 70,3694%. Це свідчить про те, що фактори правої множини, які відповідають показникам рівня протидії фінансовим шахрайствам країни, на 70,3694% пояснюють мінливість індикаторів кібербезпеки, що свідчить про високе значення впливу. Розвиток системи протидії процесам відмивання коштів в країні в певній мірі залежить від стану її кібербезпеки, оскільки фактори кібербезпеки на 47,9580% пояснюють мінливість факторів, які характеризують рівень протидії фінансовим шахрайствам. Хоча отримане значення є помірним, але воно є достатнім для обґрунтування впливу таких показників, як кібербезпека, на економічні процеси в країні.

Визначені значення канонічних коренів, а також отримані статистичні характеристики, дозволили зробити висновок, що значущими є 3 канонічні корені. Але для того, щоб одержати достовірні оцінки їх навантажень для трьох

пар канонічних змінних, необхідно мати вибірку, яка буде перевищувати в 40-60 раз кількість початкових даних [**Ошибка! Источник ссылки не найден.**, с. 190]. Тому прийнято рішення, що для визначення вагів доцільно використати значення тільки першого канонічного кореня, для якого канонічний R^2 буде мати найбільше значення 0,8791. Виходячи з даних міркувань для подальшого розгляду використаємо канонічні ваги, визначені для першого кореня (рисунки 1.55-1.56).

Variable	Canonical Weights, left set (Konvergentcia2.sta)			
	Root 1	Root 2	Root 3	**Root 4**
Global Cybersecurity Index	0,313261	-0,781709	0,63400	1,14199
Networked Readiness Index	0,264381	-0,713150	-1,56282	-0,64848
National Cyber Security Index	-0,021339	0,026080	0,91519	-1,29626
Digital Development Level	0,557799	1,355225	0,21392	0,67528

Рисунок 1.55 – Канонічні ваги для показників кібербезпеки

Variable	Canonical Weights, right set (Konvergentcia2.sta)			
	Root 1	Root 2	Root 3	Root 4
Political stability index	-0,269140	0,923250	1,32088	0,990101
Government effectiveness index	0,780788	0,200480	-1,68893	0,203985
Ease of doing business	0,341713	-0,672184	0,64477	-0,816572
Crime Index	0,050110	0,111717	0,73583	-0,231753
Global Terrorism Index	0,009265	-0,080100	1,17396	1,219424
Financial Secrece Index	-0,091481	0,070369	0,35190	-0,048788

Рисунок 1.56 – Канонічні ваги для показників, що характеризують рівень протидії легалізації кримінальних доходів

Виявилося, що отримані канонічні ваги є як додатними, так і від'ємними, що свідчить про позитивний та негативний внесок показників у значення кореня. Але для визначення узагальненої функції необхідно, щоб їх значення варіювалися від 0 до 1, тому відповідні від'ємні ваги будуть узяті по їх модулю.

На сьомому етапі обчислюються два інтегральні індекси для оцінювання рівня розвитку системи кібербезпеки та протидії легалізації кримінальних доходів. Для цього необхідно використати формули (1.40)-(1.41):

$$IC_i = \sqrt[\Sigma_{j=1}^{n} a_j]{\prod_{j=1}^{n} \left(d_{ij}^*\right)^{a_j}}, \quad (1.40)$$

$$IP_i = \sqrt[\Sigma_{j=1}^{m} a_j]{\prod_{j=1}^{m} \left(d_{ij}^*\right)^{a_j}}, \quad (1.41)$$

де IC_i – інтегральний індекс, що характеризує рівень розвитку системи кібербезпеки для і-тої країни;

IP_i – інтегральний індекс, що характеризує рівень розвитку системи протидії легалізації кримінальних доходів для і-тої країни;

n – кількість показників кібербезпеки країни (n = 4);

m – кількість показників, що характеризують рівень розвитку системи протидії легалізації кримінальних доходів (m = 6);

a_j – ваги відповідного j-го вхідного показника кібербезпеки або протидії легалізації кримінальних доходів;

d_{ij}^* - проміжне значення j-го показника кібербезпеки або протидії легалізації кримінальних доходів в розрізі і-ої країни, приведене до безрозмірної шкали бажаності Харрінгтона-Менчера.

Розраховані значення інтегральних показників інтерпретуємо із використанням якісної оцінки, а саме: якщо отримане значення знаходиться в межах 0,80 – 1,00, то стан розвитку країни відповідає оцінці «дуже добре»; від 0,63 до 0,80 – «добре»; від 0,37 до 0,63 – «задовільно»; від 0,20 до 0,37 – «погано»; від 0,00 до 0,20 – «дуже погано».

Візуалізуємо отримані значення із використанням діаграм з картами, які можна побудувати за допомогою програмного продукту MS Excel. Результати представлено на рисунках 1.57 - 1.58.

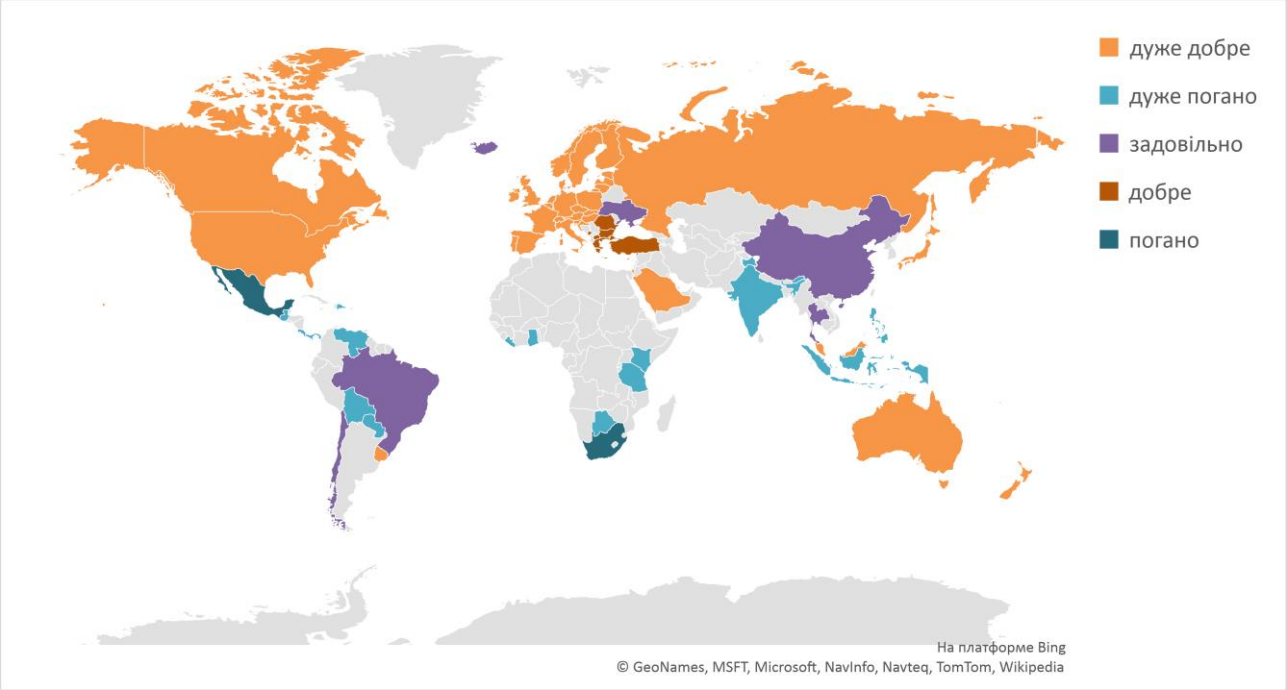


Рисунок 1.57 – Карта розподілу країн за інтегральним індексом, що характеризує рівень розвитку їх системи кібербезпеки

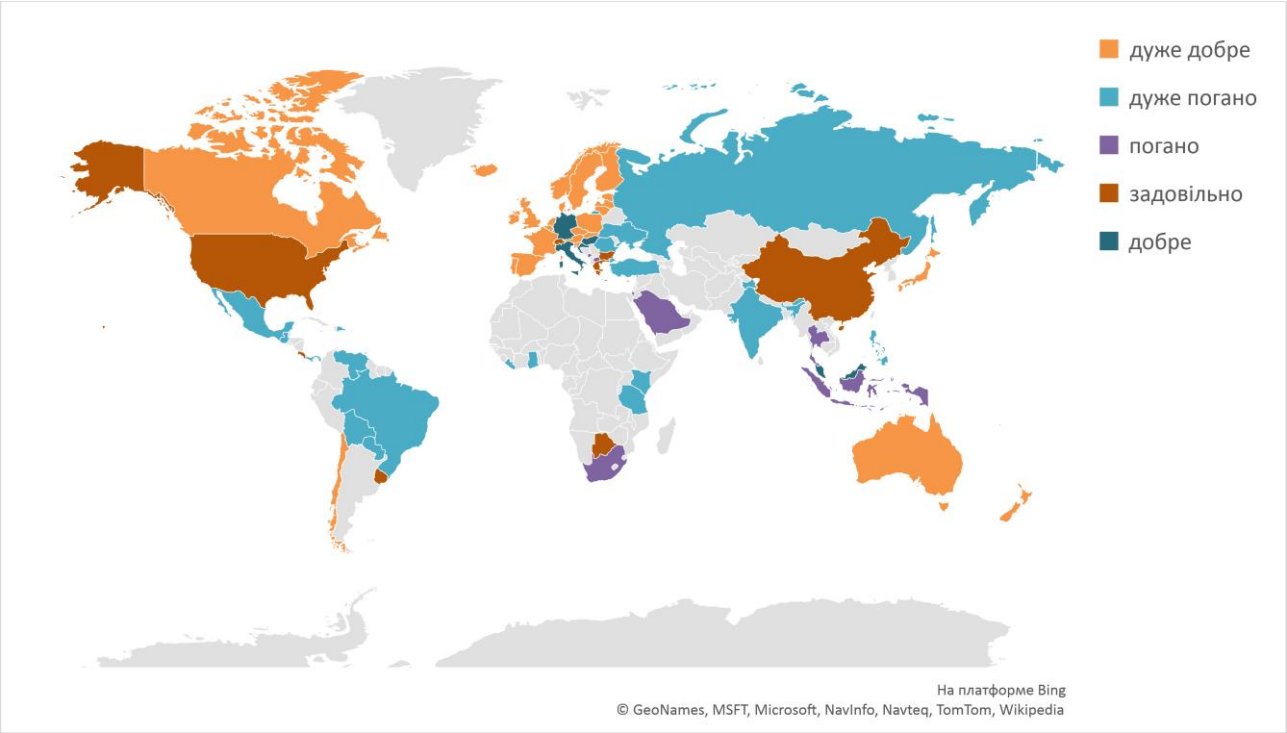


Рисунок 1.58 – Карта розподілу країн за інтегральним індексом, що характеризує рівень розвитку системи протидії легалізації кримінальних доходів

За інтегральним рівнем кібербезпеки виявилося, що оцінку «дуже добре» мають 38 країн, таких як: Австрія, Австралія, Канада, Данія, Естонія, Фінляндія, Німеччина, Великобританія, США та інші (див. рис. 1.57), тобто переважна більшість цих країн є розвиненими. Болгарія, Греція, Маврикій, Чорногорія, Північна Македонія, Туреччина та Румунія мають рівень кібербезпеки, який відповідає оцінці «добре». Задовільний рівень характерний для таких країн, як Україна, Бразилія, Чилі, Китай, Ісландія, Мальта та Тайланд. Оцінку «погано» та «дуже погано» отримали 24 країни: Барбадос, Болівія, Ботсвана, Домініканська республіка, Гана, Гватемала, Індія, Індонезія, Кенія, Ліберія та інші країни, що розвиваються або є найменш розвиненими. В цілому, рівень кібербезпеки відповідає рівню економічного розвитку країни. Ті, що є розвиненими, відповідно, мають потужні можливості для створення умов кіберзахисту різних об’єктів. Країни, що розвиваються та є найменш розвиненими, мають проблеми в сфері кібербезпеки, викликані відсутністю висококваліфікованих фахівців в цій галузі, недостатнім рівнем інвестування, слабким рівнем правового забезпечення цієї сфери, тощо.

За інтегральним рівнем протидії фінансовим шахрайствам оцінку «дуже добре» отримали 28 країн (див. рис. 1.58): Австралія, Австрія, Бельгія, Канада, Ірландія, Нідерланди, Норвегія, Великобританія, Швеція, Чехія, та інші. Такі країни, як Хорватія, Німеччина, Угорщина, Італія, Малайзія, Мальта та Сингапур, мають рівень протидії легалізації кримінальних доходів на рівні «добре». Оцінку «задовільно» отримали Ботсвана, Болгарія, Китай, Коста Ріка, Греція, Люксембург, Сейшельські острови, Швейцарія, США та Уругвай. 9 країн отримали рівень «погано», а 22 країни – «дуже погано». До них відносяться: Болівія, Бразилія, Індія, Україна, Російська Федерація, Мексика, Південна Африка, Таїланд, Індонезія та інші. Тобто, ряд країн, які мають високий рівень злочинності та тероризму, озброєні конфлікти, низький економічний розвиток є досить привабливими для легалізації кримінальних доходів та фінансування тероризму. Тому система протидії таким операціям є досить слабкою й не розвиненою. Також країни, які мають високий рівень фінансової таємниці

створюють сприятливі умови для відмивання коштів, отриманих злочинним шляхом. На сьогодні такими є Швейцарія, Люксембург та США.

Для визначення рівня конвергенції систем кібербезпеки та протидії фінансовим шахрайствам знайдемо середньоарифметичне значення двох інтегральних індексів. Результати розрахунків представимо у вигляді карти розподілу країн за рівнем конвергенції систем кібербезпеки та протидії фінансовим шахрайствам (див. рис. 1.59).

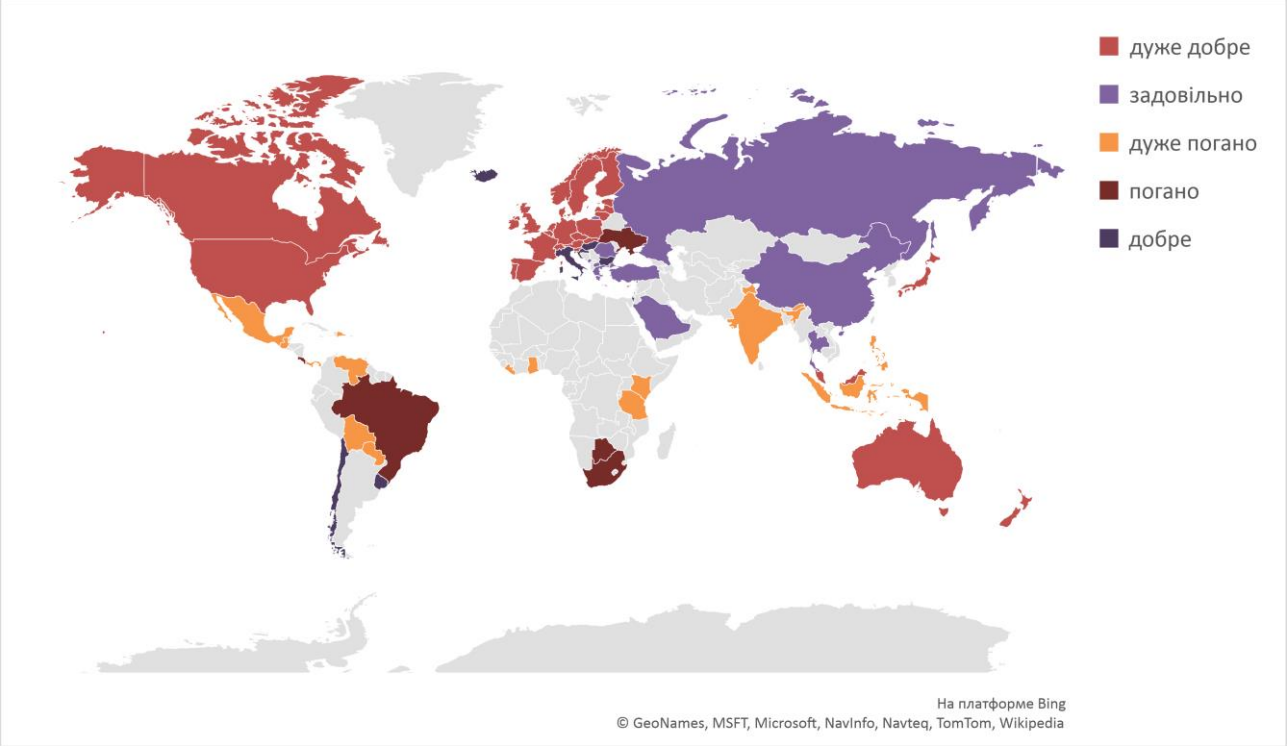


Рисунок 1.59 – Карта розподілу країн за рівнем конвергенції систем кібербезпеки та протидії легалізації кримінальних доходів

За умови конвергенції системи кібербезпеки та протидії фінансовим шахрайствам для тих країн, які мають низький рівень протидії, відбудеться посилення їх потенційних можливостей за рахунок системи кіберзахисту. Так, порівнюючи результати, представлені на рисунках 1.57-1.59, можна побачити, що такі країни, як Бахрейн, Ботсвана, Бразилія, Бруней, Болгарія, Чилі, Коста Рика, Ісландія, Ізраїль, Люксембург, Мальта, Чорногорія, Північна Македонія, Румунія, Російська Федерація, Саудівська Аравія, Сейшельські острови,

Сінгапур, Швейцарія, Таїланд, Туреччина, Україна, США та Уругвай, матимуть позитивний ефект від процесу конвергенції.

Сучасні тенденції зростання обсягів кібершахрайств та легалізації кримінальних доходів вимагають застосування нових методів і технологій в процесі боротьби з даним явищем. Це можливо тільки за рахунок системної взаємодії програмних, технічних, інформаційних, організаційних, правових та технологічних заходів, тобто конвергенції системи кібербезпеки та протидії фінансовим шахрайствам. Цей процес є доволі складним, тому потребує зваженого підходу до його здійснення. Тому здійснення попередньої оцінки рівня потенційної конвергенції цих двох систем є необхідним заходом на шляху удосконалення та підвищення ефективності боротьби із шахрайствами на світовому рівні.

В дослідженні розглянуто індикатори, які характеризують рівень розвитку кібербезпеки країни та протидії легалізації кримінальних доходів і фінансування тероризму. Використаний підхід Харрінгтона – Менчера дозволив сформувати два інтегральні показники. Оцінювання рівня кібербезпеки дозволило виявити, що розвинені країни мають високий рівень кіберзахисту. Найнижчі оцінки отримали країни, що є найменш розвиненими або розвиваються та мають низький рівень розвитку. За інтегральним оцінюванням рівня протидії легалізації кримінальних доходів встановлено, що суттєві проблеми в цій сфері мають країни із високим рівнем злочинності, тероризму, низькою якістю державного управління, а також ті, де здійснюються озброєні конфлікти та є високий рівень фінансової секретності. Це сприяє можливостям відмивання кримінальних доходів та знижує спроможності системи протидіяти таким операціям.

Визначений загальний рівень конвергенції системи кібербезпеки та протидії відмиванню кримінальних доходів дозволив зробити висновок, що цей процес матиме позитивний ефект для 32% країн з досліджуваного набору. Тобто можна говорити про те, що інтеграційні процеси є сприятливими для посилення можливостей країн у боротьбі з фінансовими та кібершахрайствами. В

подальшому, планується оцінити потенційний ефект від здійснення даних процесів для визначених груп країн.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**, **Ошибка! Источник ссылки не найден.**].

1.4 Оцінка синергетичного ефекту від конвергенції моделей фінансового моніторингу та кібербезпеки

Запропонована методологія інтегрального оцінювання рівня конвергенції системи фінансового моніторингу та кібербезпеки дозволила розробити сценарії шляхом ранжування країн, використовуючи якісну шкалу оцінювання, запропоновану Харрінгтоном-Менчером. Але інтегральне уявлення не дозволяє повністю зробити висновки щодо ефективності системи забезпечення конвергенції системи фінансового моніторингу та кібербезпеки, тобто необхідно здійснити структурний аналіз її компонентів, який дозволить оцінити не тільки поточний стан складових інтегрального показника конвергенції системи фінансового моніторингу та кібербезпеки, але й визначити резерви для його підвищення. З цією метою необхідно провести аналіз ефективності системи забезпечення конвергенції системи фінансового моніторингу та кібербезпеки та надати рекомендації щодо напрямів її покращання.

В залежності від мети аналізу у науковій літературі та на практиці застосовують різні підходи. Традиційними вважаються: класична модель Дюпона «Рентабельність капіталу», процесно-орієнтований аналіз рентабельності М. Мейера та В. Маршала, «Управління результатами» Р. Каплана та Д. Нортона, методика аналізу, заснована на аналізі грошових потоків.

Основними недоліками перелічених методів є те, що їх використання доцільно для аналізу ефективності господарської діяльності та передбачає розрахунок різних коефіцієнтів, за результатами яких робиться висновок. У

випадку оцінювання ефективності системи забезпечення конвергенції системи фінансового моніторингу та кібербезпеки доцільно використання саме математичних методів, які дозволяють проводити оцінювання параметрів відносно значень, які є найкращими у групі аналізованих об’єктів. Саме тому для проведення дослідження було використано DEA-метод (Data Envelopment Analysis), який було запропоновано А. Чарнсом, В. Купером та Е. Родесом у 1978 році [**Ошибка! Источник ссылки не найден.**]. Цей інструмент не залежить від мети аналізу та використовується у багатьох галузях для оцінки ефективності складних систем, що відбувається шляхом рішення оптимізаційної задачі лінійного програмування. Її мета – це визначення ефективності системи на основі співвідношення її виходів та входів, при цьому необхідно врахувати максимальний вихід ресурсів при заданому рівні входів, або мінімальний рівень ресурсів при заданому рівні виходів.

Для проведення дослідження було обрано вхідні дані, які було відібрано у розділі 3 на основі канонічного аналізу та розрахунку інтегрального показника конвергенції системи фінансового моніторингу та кібербезпеки, а саме: глобальний індекс кібербезпеки; національний індекс кібербезпеки; індекс мережевої готовності; рівень цифрового розвитку; індекс політичної стабільності; індекс ефективності уряду; легкість ведення бізнесу; індекс злочинності; глобальний індекс тероризму; індекс фінансової таємниці. Дані показники сформували базу вхідних даних для 76 країн світу за 2018 рік. У якості вихідного параметру, який є індикатором узагальненого рівня ефективності, виступатиме запропонований у підрозділі 3.2 інтегральний індекс конвергенції системи фінансового моніторингу та кібербезпеки.

Оскільки DEA-метод є ефективним для даних, які мають близькі характеристики, то доцільно сформувати кластери країн. Було проведено сортування за інтегральним індексом конвергенції системи фінансового моніторингу та кібербезпеки та виділено 7 груп країн.

Так, до 0-го кластеру увійшли 12 країн: Фінляндія, Австралія, Австрія, Данія, Канада, Ірландія, Швеція, Нова Зеландія, Норвегія, Естонія, Бельгія,

Португалія. До 1-го кластеру увійшли: Іспанія, Литва, Словенія, Латвія, Нідерланди, Японія, Великобританія, Франція, Кіпр, Чехія, Німеччина, Словаччина. До 2-го кластеру віднесено: Польща, Малайзія, Сінгапур, Швейцарія, США, Люксембург, Угорщина, Хорватія, Мавританія, Італія, Ісландія, Уругвай. До 3-го кластеру увійшли: Чилі, Мальта, Ізраїль, Болгарія, Греція, Саудівська Аравія, Росія, Чорногорія, Бруней, Північна Македонія, Бахрейн, Китай. До 4-го кластеру віднесено: Туреччина, Таїланд, Румунія, Коста Рика, Південна Африка, Сейшельські острови, Україна, Бразилія, Ботсвана, Мексика, Індонезія, Панама. До 5-го кластеру: Тринідад і Тобаго, Барбадос, Філіппіни, Індія, Домініканська Республіка, Гана, Домініка, Парагвай, Кенія, Гренада, Вануату, Венесуела. До 6-го кластеру: Болівія, Гватемала, Танзанія, Ліберія.

Використання DEA-методу дозволить визначити ефективність конвергенції системи фінансового моніторингу та кібербезпеки з урахуванням потенціалу країни. Ефективність буде досягатися тоді, коли рівень протидії загрозам для окремої країни не можливо збільшити, при цьому залишивши рівень розвитку та безпеки країни на тому самому рівні. Також це можливо у випадку, коли зменшення рівня розвитку та безпеки країни призводить до змін рівня протидії кіберзагрозам. Виходячи з вище сказаного, можна сформувати початкову DEA-модель [**Ошибка! Источник ссылки не найден.**], яку буде використано для проведення оцінки ефективності рівня інформаційної безпеки країни за формулою (1.42):

$$\max \theta_s = \frac{\sum_{p=1}^{z} u_{ps} y_{ps}}{\sum_{i=1}^{m} v_{is} x_{is}}$$

$$\begin{cases} \dfrac{\sum_{p=1}^{z} u_{ps} y_{pj}}{\sum_{i=1}^{m} v_{is} x_{i}\, j} \le 1, \\ s, j = \overline{1, n}, \\ u_p, v_i \ge 0, \\ y_{p,} x_i \ge 0. \end{cases} \qquad (1.42)$$

де θ – рівень ефективності конвергенції системи фінансового моніторингу та кібербезпеки для конкретної країни, визначений як коефіцієнт між зваженою сумою виходів та входів;

u_p – ваги виходів, які максимізують показник ефективності оцінюваної одиниці θ;

v_p – ваги входів, які максимізують показник ефективності оцінюваної одиниці θ;

y_p – p-та характеристика умовних виходів, тобто значень індексу конвергенції системи фінансового моніторингу та кібербезпеки для кожної країни;

x_i – i-та характеристика умовних входів, тобто значень показників системи фінансового моніторингу та кібербезпеки.

Обмеження (1.42) говорять про те, що відношення виходу до входу не може перевищувати 1 для кожної θ. Тому представлену дробову задачу слід перетворити на лінійну, що значно спрощує її подальше використання. Відповідно до цього, розрізняють два типи DEA-моделі – CCR (Charnes A., Cooper W. and Rhodes E.), яку було запропоновано Чарнсом А., Купером У. та Родесом Е. [**Ошибка! Источник ссылки не найден.**] у 1978 році, та BCC (Banker R., Charnes A. and Cooper W.), яку було розроблено на основі CCR-моделі у 1984 році Банкером Р., Чарнсом А. та Купером У. [**Ошибка! Источник ссылки не найден.**]. Кожна з цих моделей (1.43) – (1.46) орієнтована на вхід (ресурси) та вихід (результуючі показники):

$$\max_{u,v} \theta_s = \sum_{p=1}^{z} u_{ps} y_{ps} \tag{1.43}$$

$$\begin{cases} \sum_{i=1}^{m} v_{is}x_{is} = 1 \\ \sum_{p=1}^{z} u_{ps}y_{pj} - \sum_{i=1}^{m} v_{is}x_{ij} \le 0 \\ u_p, v_i \ge \gamma \end{cases}$$

$$\max_{u,v,k} \theta_s = \sum_{p=1}^{z} u_{ps}y_{ps} + k_s$$

$$\begin{cases} \sum_{i=1}^{m} v_{is}x_{is} = 1 \\ \sum_{p=1}^{z} u_{ps}y_{pj} + k_s \le \sum_{i=1}^{m} v_{is}x_{ij} \\ u_p, v_i \ge \gamma \\ k_s - \ unconstrained \end{cases} \quad (1.44)$$

$$\min_{\alpha,\beta,k} \theta_s = \sum_{i=1}^{m} \beta_i x_{is} - k_s$$

$$\begin{cases} \sum_{p=1}^{z} \alpha_p y_{ps} = 1 \\ \sum_{i=1}^{m} \beta_i x_{ij} - k_s \ge \sum_{p=1}^{z} \alpha_p y_{pj} \\ \alpha_p, \beta_i \ge \gamma \\ k_s - \ unconstrained \end{cases} \quad (1.45)$$

$$\min_{\alpha,\beta} \theta_s = \sum_{i=1}^{m} \beta_i x_{is}$$

$$\begin{cases} \sum_{p=1}^{z} \alpha_p y_{ps} = 1 \\ \sum_{i=1}^{m} \beta_i x_{ij} - \sum_{p=1}^{z} \alpha_p y_{pj} \ge 0 \\ \alpha_p, \beta_i \ge \gamma \end{cases} \quad (1.46)$$

де γ – це невелике додатне дійсне число, яке виключає можливість набуття змінними нульового значення.

Моделі CCR (1.43) та BCC (1.44) є Input-oriented моделями, тобто направлені на оцінку ефективності розподілу показників, що характеризують системи фінансового моніторингу та кібербезпеки, що сприяє виявленню структурної неефективності заданих індексів. Моделі CCR (1.45) та BCC (1.46) є Output-oriented, тобто дозволяють здійснити оцінку ефективності конвергенції системи фінансового моніторингу та кібербезпеки країни шляхом визначення максимальних значень індексу конвергенції системи фінансового моніторингу та кібербезпеки за умови заданих значень показників, що характеризують системи фінансового моніторингу та кібербезпеки.

DEA-аналіз було проведено у аналітичному пакеті “Frontier Analyst”, який дозволяє здійснювати розрахунки за моделями CCR та BCC [**Ошибка! Источник ссылки не найден.**]. Оскільки було використано демо-версію, то для дослідження в кожному кластері країн було обрано 12 представників, для яких проводився Data Envelopment Analysis. Мінімальне значення вагів у програмі було встановлено на основі результатів проведеного канонічного аналізу за допомогою аналітичної платформи “STATISTICA”, що дозволило визначити частки їх значень у загальній сукупності. Так, для глобального індексу кібербезпеки визначено вагу, що дорівнює 0,3133, індексу мережевої готовності – 0,2644, національного індексу кібербезпеки – 0,0213, рівня цифрового розвитку – 0,5578, індекс політичної стабільності – 0,2691; індекс ефективності уряду – 0,7808; легкість ведення бізнесу – 0,3417; індекс злочинності – 0,0501; глобальний індекс тероризму – 0,0093; індекс фінансової таємниці – 0,0915. Максимальне значення вагів було встановлено на рівні 100%, тому відповідні значення було перераховано у пропорції.

Модель CCR є більш обмежувальною ніж BCC. Це пов’язано із тим, що вона базується на постійності віддачі від масштабу, а також дає можливість масштабувати неефективні одиниці вибірки. BCC-модель базується на змінній

віддачі від масштабу та дозволяє оцінити технічну ефективність. Така зміна її вхідних параметрів може призводити до непропорційної зміни вихідних, що дозволяє оцінювати більшість об'єктів як ефективні. Тому для визначення синергетичних ефектів будемо використовувати тільки BCC-модель.

Проаналізуємо структурну ефективність вхідних показників для країн 0-го кластеру, отриману в результаті проведення аналізу за Input-oriented CCR-model (рисунок 1.60). Отримані значення всіх показників є від'ємними, тобто забезпечення поточного рівня конвергенції системи фінансового моніторингу та кібербезпеки країн 0-го кластеру відбувається із досягненням ефективності по кожному напряму – фінансового моніторингу та кібербезпеки. При чому можна побачити, що є потенціал для забезпечення рівня конвергенції (8,51%). Слід відмітити, що для даних країн найбільший резерв формується саме за глобальним індексом тероризму, що свідчить про те, що для даних країн відсутні сприятливі умови для легалізації кримінальних доходів та фінансування тероризму.

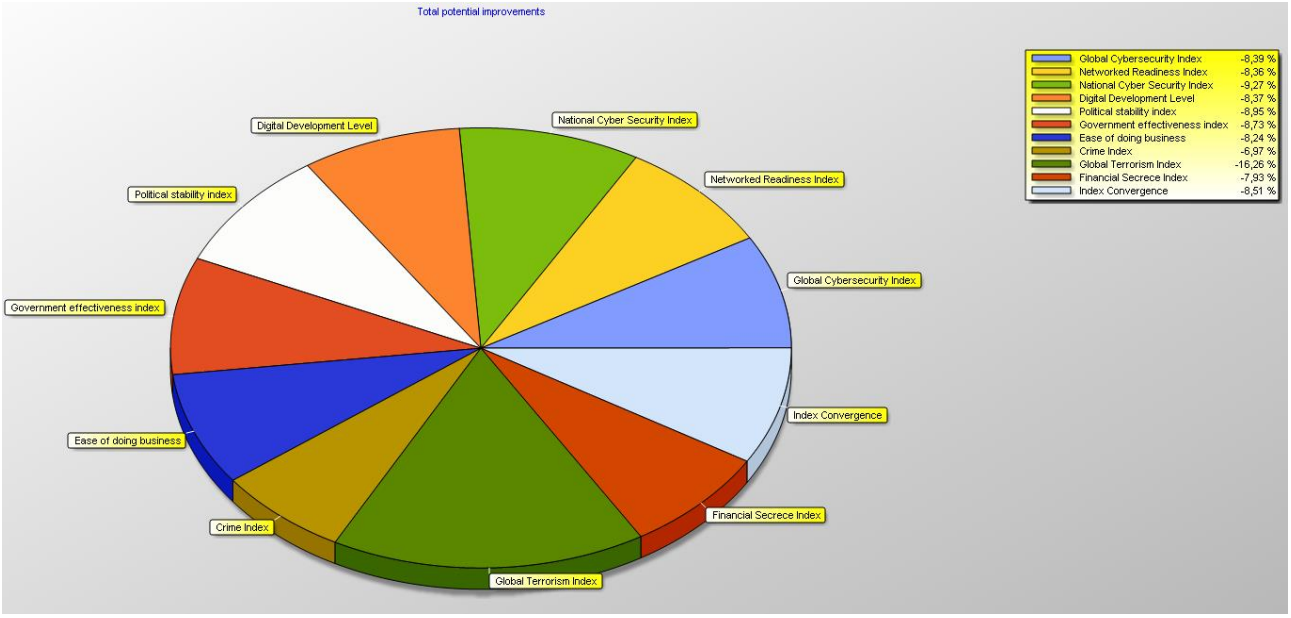


Рисунок 1.60 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 0-го кластеру (за Input-oriented CCR-моделлю)

Проведемо аналіз потенціалу покращання ефективності конвергенції системи фінансового моніторингу та кібербезпеки 0-го кластера за умови максимізації інтегрального індексу конвергенції системи фінансового моніторингу та кібербезпеки. Результати Output-oriented CCR-model представлено на рисунку 1.61, де можна побачити, що максимальне зростання індексу конвергенції системи фінансового моніторингу та кібербезпеки можливе на 6,6%. Це можливо забезпечити за рахунок резервів потенціалу за показниками: національний індекс кібербезпеки (-4,18%), індекс злочинності (-36,48%). Тобто країни 0-го кластеру мають, з одного боку, високий потенціал розвитку кібербезпеки в країні, достатній для забезпечення підвищення рівня конвергенції, а з іншого боку, низький рівень злочинності створює передумови для формування ефективної системи фінансового моніторингу.

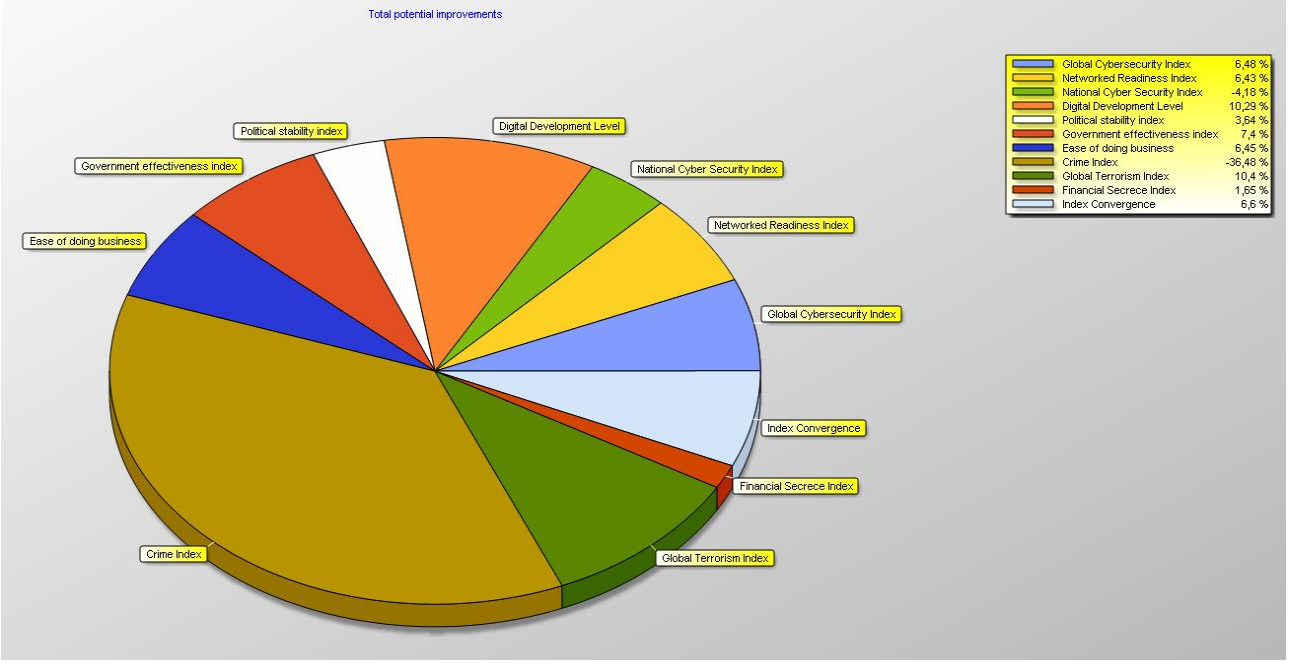


Рисунок 1.61 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 0-го кластеру (за Output-oriented CCR-моделлю)

Проаналізуємо структурну ефективність вхідних показників для країн першого кластеру (рисунок 1.62). Отримані значення всіх показників є

додатними, що свідчить про те, що не можливо досягти рівня конвергенції системи фінансового моніторингу та кібербезпеки за рахунок поточного стану їх функціонування. Хоча країни даного кластеру мають розвинену економіку, але існують проблеми на національному рівні, які переважають здійснення конвергенції систем. Найбільшого удосконалення потребує індекс ефективності уряду, який необхідно підвищити на 11,67%, а також індекс мережевої готовності (10,24%). Для забезпечення поточного рівня конвергенції систем країн 1-го кластеру необхідно підвищити ефективність функціонування всіх відповідних напрямів, які характеризують систему кібербезпеки в країні та протидії фінансовим кіберзлочинам.

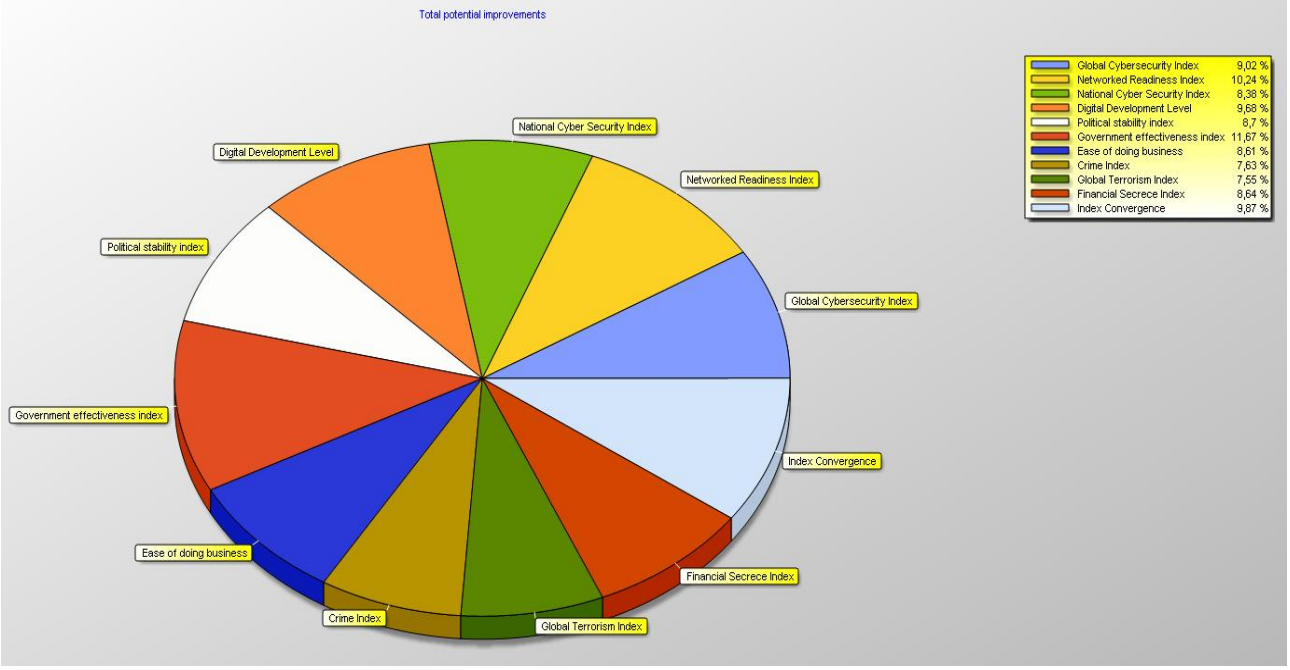


Рисунок 1.62 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 1-го кластеру (за Input-oriented CCR-моделлю)

Проаналізуємо структурну ефективність вихідних показників для країн першого кластеру (рисунок 1.63).

Аналіз потенціалу покращання ефективності конвергенції систем фінансового моніторингу та кібербезпеки першого кластера за умови

максимізації інтегрального індексу конвергенції показує, що його максимальне зростання можливе на 8,01%. Це можливо забезпечити за рахунок резервів потенціалу за показниками: глобальний індекс кібербезпеки (-7,74%), мережевий індекс готовності (-7,74%), рівень цифрового розвитку (-7,78%) та національний індекс кібербезпеки (-9,04%). Тобто країни першого кластеру мають значний потенціал системи кібербезпеки, достатній для забезпечення підвищення рівня конвергенції систем. Щодо системи протидії фінансовим злочинам, то ці країни в даній сфері також мають значний потенціал, а саме: індекс політичної стабільності (-8,61%); індекс ефективності уряду (-8,65%); легкість ведення бізнесу (-7,61%); індекс злочинності (-10,56%); глобальний індекс тероризму (-14,98%); індекс фінансової таємниці (-9,29%).

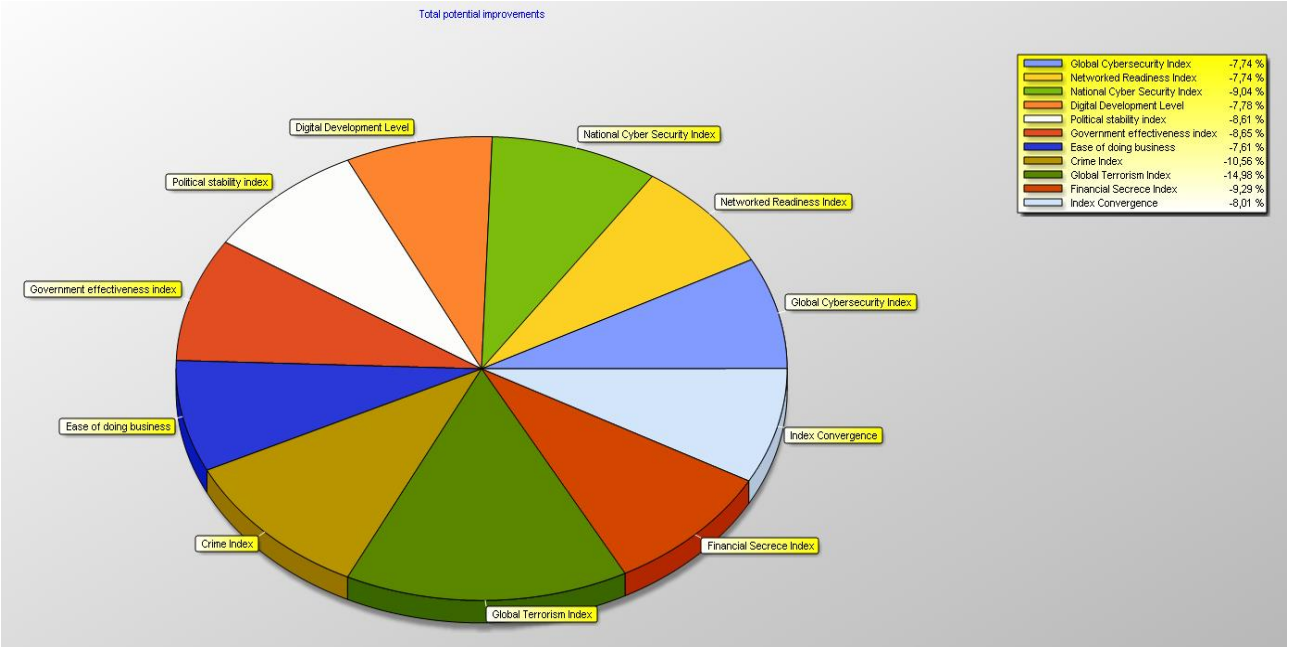


Рисунок 1.63 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 1-го кластеру (за Output-oriented CCR-моделлю)

Тобто, країни 1-го кластеру не можуть досягнути поточного рівня конвергенції, але існуючий потенціал може забезпечити максимальний рівень, який є нижче поточного.

Проаналізуємо структурну ефективність вхідних показників для країн другого кластеру (рисунок 1.64). Отримані значення всіх показників є додатними, що свідчить також про те, що країни цього кластеру не можуть досягти поточного рівня конвергенції системи фінансового моніторингу та кібербезпеки, який потребує зростання на 12,4%. Найбільшого удосконалення потребують індекси кібербезпеки, які необхідно підвищити вище ніж на 10% кожний. Результати показують, що для забезпечення ефективного процесу інтеграції необхідно звернути увагу саме на рівень кібербезпеки. Наприклад, в даному кластері знаходиться США, які сьогодні займають перше місце серед країн, які є атакованими з боку інших країн. При цьому ця країна є також лідером щодо здійснення кібератак на інші країни. Тому проблема, пов'язана із кіберзахистом, є актуальною для країн даного кластеру.

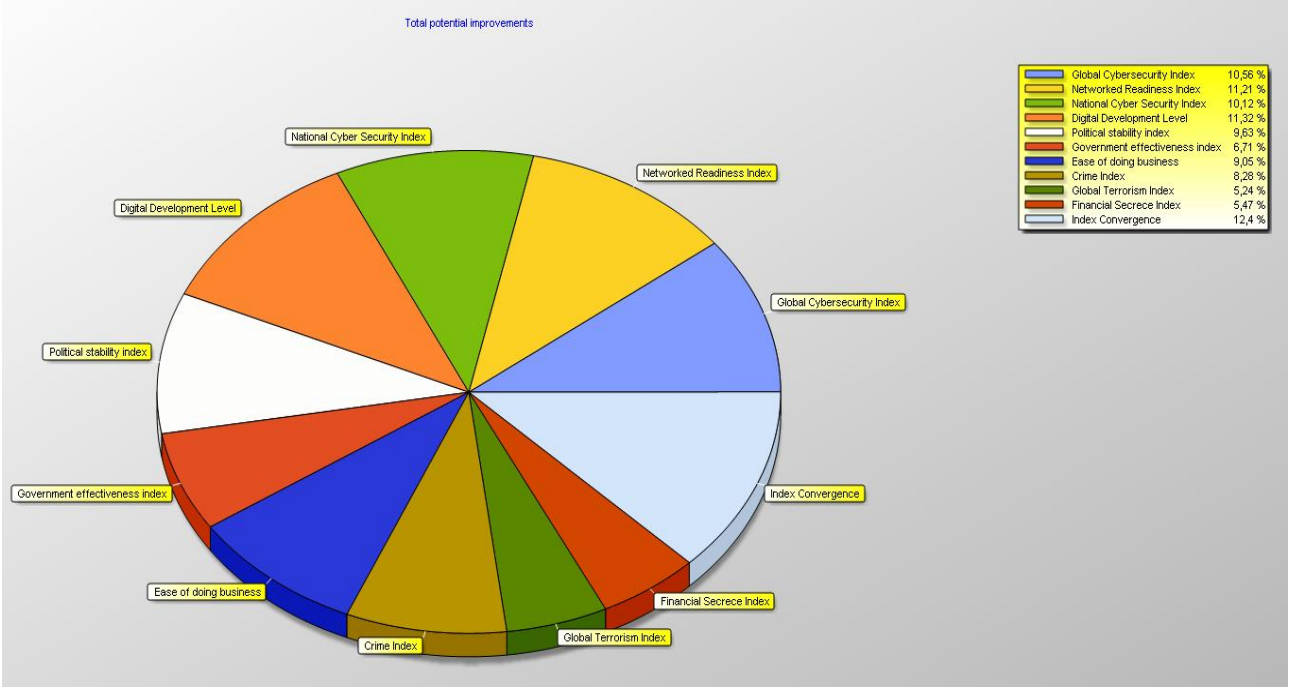


Рисунок 1.64 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 2-го кластеру (за Input-oriented CCR-моделлю)

Проаналізуємо структурну ефективність вихідних показників для країн другого кластеру (рисунок 1.65).

Аналіз потенціалу покращання ефективності конвергенції систем фінансового моніторингу та кібербезпеки другого кластера за умови максимізації інтегрального індексу конвергенції показує, що його максимальне зростання можливе тільки на 0,54%. Це можливо забезпечити за рахунок резервів потенціалу за показниками: глобальний індекс кібербезпеки (-7,85%), мережевий індекс готовності (-4,06%), рівень цифрового розвитку (-12,48%) та національний індекс кібербезпеки (-13,59%). Тобто країни даного кластеру мають потенціал системи кібербезпеки, достатній для незначного підвищення рівня конвергенції систем. Дані країни мають потенціал системи протидії фінансовим злочинам, а саме: індекс політичної стабільності (-3,73%); індекс ефективності уряду (-7,45%); індекс злочинності (-9,55%); глобальний індекс тероризму (-21,27%); індекс фінансової таємниці (-18,03%).

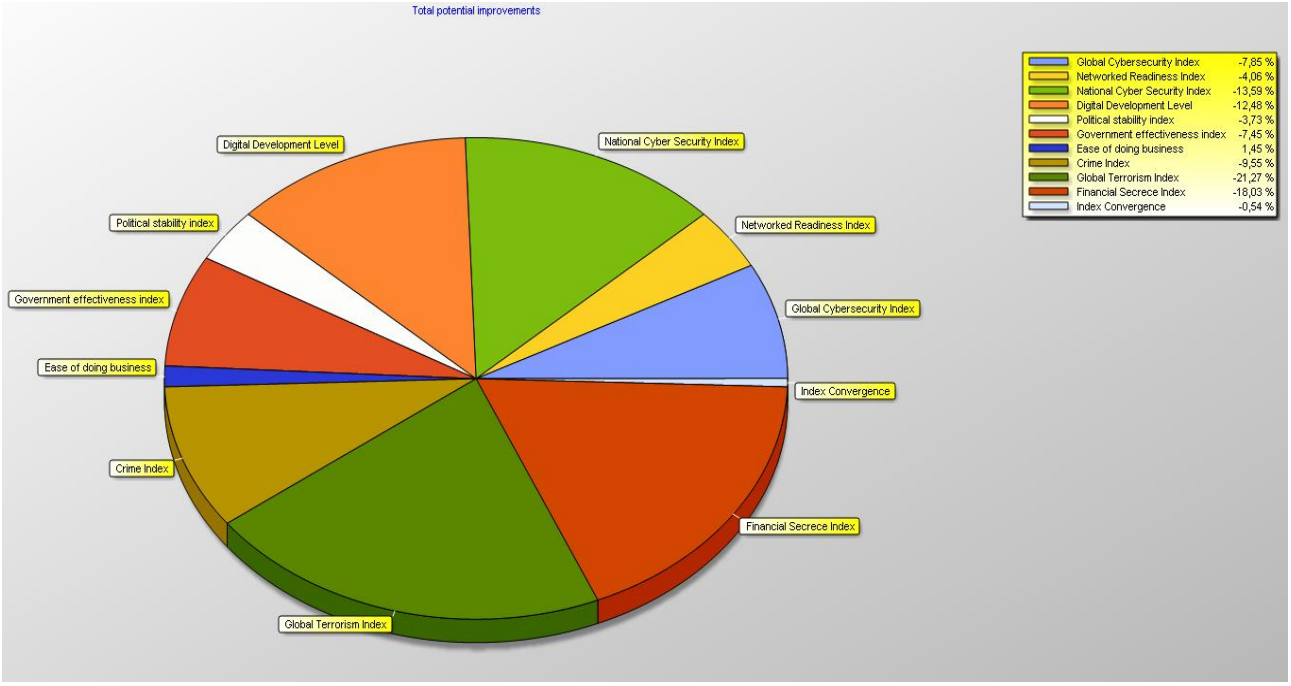


Рисунок 1.65 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 2-го кластеру (за Output-oriented CCR-моделлю)

Тобто, країни 2-го кластеру не можуть досягнути поточного рівня конвергенції, але досягнення максимального рівня ефективності є досить незначним, оскільки спостерігається розбалансованість систем фінансового моніторингу і кібербезпеки.

Проаналізуємо структурну ефективність вхідних показників для країн 3-го кластеру (рисунок 1.66). Отримані значення всіх показників є від'ємними, що свідчить не тільки про забезпечення поточного рівня конвергенції системи фінансового моніторингу та кібербезпеки, але й його перевищення на 10,94%. При цьому спостерігається досягнення ефективності по всім показникам, що характеризують систему фінансового моніторингу та кібербезпеки. Найбільше значення характерне для індексу ефективності уряду, яке дозволяє його підвищення на 11,22%. Тобто політика уряду цих країн є настільки ефективною, що створюються можливості для протидії легалізації кримінальних доходів. Але, оскільки рівень конвергенції країн цього кластеру є нижче, ніж для країн 2-го кластеру, то отриманий результат говорить тільки про те, що країни досягли певного рівня конвергенції, який відповідає рівню їх економічного розвитку.

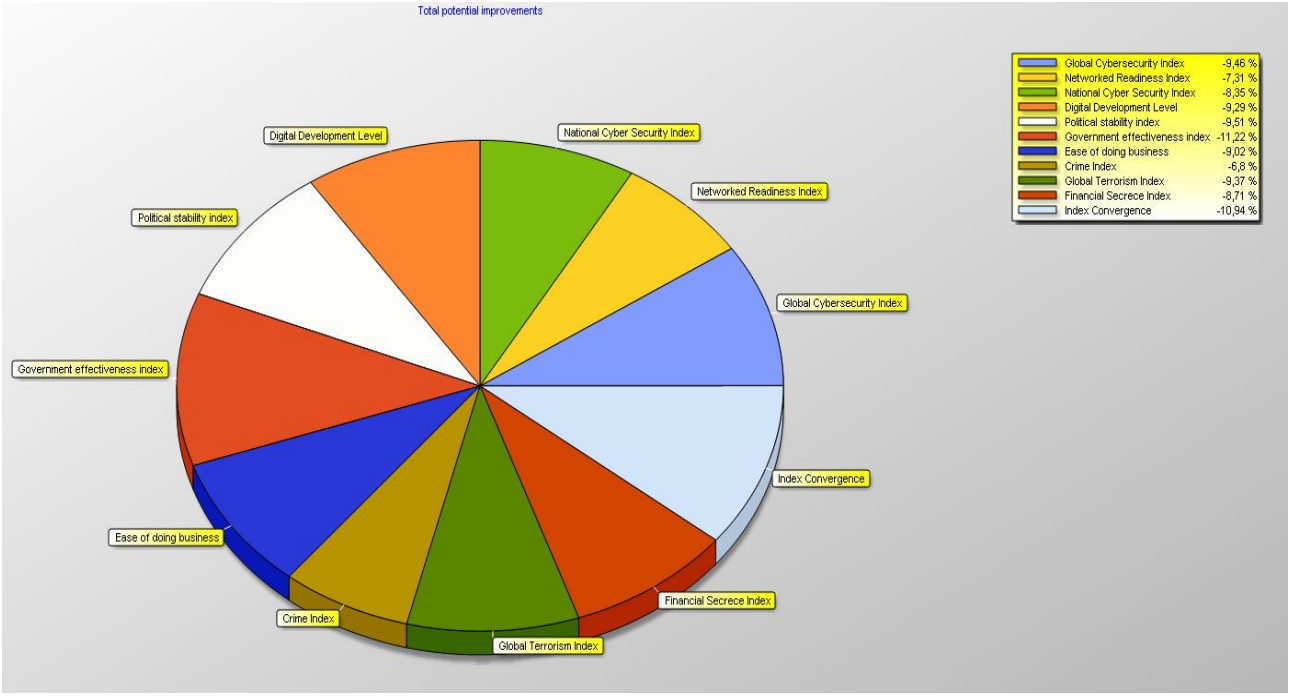


Рисунок 1.66 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 3-го кластеру (за Input-oriented CCR-моделлю)

Проаналізуємо структурну ефективність вихідних показників для країн третього кластеру (рисунок 1.67).

Результати Output-oriented CCR-model (рис. 4.8) показують, що максимальне зростання індексу конвергенції системи фінансового моніторингу та кібербезпеки можливе на 15,79%. Це можливо забезпечити за рахунок резервів за показниками: національний індекс кібербезпеки (-14,57%), індекс політичної стабільності (-12,13%) та індексу злочинності (-5,03%). Всі інші фактори потребують відповідного покращення, щоб забезпечити максимальний рівень конвергенції систем. Оскільки фактичний рівень конвергенції забезпечується та спостерігається його перевищення, то країни 3-го кластеру мають значний потенціал для ефективного рівня інтеграції двох систем.

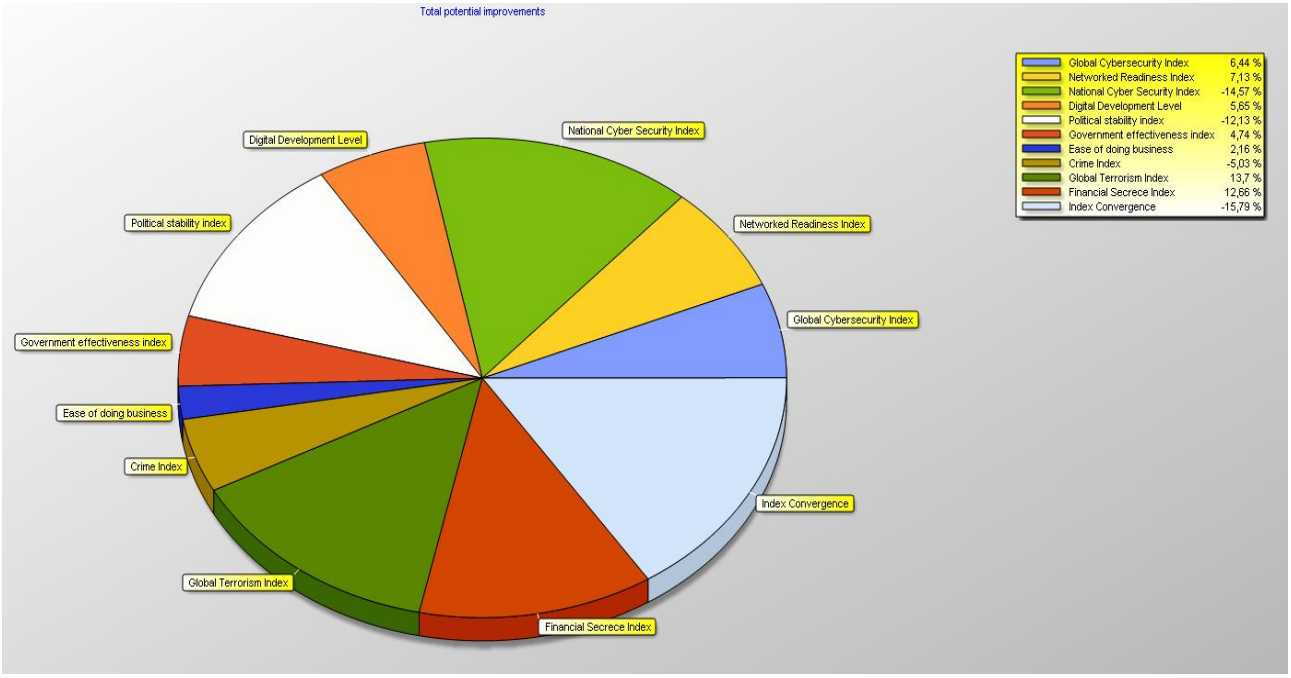


Рисунок 1.67 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 3-го кластеру (за Output-oriented CCR-моделлю)

Проаналізуємо структурну ефективність вхідних показників для країн четвертого кластеру (рисунок 1.68). Отриманий інтегральний рівень свідчить про те, що його фактичне значення не може бути забезпечено на 0,12%. Це відбувається за рахунок високого рівня злочинності в даних країнах (93,69%). При чому показники, що характеризують систему фінансового моніторингу є додатними, тобто для даних країн характерні сприятливі умови для легалізації кримінальних доходів та фінансування тероризму. Але система кібербезпеки має незначний резерв в межах 1% по кожному показнику безпеки, що свідчить про можливості підтримки системи фінансового моніторингу за рахунок системи кібербезпеки.

Проаналізуємо структурну ефективність вихідних показників для країн другого кластеру (рисунок 1.69). Результати Output-oriented CCR-model показують, що можливе тільки максимальне зниження індексу конвергенції системи фінансового моніторингу та кібербезпеки на 13,23%. Показники не мають резервів зростання, оскільки отримані значення є додатними. Така ситуація може свідчити тільки про те, що країни цього кластеру мають серйозні проблеми, пов’язані із організацією системи протидії фінансовим кібершахрайствам, а також забезпечення кіберзахисту. Тому інтеграція цих систем не сприятиме отриманню синергетичного ефекту.

До даного кластеру відноситься Україна. Показники її ефективності представлені на рисунках 1.70-1.71.

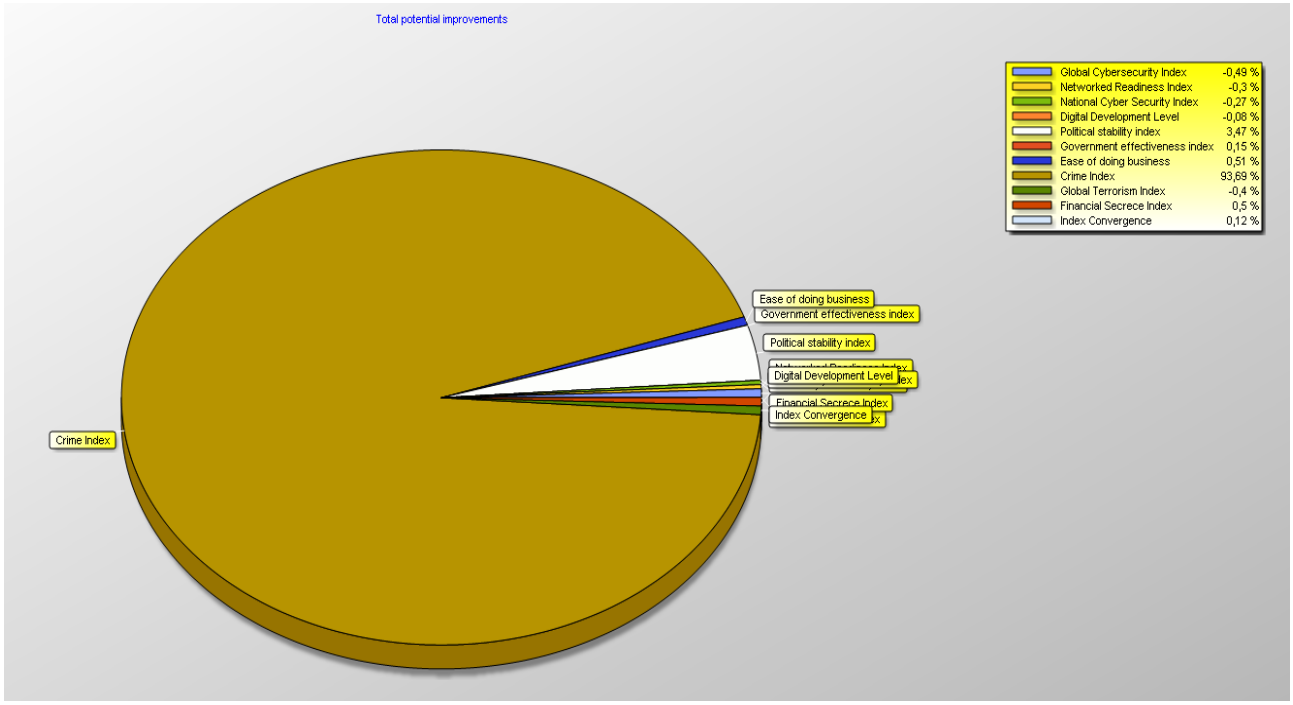


Рисунок 1.68 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 4-го кластеру (за Input-oriented CCR-моделлю)

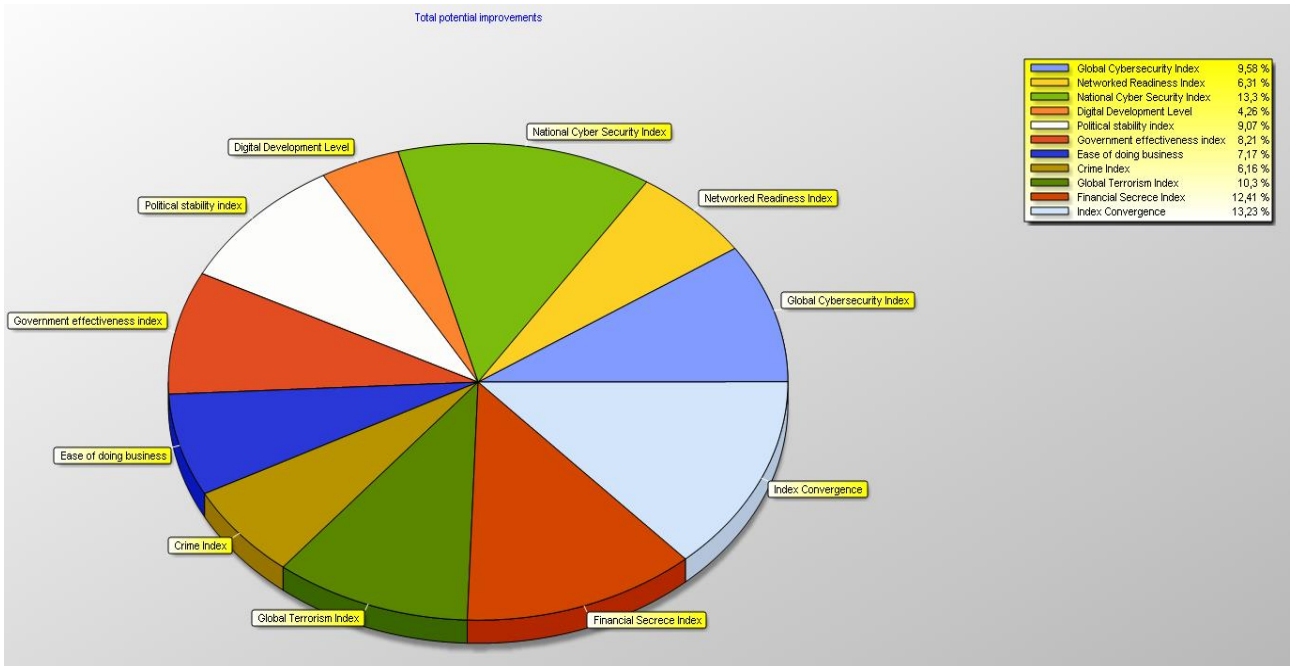


Рисунок 1.69 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 4-го кластеру (за Output-oriented CCR-моделлю)

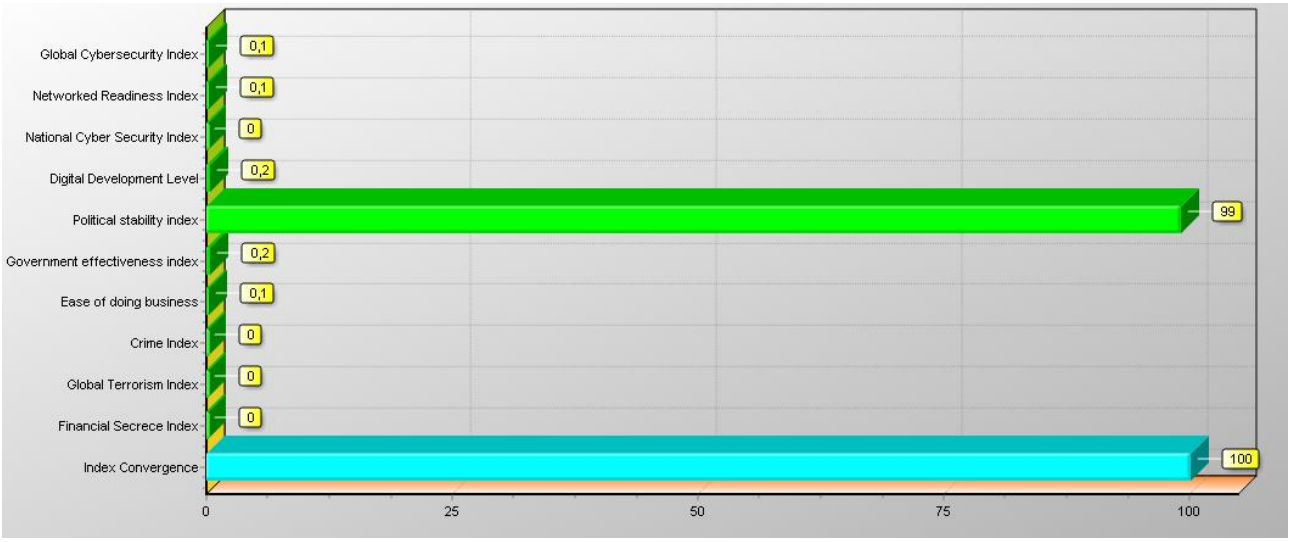


Рисунок 1.70 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки України (за Input-oriented CCR-моделлю)

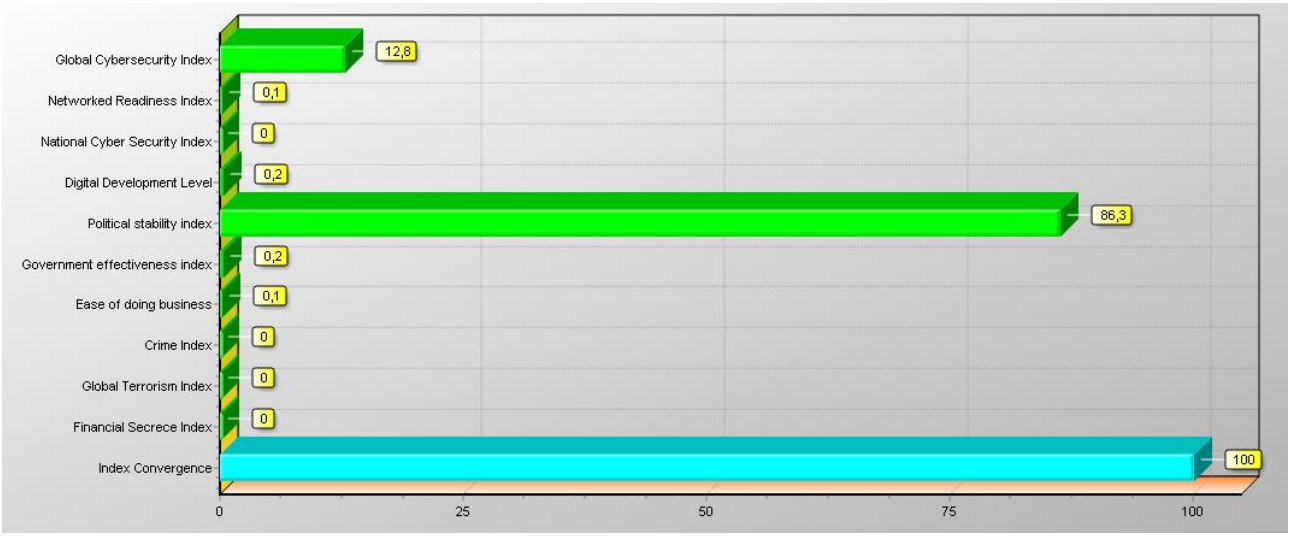


Рисунок 1.71 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки України (за Output-oriented CCR-моделлю)

Україна є представником 4-го кластеру. Ефективність конвергенції її системи фінансового моніторингу та кібербезпеки забезпечується на рівні 100,0% по відношенню до інших країн кластеру. Всі показники або близькі до 0 або є додатними, що свідчить про відсутність резервів зростання рівня інтеграції для України. Але такий показник, як індекс політичної стабільності, має

критичне значення - 99%. Тобто він потребує значного покращання для забезпечення ефективності системи інформаційної безпеки на фактичному рівні. Дана ситуація обумовлюється кризою політичної влади та наявністю військового конфлікту в країні. Оскільки вплив даного показника є досить вагомим, то першочерговим завданням для забезпечення ефективності конвергенції систем повинно бути саме урегулювання даної ситуації.

Що стосується максимального рівня, який може досягнути Україна, то на слайді можна побачити, що на даному етапі Україна досягла максимального рівня конвергенції. Не можливо покращити це значення за рахунок глобального рівня кібербезпеки (12,8%) та індексу політичної стабільності (86,3%). Це є цілком логічним, оскільки ті ризики, які на сьогодні сформовані в країні, мають значний вплив й на глобальне середовище. Відповідно така ситуація вимагає розробки спеціальних заходів кібербезпеки, а також налагодження політичної ситуації в країні.

Аналіз структурної ефективність вхідних показників для країн 5-го кластеру (рисунок 1.72) показав, що в даних країнах забезпечується поточний рівень конвергенції системи фінансового моніторингу та кібербезпеки та відбувається його перевищення на 4,7%. Ефективність досягається за такими показниками, як: мережевий індекс готовності (-27,41%), національний індекс кібербезпеки (-15,22%), індекс ефективності уряду (-3,75%); легкість ведення бізнесу (-5,4%); глобальний індекс тероризму (-2,89%). Тобто країни даного кластеру мають потенціал розвитку національної системи кібербезпеки, а також можливості до впровадження сучасних інформаційних технологій, що необхідно для забезпечення відповідного рівня конвергенції систем. Також сформовані умови, сприятливі для ведення бізнесу. Рівень конвергенції країн цього кластеру є нижчим у порівнянні з країнами попередніх кластерів. Отриманий результат свідчить тільки про невеликі позитивні кроки, необхідні для конвергенції систем.

Проаналізуємо структурну ефективність вихідних показників для країн 5-го кластеру (рисунок 1.73).

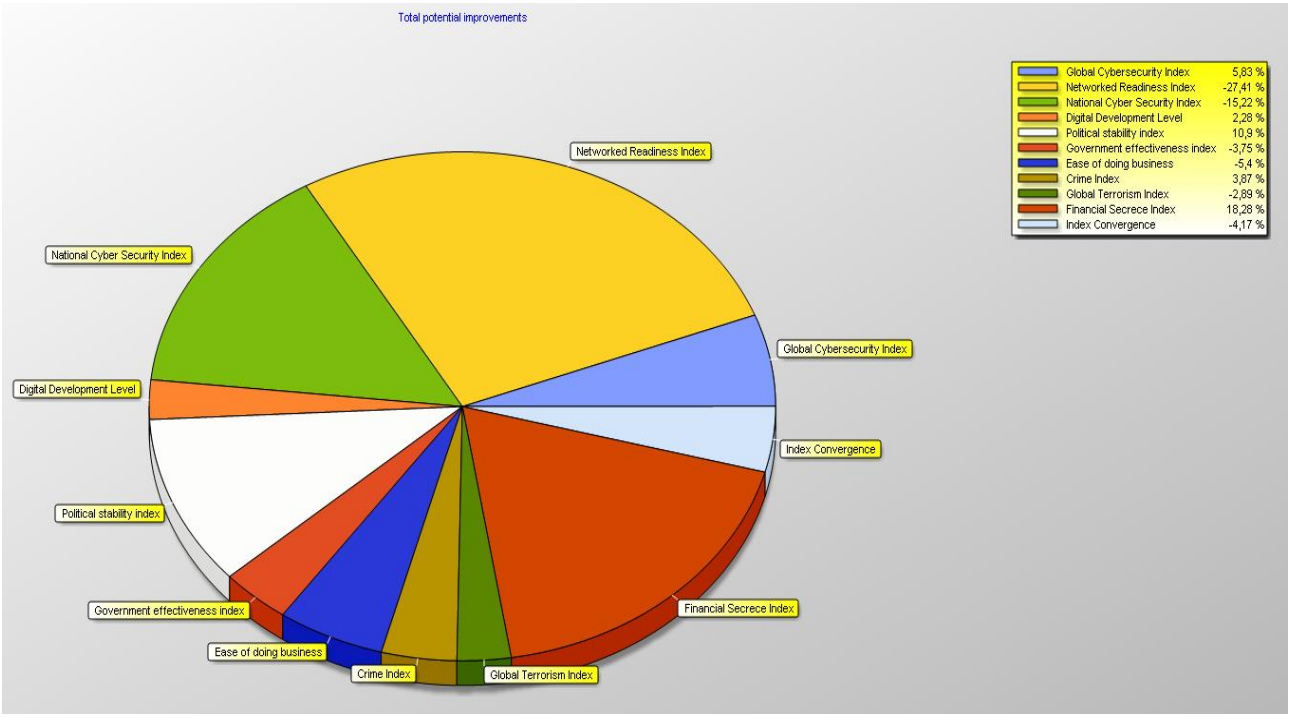


Рисунок 1.72 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 5-го кластеру (за Input-oriented CCR-моделлю)

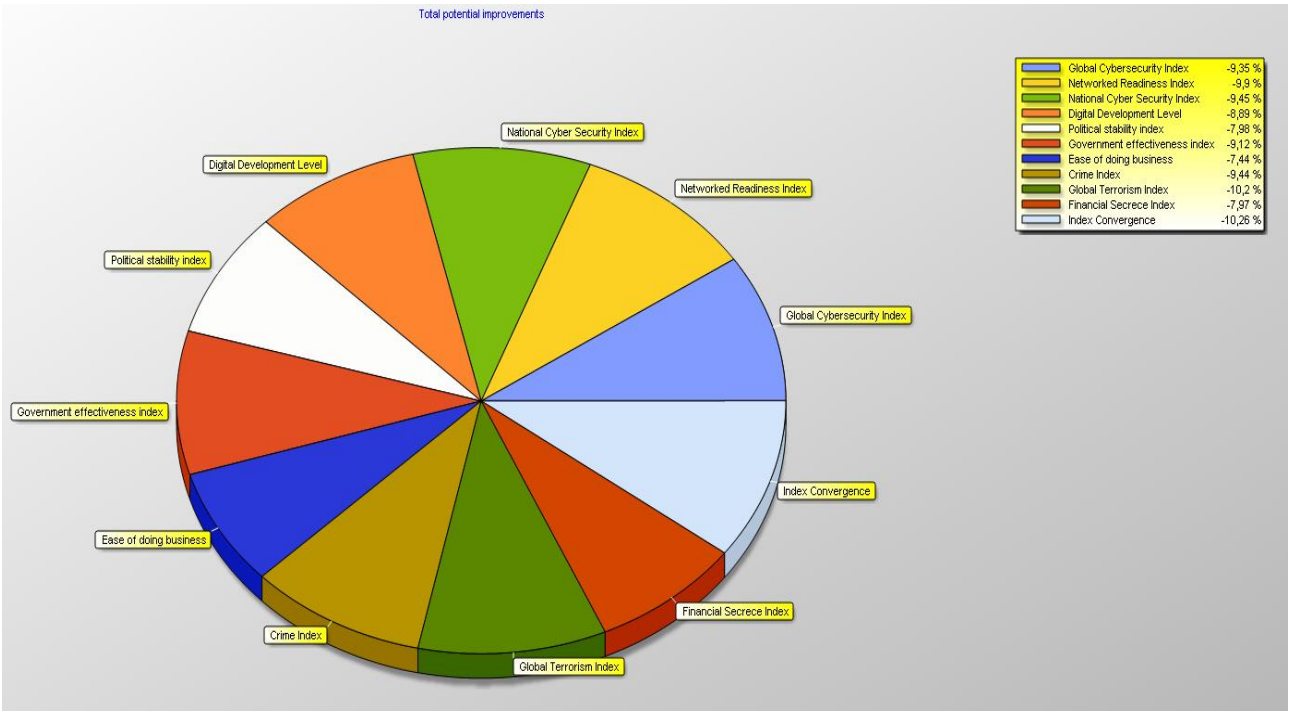


Рисунок 1.73 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 5-го кластеру (за Output-oriented CCR-моделлю)

Результати Output-oriented CCR-model (рис. 1.73) показують, що максимальне зростання індексу конвергенції системи фінансового моніторингу та кібербезпеки можливе на 10,26%. Це можливо забезпечити за рахунок резервів за всіма показниками. Оскільки фактичний рівень конвергенції забезпечується та спостерігається його перевищення, то країни 5-го кластеру також мають потенціал для ефективної інтеграції двох систем.

Шостий кластер сформували тільки 4 країни із найнижчим рівнем конвергенції систем. Аналіз структурної ефективність їх вхідних показників для (рисунок 1.74) показав, що в даних країнах забезпечується поточний рівень конвергенції системи фінансового моніторингу та кібербезпеки. Ефективність досягається за такими показниками, як: мережевий індекс готовності (-0,14%), національний індекс кібербезпеки (-0,36%), рівень цифровізації (-0,35%), індекс ефективності уряду (-0,24%), індекс фінансової таємниці (-0,13%). Тобто країни даного кластеру можуть мати певний успіх в розвитку, який буде досягатися шляхом системної інтеграції системи фінансового моніторингу і кібербезпеки. Але значна проблема, пов’язана з тероризмом та забезпеченням глобального рівня кібербезпеки сигналізує про необхідність прийняття чітких заходів на рівні держави щодо усунення цих питань або зменшення їх впливів.

Проаналізуємо структурну ефективність вихідних показників для країн 6-го кластеру (рисунок 1.75). Результати Output-oriented CCR-model показують, що максимальне зростання індексу конвергенції системи фінансового моніторингу та кібербезпеки можливе на 8,66%. Це можливо забезпечити за рахунок резервів за всіма показниками. Тобто для країн даного кластеру є можливості розвитку за рахунок підвищення ефективності від конвергенції систем протидії фінансовим злочинам та кібершахрайствам.

Питання підвищення ефективності системи національної безпеки у частині конвергенції систем фінансового моніторингу і кібербезпеки є досить актуальним, що пов’язано із зростанням рівня інформатизації, цифровізації та комп’ютеризації суспільства.

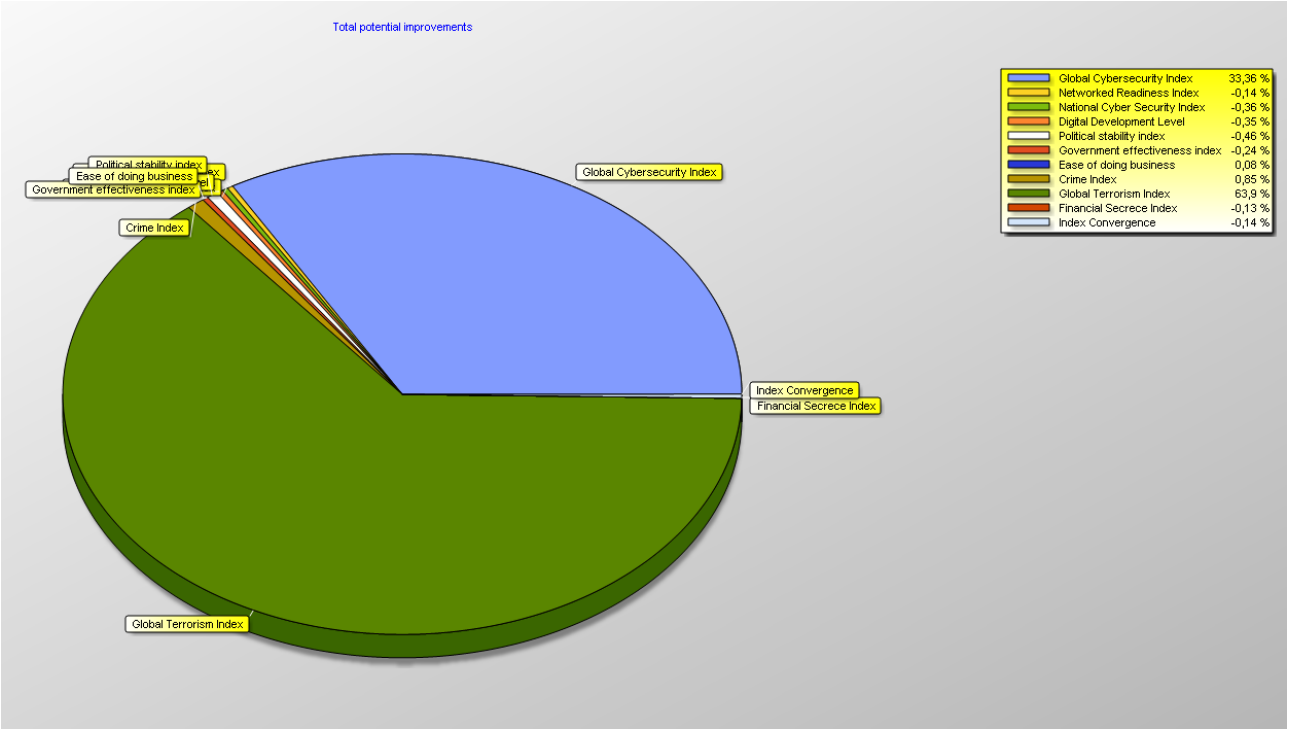


Рисунок 1.74 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 6-го кластеру (за Input-oriented CCR-моделлю)

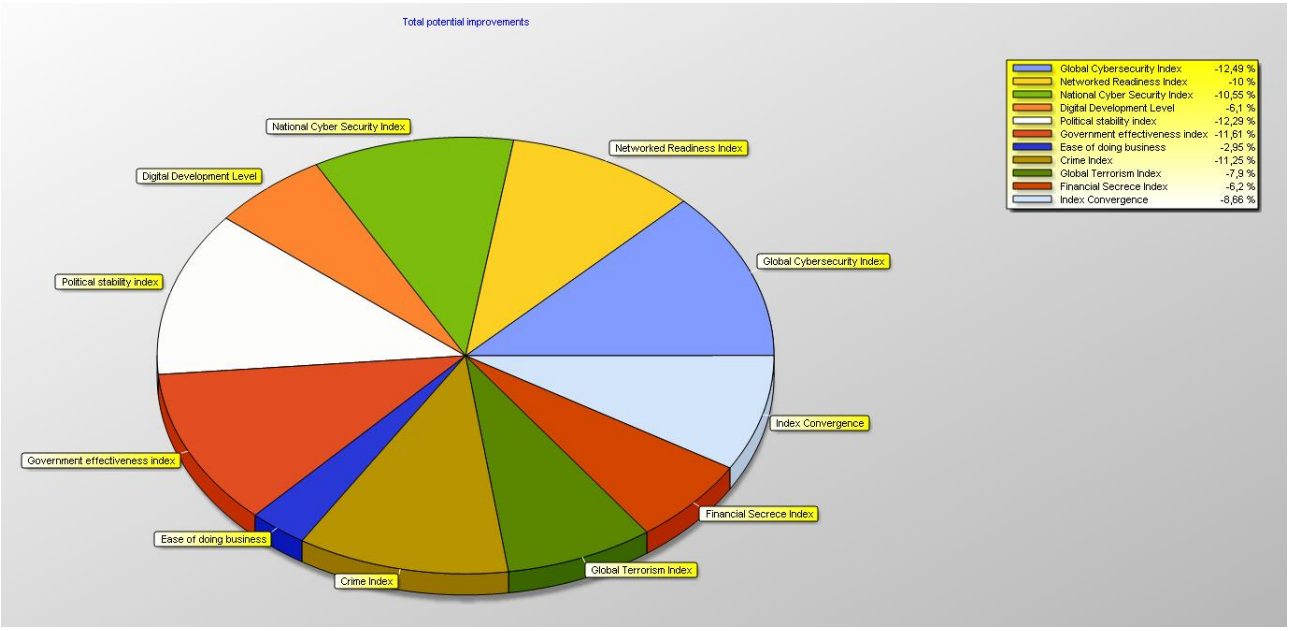


Рисунок 1.75 – Результати ефективності за складовими конвергенції системи фінансового моніторингу та кібербезпеки країн 6-го кластеру (за Output-oriented CCR-моделлю)

Застосування Data Envelopment Analysis у даному дослідженні дозволило визначити ефективність таких процесів. Використана модель CCR надала можливість проаналізувати структурну ефективність показників, що характеризують рівень розвитку системи кіберзахисту та протидії легалізації кримінальних доходів. Також дана модель дозволили оцінити максимальний рівень його зростання за наявного ресурсного потенціалу країни. Модель CCR є більш обмежуючою, ніж BCC, для визначення ефективності, що сприяло формуванню більш критичної оцінки щодо існуючих резервів країн, необхідних для забезпечення інтеграційних процесів. Саме тому вона була використана для проведення аналізу усіх кластерів країн.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**].

2 МОДЕРНІЗАЦІЯ ІНСТРУМЕНТАРІЮ ПРОТИДІЇ ЛЕГАЛІЗАЦІЇ КРИМІНАЛЬНИХ ДОХОДІВ ТА КІБЕРШАХРАЙСТВАМ

2.1 Структурний багатошаровий аналіз джерел кібератак

2.1.1 Базовий, статистичний, кореляційний, декомпозиційний аналіз часових рядів

Зростання рівня інформатизації та комп'ютеризації багатьох сфер життєдіяльності суспільства призвело до появи та розповсюдження такого явища як кіберзлочинність. Кіберзлочином вважається дія особи чи групи осіб, направлена на незаконне отримання персональних даних іншої фізичної особи, суб'єктів господарювання, державних органів, або порушення функціонування їх програмних та технічних засобів. Як правило, даний вид злочину здійснюються за допомогою комп'ютерних засобів та технологій.

Найбільш розповсюдженим видом кіберзлочину є кібератаки, які провадяться хакерами для досягнення економічних, політичних та соціальних цілей особи, груп осіб або держави. У 2020 році вони посідали п'яте місце у світі серед таких видів ризиків, як геополітичні, економічні, соціальні та навколишнього середовища, що робить їх досить серйозною проблемою для суспільства [**Ошибка! Источник ссылки не найден.**]. Серед кібератак виділяються такі види, як фішинг, DoS-атаки, розповсюдження шкідливого програмного забезпечення, Man-in-the-Middle, Zero-day exploit атаки, міжсайтовий скриптинг, логічні бомби, тощо. Їх головними характеристиками є непередбачуваність, стрімкість здійснення, масове охоплення об'єктів, висока ймовірність досягнення цілей, що робить їх швидкою та небезпечною зброєю в руках злочинців.

Результатом кібератак, як правило, є витік або втрата інформації. Так, у 2022 році найбільші втрати від кіберзлочинів відбулися у сфері охорони здоров'я (10,10 млн. дол. США), фінансовій індустрії (5,97 млн. дол. США), фармацевтичній галузі (5,01 млн. дол. США), технологічній сфері (4,97 млн. дол.

США), енергетиці (4,72 млн. дол. США) та інших [**Ошибка! Источник ссылки не найден.**]. Також прогнозується, що у 2025 році кіберзлочинність буде коштувати компаніям приблизно 10,5 трлн. дол. США, що перевищуватиме втрати у 3,5 рази в порівнянні з 2015 роком [**Ошибка! Источник ссылки не найден.**]. Також слід зазначити, що кількість кібератак невпинно зростає. Наприклад, кількість їх випадків в результаті пандемії COVID-19 зросла на 600% [**Ошибка! Источник ссылки не найден.**].

Таким чином, проблема кіберзлочинів в цілому та кібератак зокрема є досить актуальною, потребує пошуку різних інструментів і методів її дослідження та протидії. Для цього ефективними є не тільки технічні та програмні засоби але й управлінські інструменти, такі як прогнозування трендів. Процес прогнозування є складним і включає різні етапи реалізації, одним з яких є підбір даних, здійснення їх попереднього аналізу та підготовки до розробки ефективних прогнозних моделей. Реалізації даного етапу й буде присвячене це дослідження.

Питання виявлення та протидії кібератак є актуальним перед усім для науковців, які займаються питаннями кібербезпеки. Але сьогодні дана проблема набуває міждисциплінарного значення, оскільки її наслідки спостерігаються в економіці, бізнесі, суспільстві, політиці, охороні здоров’я та інше. Тому вчені з різних наукових шкіл та напрямків намагаються вирішувати її з різних точок зору.

Стейсі П., Тейлор Р., Спанакі К. досліджували психологічні аспекти впливу кібератак, а саме емоційні реакції персоналу компаній [**Ошибка! Источник ссылки не найден.**]. Шендлер Р. та Гомес М. А. виявили, що кібератаки є джерелом суспільного ризику, що проявляється у зростанні рівня суспільної недовіри до уряду у випадку кіберзагроз [**Ошибка! Источник ссылки не найден.**]. Лонсдейл Д. Дж. перевіряв кібератаки на предмет їх благ для суспільства, що визначалося з точки зору поваги до людини, соціального благополуччя, безпеки та миру, а також солідарності [**Ошибка! Источник ссылки не найден.**]. Болпагні М. запропонував зведений індекс для

вимірювання кіберризику та оцінив вплив соціо-економічних факторів на його зміни [**Ошибка! Источник ссылки не найден.**]. Сімонс Г., Даник Ю., Малярчук Т. намагалися вирішити дилему, породжену суперечністю правового регулювання із політичною та оперативною необхідністю управління ситуаціями, пов'язаними із кіберзагрозами [**Ошибка! Источник ссылки не найден.**].

Вівер Г. А., Феддерсен Б., Марла Л., Вей Д., Роуз А., Ван Моер М. вивчали економічні наслідки кібератак на прикладі морської транспортної системи та застосували оптимізаційний підхід для оцінки взаємодії між кібератаками та відповідними інформаційними технологіями компанії [**Ошибка! Источник ссылки не найден.**]. Лерой І. запропонувала застосовувати інструменти управління репутацією компанії для відновлення вартості її акцій після здійснення кібератак [**Ошибка! Источник ссылки не найден.**]. Акото У. дослідив позитивний вплив кібератак на торгівельні операції країни, що проявляється у використанні секретів, які добуваються в результаті кібершпигунства на користь держави [**Ошибка! Источник ссылки не найден.**]. Лаллі Г. С., Шеперд Л. А., Медсестра Дж. Р. К., Ерола А., Епіфаніу Г., Мейпл К., Беллекенс Х. проаналізували період пандемії COVID-19 з точки зору кібератак та виявили тенденцію їх щоденного зростання [**Ошибка! Источник ссылки не найден.**].

Не дивлячись на те, що проблема кібератак є досить актуальною та практично значущою, існує потреба у розробці прогнозних моделей, які дозволять виявляти потенційні кіберзагрози для певних країн та застосовувати контрзаходи щодо їх попередження.

Базою для розробки будь-якої прогнозної моделі є побудова її концептуальної моделі. Вона представляє собою зображення процесу моделювання, як сукупності етапів, починаючи з виявлення та аналізу вхідних даних, які відображають проблеми дійсності, та завершуючи розрахунком прогнозів за обраною моделлю та перевіркою їх якості. Для прогнозування

інформаційних трендів кібератак пропонуємо наступну концептуальну модель (Рисунок 2.1).

Представлена на рисунку 2.1 концептуальна модель прогнозування трендів кібератак передбачає виконання двох процесів – попереднього аналізу і підготовки даних та прогнозування. Перший процес є необхідним для досліджень подібного характеру, оскільки він дозволяє сформувати такий набір даних, від якості якого залежатимуть подальші дії щодо створення прогнозної моделі та отримання адекватних та точних прогнозів. Другий процес передбачає вибір математичних моделей, які відповідатимуть результатам, отриманим після попереднього аналізу та підготовки даних. Дане дослідження буде охоплювати результати здійснення першого процесу, передбаченого концептуальною моделлю. Другий процес висвітлюватиметься у наступному дослідженні.

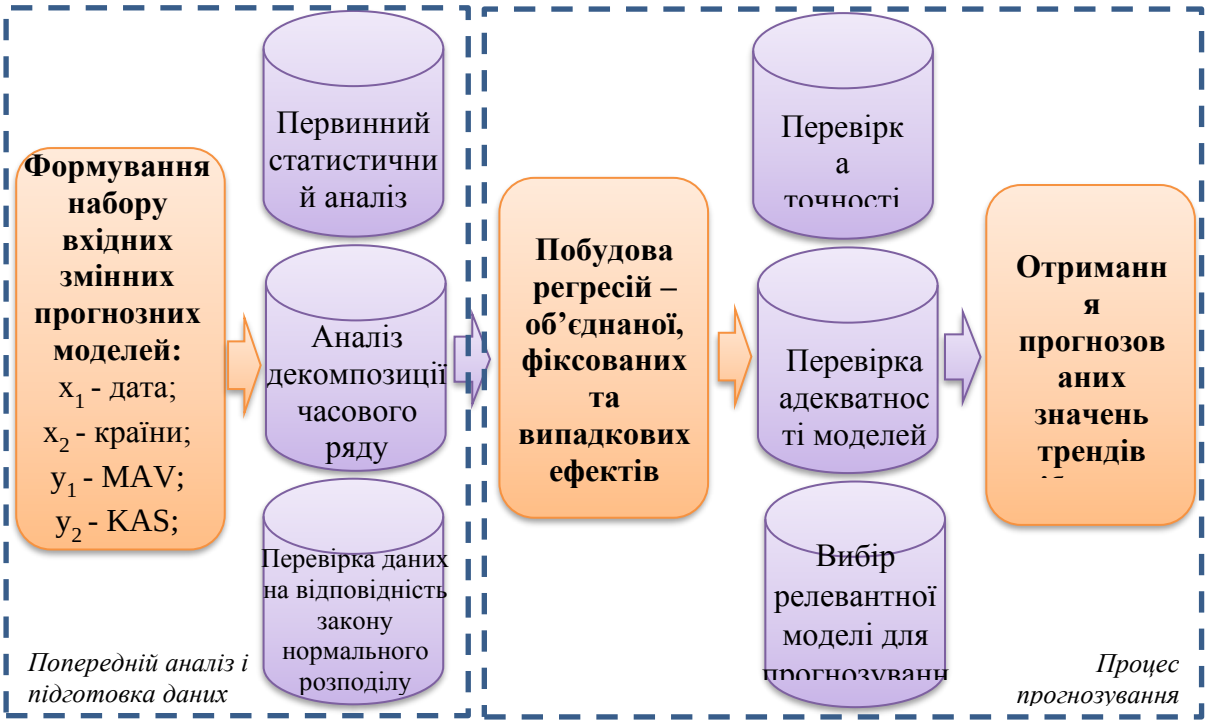


Рисунок 2.1 – Концептуальна модель прогнозування трендів кібератак

На першому етапі було сформовано набір змінних для розробки прогнозної моделі трендів кібератак. Вхідною інформацією було обрано статистичні дані 40 країн світу (по 10 країн з Європи, Азії, Африки та по 5 країн з Північної та Південної Америки) за період з 14 серпня 2022 року до 13 вересня 2022 року,

узятих з відкритого доступу Лабораторії Касперського. Вони представляють собою щоденну статистику про кількість кібератак, виявлених за допомогою спеціальних інструментів їх протидії, а саме:

- MAV (Mail Anti Virus) – поштовий антивірус, який показує потік даних шкідливих програм, виявлених серед нових об'єктів у поштових додатках. Він перевіряє вхідні повідомлення та запускає автоматичну перевірку при збереженні вкладених файлів на диск;
- KAS (Kaspersky Anti-Spam) – Касперський Анти-Спам, який показує підозрілий та небажаний поштовий трафік, виявлений за допомогою технологій репутаційної фільтрації «Лабораторії Касперського»;
- IDS (Intrusion Detection Scan) – система виявлення вторгнень, яка показує потік даних з виявлених мережевих атак.

Обрані дані є панельними, оскільки містять інформацію про одну і ту ж множину об’єктів за ряд послідовних періодів часу. Тобто маємо одні й ті самі дані щодо трьох видів кібератак для сорока країн за 30 днів. Відповідно, для кожного спостереження будуть вимірюватися декілька параметрів (регресійні змінних або ефектів) за кожен період часу. Оскільки в даному випадку всі показники відстежуються протягом однакової кількості періодів часу, то така панель є збалансованою.

На наступному кроці необхідно провести первинний аналіз початкових даних та здійснити відповідні маніпуляції для його підготовки до безпосередньої побудови прогнозної моделі. Розрахунки для даного дослідження проводилися із використанням мови програмування Python. Для цього було використано ряд стандартних бібліотек для аналізу, візуалізації і моделювання даних, таких як: Pandas, Numpy, Scipy.stats, Statsmodels, Matplotlib, Seaborn, Linearmodels та інші. Спочатку була проведена перевірка набору даних щодо наявності пропущених значень за допомогою функції isna(). Дана процедура необхідна для виявлення відсутніх даних, що робить вибірку неоднорідною. Результат її проведення показав, що набір не має пропущених даних і не потребує додаткових маніпуляцій по їх відновленню чи заміні.

Далі була проведена оцінка базових статистик, результати якої представлені на рисунку 2.2.

index	MAV	KAS	IDS
count	1240.0	1240.0	1240.0
mean	4647.270967741935	7617245.080645162	152111.00161290323
std	7700.3844928245235	20092268.916028455	251532.5601747635
min	1.0	3500.0	2.0
25%	286.0	140375.0	8927.0
50%	1630.5	764000.0	46237.0
75%	5174.0	4785625.0	213360.75
max	77612.0	181005000.0	2643943.0

Рисунок 2.2 – Результати розрахунку базових статистик для початкових даних

На рисунку 2.2 можна побачити, що набір даних складається з 1240 спостережень. Значення середньоквадратичного відхилення по всім трьом видам кібератак є дуже високим і значно перевищує середнє значення ряду, що говорить пре неоднорідність даних. Це пов’язано із тим, що деякі країни, які увійшли у вибірку, є більш атакованими, ніж інші. Також мінімальні та максимальні значення для всіх трьох видів кібератак мають суттєвий розкид – мінімальне є дуже маленьким числом, що свідчить про відсутність кібератак в даний момент часу для певної країни, а максимальне – дуже великим числом, що свідчить про значну активність кібератак в певному регіоні. У випадку MAV та IDS кібератак їх середні значення відповідають третьому квартилю, а у випадку KAS атаки – четвертому квартилю. Це свідчить про те, що кількість спостережень, відповідних найбільш активним фазам кібератак, складає приблизно 20-30% від загальної кількості. Тобто вони носять періодичний характер, який ймовірно залежить від часового періоду та від самої країни, на яку спрямована кібератака.

Оскільки панельні дані являються часовим рядом, необхідно дослідити їх декомпозицію та перевірити на відповідність нормальному розподілу. На рисунку 2.3 представлено декомпозицію інформаційних трендів кібератак, яка включає побудову графіків фактичних даних, трендової, сезонної та залишкової компонент.

Декомпозиції, представлені на рисунку 2.3, побудовано за адитивною моделлю, оскільки обрані тренди відповідають саме адитивному процесу. Це підтверджує випадковий розподіл їх залишків, які коливаються біля нуля. Візуальний аналіз трендової складової свідчить про її відсутність, але в даному випадку потрібні додаткові перевірки на стаціонарність, для чого використано тест Дики-Фулера. Графіки, які відповідають сезонним складовим вказують на можливість наявності даної компоненти в досліджуваних часових рядах.

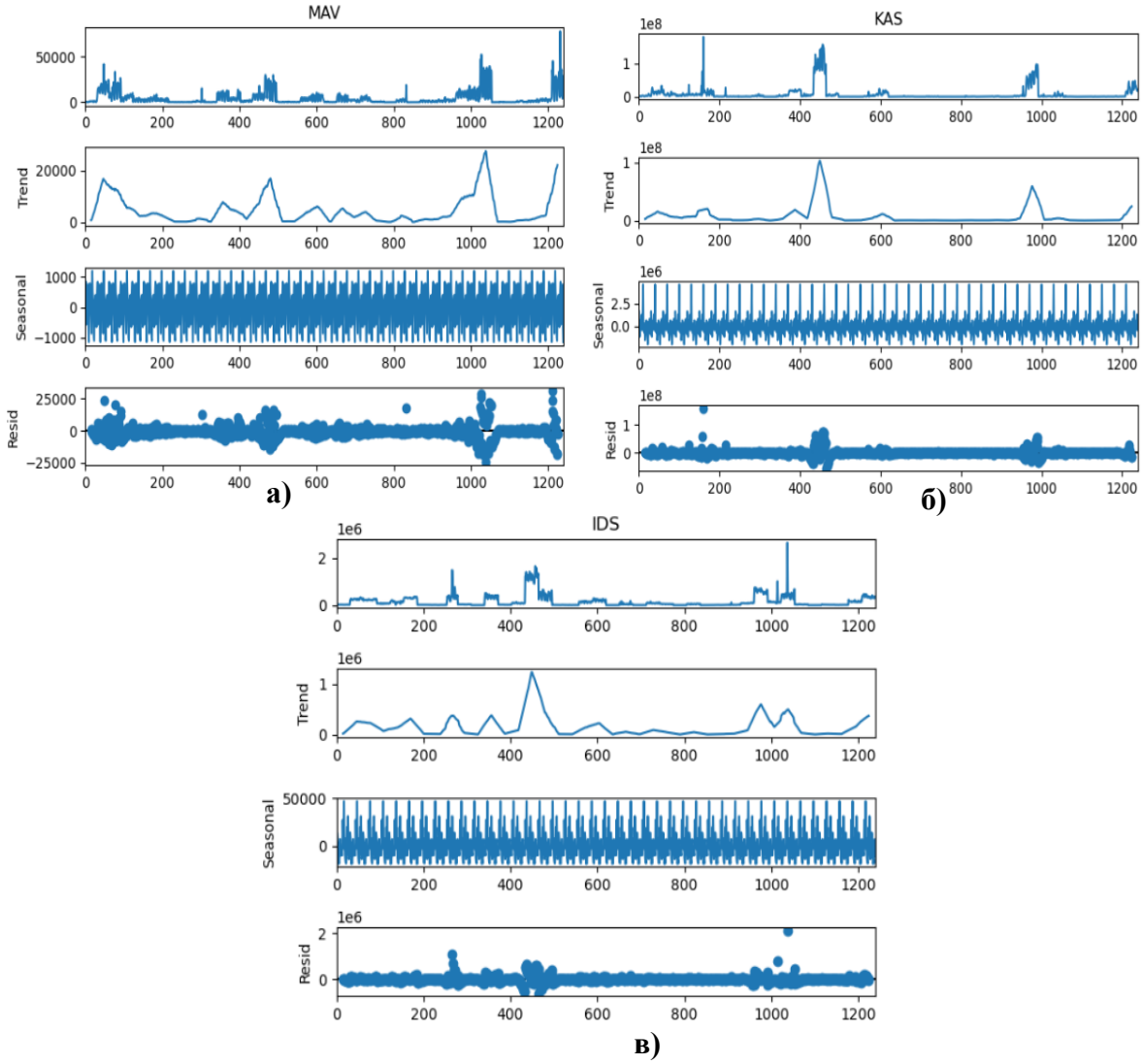


Рисунок 2.3 – Декомпозиція трендів для змінних: а) MAV; б) KAS; в) IDS

Результати перевірки досліджуваних рядів на стаціонарність представлені у таблиці 2.1.

Таблиця 2.1 – Результати тесту Дики-Фулера

Показники тесту	MAV	KAS	IDS
ADF	-7,1100	-36,6960	-36,0314
P-value	0,0000	0,0000	0,0000
Critical value 1%	-3,4357	-3,4356	-3,4356
Critical value 5%	-2,8639	-2,8639	-2,8639
Critical value 10%	2,5680	2,5680	2,5680
Висновок тесту	одиничних коренів немає, ряд є стаціонарним	одиничних коренів немає, ряд є стаціонарним	одиничних коренів немає, ряд є стаціонарним

Проведені тести перевірки рядів на стаціонарність показали, що вони є стаціонарними, тобто значення рядів не мають трендової складової. Для панельних даних у нашому випадку це означає, що у нас не буде виявлено ефекту хибної регресії, що дозволить будувати такі її різновиди, як об’єднана регресія, регресія з фіксованими та випадковими ефектами.

На наступному кроці проведемо перевірку часових рядів на нормальність, а саме їх відповідність нормальному розподілу. Для цього застосуємо два методи: метод побудови гістограм та обчислення тесту Харке-Бера. Результати проведеної процедури представлено на рисунку 2.4. Обидва методи показали, що вхідні дані не відповідають нормальному розподілу. Статистика тесту Харка-Бера завжди є позитивним числом, і якщо вона далека від нуля, а значення p-value менше 0,05, то це вказує на те, що вибіркові дані не відповідають нормальному розподілу. Також й візуальний аналіз графіків підтверджує даний висновок. Якщо гістограма має приблизно «дзвіноподібну форму», то дані вважаються нормально розподіленими. В нашому випадку змінні не мають такої форми, що свідчить про наявність асиметрії в даних і їх не відповідність нормальному розподілу.

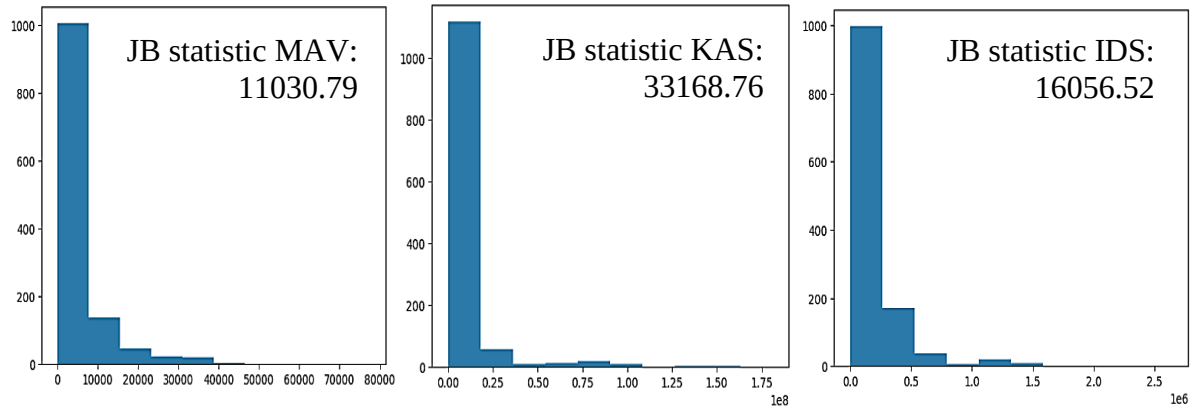


Рисунок 2.4 – Перевірка даних на відповідність закону нормального розподілу змінних

Для наближення даних до нормального розподілу можна виконати процедуру їх логарифмування, тобто здійснити перетворення незалежних змінних “x” із використанням log(x). Отримані трансформовані дані візуалізовані на рисунку 2.5.

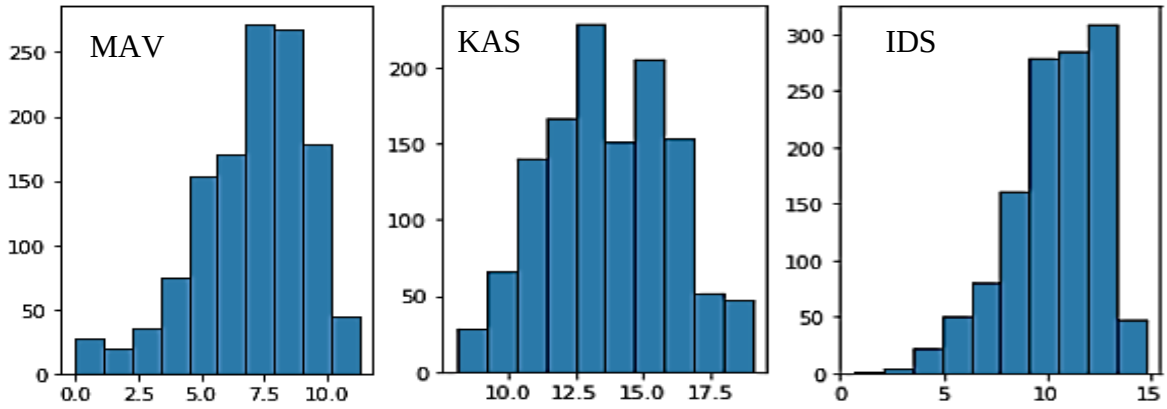


Рисунок 2.5 – Результат трансформування незалежних змінних “x”

Хоча трансформація даних й не призвела до повній відповідності даних нормальному закону, але отримані розподіли, зображені на рисунку 2.5, є досить близькими, що, в принципі, дозволяє їх використання для побудови прогнозних моделей.

Дане дослідження присвячене проблематиці кібератак, кількість випадків яких зросла за останні роки. Їх наслідки є катастрофічними для бізнесу, фізичних осіб та держав в цілому. Саме тому виникає потреба у використанні інструментів

щодо їх попередження та протидії, в якості яких можуть виступати моделі прогнозування. Для їх реалізації важливим етапом є аналіз та підготовка вхідних даних, що було проведено у даному досліджені. В якості бази емпіричних даних виступили три види часових трендів кібератак, які відслідковувалися за допомогою поштового антивірусу, Касперського Анти-Спаму та системи виявлення вторгнень.

У дослідженні було запропоновано концептуальну модель розробки прогнозних моделей кібератак, яка показує всі етапи процесу прогнозування. Розраховані базові статистики дозволили виявити неоднорідність даних. Встановлено, що це пов’язано із різним рівнем економічного розвитку країн, які було обрано для аналізу. Відповідно, деякі з них в більшій мірі ставали об’єктами кіберзагроз, інші – в меншій мірі. Проведена декомпозиція трендів дозволила виявити, що дані не містять трендової складової, мають сезонність та зв'язок між змінними є адитивним. Перевірка на стаціонарність за допомогою розширеного тесту Дики-Фулера встановила, що аналізовані тренди є стаціонарними, тобто був підтверджений попередній висновок щодо відсутності трендової складової. Оскільки дані характеризуються нерівномірністю, то проведена перевірка на відповідність нормальному розподілу за допомогою тесту Харка-Бера підтвердила, що їх невідповідність. Для їх наближення до умов нормального розподілу було проведено трансформацію змінних “x” шляхом логарифмування.

Таким чином, проведені в статті процедури підготовки даних дозволяють побудувати прогнозні моделі, такі як об’єднану регресію, регресію з випадковим та фіксованим ефектами. Дані побудові буде реалізовано у подальших дослідженнях.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**].

2.1.2 Побудова регресійних моделей для змінної «Mail Anti Virus»

Регресійна модель Pooled OLS часто є хорошою відправною точкою та еталонною моделлю для кількох наборів панельних даних. Для побудови використовуємо `OLS` клас statsmodels для побудови та адаптації регресійної моделі OLS [**Ошибка! Источник ссылки не найден.**].

Для початку необхідно визначити залежну та незалежні змінні. Залежна змінна – це MAV, яка показує потік даних за шкідливими програмами, виявленими серед нових об'єктів у поштових додатках, незалежними змінними Data_num – показує порядок днів у місяці, так як Python не розуміє типу даних Дата, та перетворює їх в числа з 0 до кінцевого значення по порядку.

Також необхідно створити dummy змінні, які будуть виступати незалежними змінними, тобто кожна країна – це окрема булева змінна та приєднуємо їх до основної бази даних (рис. 2.6).

```
df_dummies = pd.get_dummies(df[unit_col_name])
df_panel_with_dummies = df.join(df_dummies)
df_panel_with_dummies
```

	Date	Country	Date_num	Country_Id	MAV	KAS	IDS	Afghanistan	Armenia	Azerbaijan	...	Togo	Tynisia	UK	Uganda	Ukraine	Uni Sta
0	2022-08-14	Ukraine	1	1	4.795791	13.978904	9.912150	0	0	0	...	0	0	0	0	1	
1	2022-08-15	Ukraine	2	1	5.379897	14.321123	9.913190	0	0	0	...	0	0	0	0	1	
2	2022-08-16	Ukraine	3	1	5.493061	14.300403	9.952897	0	0	0	...	0	0	0	0	1	
3	2022-08-17	Ukraine	4	1	6.345636	14.399958	9.977481	0	0	0	...	0	0	0	0	1	
4	2022-08-18	Ukraine	5	1	6.077642	14.313859	9.906084	0	0	0	...	0	0	0	0	1	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
1235	2022-09-09	Brazil	27	40	10.477851	17.168917	12.671064	0	0	0	...	0	0	0	0	0	
1236	2022-09-10	Brazil	28	40	8.645059	16.921710	12.578605	0	0	0	...	0	0	0	0	0	
1237	2022-09-11	Brazil	29	40	8.308966	16.747717	12.606053	0	0	0	...	0	0	0	0	0	
1238	2022-09-12	Brazil	30	40	10.256290	16.911541	12.725729	0	0	0	...	0	0	0	0	0	
1239	2022-09-13	Brazil	31	40	10.105653	16.701400	12.790911	0	0	0	...	0	0	0	0	0	

1240 rows × 47 columns

Рисунок 2.6 – Створення dummy змінних

Визначаємо залежну та незалежні змінні (рис. 2.7).

```
y = 'MAV'
x = df_panel_with_dummies.drop(['MAV', 'KAS', 'IDS', 'Date', 'Country'], axis=1)
```

Рисунок 2.7 – Вхідні змінні моделі

Будуємо об’єднану модель для змінної MAV (рис. 2.8).

Після побудови об’єднаної регресійної моделі отримуємо наступне рівняння регресії (2.1):

$$MAV = 6.77 + 0.0038 \cdot Data_{num} - 4.83 \cdot Afganistan + 5.11 \cdot Armenia + 5.22 \cdot Azerbaijan \ldots + Zandia \cdot 5.05 + 7.09 \cdot Zimbabwe + \epsilon \quad (2.1)$$

OLS Regression Results

Dep. Variable:	MAV	R-squared:	0.776
Model:	OLS	Adj. R-squared:	0.769
Method:	Least Squares	F-statistic:	104.0
Date:	Sun, 27 Nov 2022	Prob (F-statistic):	0.00
Time:	12:52:01	Log-Likelihood:	-1785.7
No. Observations:	1240	AIC:	3653.
Df Residuals:	1199	BIC:	3863.
Df Model:	40		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	6.7728	0.059	114.842	0.000	6.657	6.888
Date_num	0.0038	0.003	1.147	0.251	-0.003	0.010
Afghanistan	-1.9391	0.184	-10.527	0.000	-2.301	-1.578
Armenia	-1.6596	0.184	-9.010	0.000	-2.021	-1.298
Azerbaijan	-1.5556	0.184	-8.445	0.000	-1.917	-1.194
Brazil	2.7729	0.184	15.053	0.000	2.412	3.134
Canada	0.7021	0.184	3.812	0.000	0.341	1.064
Chile	0.8461	0.184	4.593	0.000	0.485	1.207
China	1.8793	0.184	10.202	0.000	1.518	2.241
Colombia	2.1974	0.184	11.929	0.000	1.836	2.559
Costa Rica	-0.7608	0.184	-4.130	0.000	-1.122	-0.399
Cuba	-2.4545	0.184	-13.325	0.000	-2.816	-2.093
Czech Republic	-0.2331	0.184	-1.266	0.206	-0.595	0.128
Egypt	1.2235	0.184	6.642	0.000	0.862	1.585
France	0.9883	0.184	5.365	0.000	0.627	1.350
Germany	2.8093	0.184	15.251	0.000	2.448	3.171
Hungary	0.7274	0.184	3.949	0.000	0.366	1.089
India	1.5464	0.184	8.395	0.000	1.185	1.908
Indonezia	1.8612	0.184	10.104	0.000	1.500	2.223
Iran	1.2874	0.184	6.989	0.000	0.926	1.649
Italy	2.3567	0.184	12.794	0.000	1.995	2.718
Japan	1.4038	0.184	7.620	0.000	1.042	1.765
Kenya	1.2794	0.184	6.945	0.000	0.918	1.641
Mexico	3.1120	0.184	16.894	0.000	2.751	3.473
Moldova	-2.0779	0.184	-11.280	0.000	-2.439	-1.716
Paraguay	-0.1887	0.184	-1.025	0.306	-0.550	0.173
Poland	0.6754	0.184	3.667	0.000	0.314	1.037
Slovakia	-2.0631	0.184	-11.200	0.000	-2.424	-1.702
Somalia	-5.6288	0.184	-30.557	0.000	-5.990	-5.267
South Korea	0.0173	0.184	0.094	0.925	-0.344	0.379
Sudan	-0.9073	0.184	-4.925	0.000	-1.269	-0.546
Tanzania	-0.1039	0.184	-0.564	0.573	-0.465	0.257
Togo	-3.0726	0.184	-16.680	0.000	-3.434	-2.711
Tynisia	0.2607	0.184	1.415	0.157	-0.101	0.622
UK	1.6226	0.184	8.808	0.000	1.261	1.984
Uganda	-0.3444	0.184	-1.870	0.062	-0.706	0.017
Ukraine	-0.5675	0.184	-3.081	0.002	-0.929	-0.206
United States	2.2485	0.184	12.206	0.000	1.887	2.610
Venezuala	-2.6326	0.184	-14.291	0.000	-2.994	-2.271
Vietnam	2.5526	0.184	13.857	0.000	2.191	2.914
Zamdia	-1.7266	0.184	-9.373	0.000	-2.088	-1.365
Zimbabwe	0.3189	0.184	1.731	0.084	-0.042	0.680

Omnibus:	245.994	Durbin-Watson:	1.499
Prob(Omnibus):	0.000	Jarque-Bera (JB):	441.868
Skew:	-1.216	Prob(JB):	1.12e-96
Kurtosis:	4.623	Cond. No.	1.72e+16

Рисунок 2.8 – Результати побудови об’єднаної моделі для змінної MAV

Щоб проаналізувати, чи є об’єднана модель OLS адекватною моделлю для нашої проблеми регресії, необхідно провести аналіз таких показників **R-квадрат** і **F-тест**, логарифм правдоподібності та балів **AIC**, а також опосередковано через аналіз залишків.

Скоригований R-квадрат, який вимірює частку загальної дисперсії в *y*, яка пояснюється *X після врахування ступенів свободи, втрачених через включення змінних регресії*, становить 0.769 або близько 76.9 %. Це, безумовно гарний результат.

F - **тест для регресії**, який вимірює спільну значущість параметрів моделі, дав тестову статистику 104.0 із значенням p = 0.00, що дозволяє зробити висновок, що оцінки коефіцієнтів моделі є спільно значущими при p < 0.001 [**Ошибка! Источник ссылки не найден.**].

Log-правдоподібність регресійної моделі - це спосіб вимірювання користі придатності для моделі. Чим вище значення лог-ймовірності, тим краще модель підходить для набору даних [**Ошибка! Источник ссылки не найден.**]. Log-правдоподібність моделі становить 1785.7, а показник **AIC** 3653. Ці значення придатності самі по собі не мають сенсу, якщо ми не порівняємо їх із показниками конкуруючої моделі [**Ошибка! Источник ссылки не найден.**].

Проаналізуємо **залишкові похибки** моделі для **нормальності**, **гетероскедастичності** та **кореляції** - трьох властивостей, які впливають на відповідність лінійної моделі.

Залишкова стандартна похибка - це міра, яка використовується для оцінки того, наскільки добре модель лінійної регресії відповідає даним [**Ошибка! Источник ссылки не найден.**]. Її результат представлений на рисунках 2.9-2.10.

```
print(pooled_olsr_model_results.resid)
```

```
0      -1.413282
1      -0.832959
2      -0.723579
3       0.125212
4      -0.146566
          ...
1235    0.829994
1236   -1.006582
1237   -1.348459
1238    0.597081
1239    0.442660
Length: 1240, dtype: float64
```

Рисунок 2.9 – Залишкові похибки моделі

```
print('Mean value of residual errors='+str(pooled_olsr_model_results.resid.mean()))
```

```
Mean value of residual errors=5.80431760003223e-15
```

Рисунок 2.10 – Середні значення похибок моделі

Це говорить нам про те, що регресійна модель прогнозує MAV із середньою похибкою близько 5.80 e-15.

Для побудови моделі регресії фіксованих ефектів, необхідно створити фіктивні змінні (рис. 2.11).

```
unit_col_name= 'Country'
time_period_col_name='Date'
```

Рисунок 2.11 – Створення фіктивних змінних

Здійснюємо побудову регресії (рис. 2.12-2.13).

```
unit_names = ['Germany','Italy','UK' ,'Poland','France' ,'Hungary' ,'Moldova', 'Slovakia', 'Afghanistan','Indonezia','Japan','Chi
```

```
lsdv_expr = y_var_name + ' ~ '
i = 0
for X_var_name in X_var_names:
    if i > 0:
        lsdv_expr = lsdv_expr + ' + ' + X_var_name
    else:
        lsdv_expr = lsdv_expr + X_var_name
    i = i + 1
for dummy_name in unit_names[:-1]:
    lsdv_expr = lsdv_expr + ' + ' + dummy_name

print('Regression expression for OLS with dummies=' + lsdv_expr)
```

```
Regression expression for OLS with dummies=MAV ~ Date_num + Germany + Italy + UK + Poland + France + Hungary + Moldova + Slovak
ia + Afghanistan + Indonezia + Japan + China + Vietnam + Armenia + Azerbaijan + Iran + India + Zamdia + Kenya + Zimbabwe + Egyp
t + Sudan + Somalia + Tynisia + Togo + Uganda + Tanzania + Canada + Colombia + Mexico + Cuba + Paraguay + Chile + Brazil
```

Рисунок 2.12 – Побудова рівняння регресії

OLS Regression Results

Dep. Variable:	MAV	R-squared:	0.710
Model:	OLS	Adj. R-squared:	0.702
Method:	Least Squares	F-statistic:	84.29
Date:	Fri, 25 Nov 2022	Prob (F-statistic):	6.31e-295
Time:	13:51:22	Log-Likelihood:	-1946.1
No. Observations:	1240	AIC:	3964.
Df Residuals:	1204	BIC:	4149.
Df Model:	35		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	6.4514	0.105	61.309	0.000	6.245	6.658
Date_num	0.0038	0.004	1.010	0.313	-0.004	0.011
Germany	3.1307	0.229	13.680	0.000	2.682	3.580
Italy	2.6781	0.229	11.702	0.000	2.229	3.127
UK	1.9439	0.229	8.494	0.000	1.495	2.393
Poland	0.9968	0.229	4.356	0.000	0.548	1.446
France	1.3097	0.229	5.723	0.000	0.861	1.759
Hungary	1.0487	0.229	4.583	0.000	0.600	1.498
Moldova	-1.7565	0.229	-7.675	0.000	-2.206	-1.308
Slovakia	-1.7417	0.229	-7.611	0.000	-2.191	-1.293
Afghanistan	-1.6177	0.229	-7.069	0.000	-2.067	-1.169
Indonezia	2.1826	0.229	9.537	0.000	1.734	2.632
Japan	1.7251	0.229	7.538	0.000	1.276	2.174
China	2.2007	0.229	9.616	0.000	1.752	2.650
Vietnam	2.8740	0.229	12.558	0.000	2.425	3.323
Armenia	-1.3383	0.229	-5.848	0.000	-1.787	-0.889
Azerbaijan	-1.2342	0.229	-5.393	0.000	-1.683	-0.785
Iran	1.6088	0.229	7.030	0.000	1.160	2.058
India	1.8678	0.229	8.162	0.000	1.419	2.317
Zandia	-1.4052	0.229	-6.140	0.000	-1.854	-0.956
Kenya	1.6007	0.229	6.995	0.000	1.152	2.050
Zimbabwe	0.6403	0.229	2.798	0.005	0.191	1.089
Egypt	1.5449	0.229	6.751	0.000	1.096	1.994
Sudan	-0.5859	0.229	-2.560	0.011	-1.035	-0.137
Somalia	-5.3074	0.229	-23.192	0.000	-5.756	-4.858
Tynisia	0.5821	0.229	2.544	0.011	0.133	1.031
Togo	-2.7512	0.229	-12.022	0.000	-3.200	-2.302
Uganda	-0.0231	0.229	-0.101	0.920	-0.472	0.426
Tanzania	0.2175	0.229	0.950	0.342	-0.232	0.666
Canada	1.0235	0.229	4.472	0.000	0.575	1.472
Colombia	2.5187	0.229	11.006	0.000	2.070	2.968
Mexico	3.4334	0.229	15.003	0.000	2.984	3.882
Cuba	-2.1331	0.229	-9.321	0.000	-2.582	-1.684
Paraguay	0.1327	0.229	0.580	0.562	-0.316	0.582
Chile	1.1674	0.229	5.101	0.000	0.718	1.616
Brazil	3.0943	0.229	13.521	0.000	2.645	3.543

Omnibus:	246.665	Durbin-Watson:	1.173
Prob(Omnibus):	0.000	Jarque-Bera (JB):	603.618
Skew:	-1.067	Prob(JB):	8.43e-132
Kurtosis:	5.670	Cond. No.	303.

Рисунок 2.13 – Регресія фіксованих змінних для змінної MAV

Рівняння регресії з фіксованими змінними виглядає наступним чином (2.2):

$$\begin{aligned} \text{MAV} = 6.45 \ &+ 0.0038 \cdot \text{Data}_{\text{num}} - 3.13 \cdot \text{Germany} + 2.68 \cdot \text{Italy} + \cdots \\ &+ 2.17 \cdot \text{Cuba} + 1.18 \cdot \text{Chile} + 3.09 \cdot \text{Brazil} + \ \epsilon \end{aligned} \tag{2.2}$$

Скоригований R-квадрат**,** який вимірює частку загальної дисперсії в *y*, яка пояснюється X після врахування ступенів свободи, втрачених через включення

змінних регресії, становить 0.702 або близько 70.2 %. Це, безумовно гарний результат.

F - тест для регресії, який вимірює спільну значущість параметрів моделі, дав тестову статистику 84.29 із значенням p = 0.00, що дозволяє зробити висновок, що оцінки коефіцієнтів моделі є спільно значущими при $p < 0.001$.

Log-правдоподібність моделі становить 1946.1, а показник AIC 3964. Ці значення придатності самі по собі не мають сенсу, якщо ми не порівняємо їх із показниками конкуруючої моделі.

Проаналізуємо залишкові похибки підігнаної моделі для нормальності, гетероскедастичності та кореляції - трьох властивостей, які впливають на відповідність лінійної моделі [**Ошибка! Источник ссылки не найден.**] (рис. 2.14).

```
print(lsdv_model_results.resid)
print('Mean value of residual errors='+str(lsdv_model_results.resid.mean()))
```

```
0      -1.659387
1      -1.079064
2      -0.969684
3      -0.120892
4      -0.392670
          ...
1235    0.829994
1236   -1.006582
1237   -1.348459
1238    0.597081
1239    0.442660
Length: 1240, dtype: float64
Mean value of residual errors=4.410210548010421e-13
```

Рисунок 2.14 - Залишкові похибки моделі

Першим кроком для побудови моделі випадкових ефектів, необхідно розрахувати σ^2_ϵ і σ^2_μ - дисперсії компонентів похибки μ і ϵмоделі фіксованих ефектів та об’єднаної моделі та знайти різницю між ними (рис. 2.15).

```
sigma2_epsilon = lsdv_model_results.ssr/(n*T-(n+k+1))
print('sigma2_epsilon = ' + str(sigma2_epsilon))
```

```
sigma2_epsilon = -45.283594237287836
```

```
sigma2_pooled = pooled_olsr_model_results.ssr/(n*T-(k+1))
print('sigma2_pooled = ' + str(sigma2_pooled))
```

```
sigma2_pooled = -646.7554791838006
```

```
sigma2_u = sigma2_pooled - sigma2_epsilon
print('sigma2_u = ' + str(sigma2_u))
```

```
sigma2_u = -601.4718849465128
```

Рисунок 2.15 – Розрахунок значень дисперсії компонентів похибки моделей

Обчислюємо середні значення *y* та *X* для кожної групи (тобто кожної одиниці *i*) на панелі даних таким чином (рис. 2.16).

```
df_group_means = df_panel_with_dummies.groupby(unit_col_name).mean()

print(df_group_means)
```

Country	Date_num	Country_id	MAV	KAS	IDS
Afghanistan	16.0	11.0	4.894210	9.706546	7.178635
Armenia	16.0	17.0	5.173679	11.371538	8.767991
Azerbaijan	16.0	18.0	5.277701	11.792697	8.177036
Brazil	16.0	40.0	9.606234	16.871789	12.799997
Canada	16.0	31.0	7.535446	14.592936	11.297106
Chile	16.0	39.0	7.679380	13.193566	11.915792
China	16.0	15.0	8.712614	18.418690	14.017562
Colombia	16.0	33.0	9.030679	13.974321	11.785556
Costa Rica	16.0	37.0	6.072505	11.769029	9.744092
Cuba	16.0	36.0	4.378815	9.496146	6.163067
Czech Republic	16.0	10.0	6.600173	14.659354	10.278660
Egypt	16.0	24.0	8.056791	12.712312	11.350442
France	16.0	6.0	7.821634	16.351361	12.645950
Germany	16.0	2.0	9.642643	16.479932	12.442787
Hungary	16.0	7.0	7.560679	13.582649	9.446409
India	16.0	20.0	8.379729	16.105214	12.285693
Indonezia	16.0	12.0	8.694488	14.663417	12.816168

Рисунок 2.16 – Обчислення середніх значень

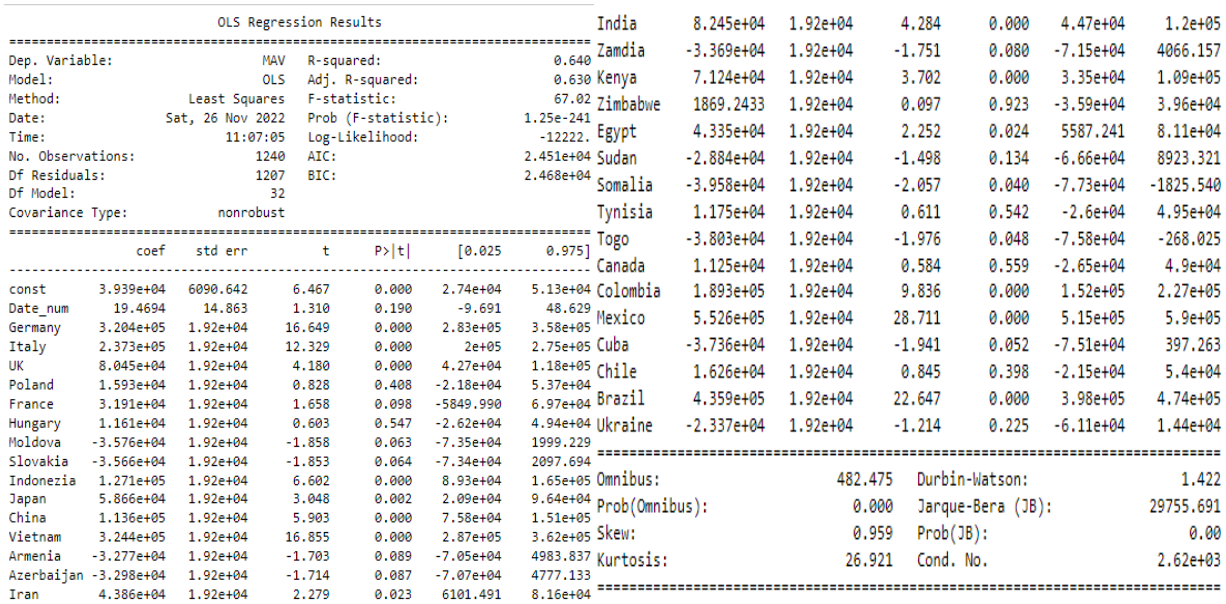
Наступним кроком є зменшення середніх значень всіх значень *y* та *X* для кожної одиниці, використовуючи масштабовану версію відповідного середнього значення для конкретної групи, обчислених на другому кроці (рис. 2.17).

```
theta = 1 - math.sqrt(c/(u))
print('theta = ' + str(theta))
```

```
theta = 0.9539504917492148
```

Рисунок 2.17 – Обчислення показника Тета

Тепер будуємо модель випадкових ефектів (рис. 2.18):



OLS Regression Results

Dep. Variable:	MAV	R-squared:	0.640
Model:	OLS	Adj. R-squared:	0.630
Method:	Least Squares	F-statistic:	67.02
Date:	Sat, 26 Nov 2022	Prob (F-statistic):	1.25e-241
Time:	11:07:05	Log-Likelihood:	-12222.
No. Observations:	1240	AIC:	2.451e+04
Df Residuals:	1207	BIC:	2.468e+04
Df Model:	32		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	3.939e+04	6090.642	6.467	0.000	2.74e+04	5.13e+04
Date_num	19.4694	14.863	1.310	0.190	-9.691	48.629
Germany	3.204e+05	1.92e+04	16.649	0.000	2.83e+05	3.58e+05
Italy	2.373e+05	1.92e+04	12.329	0.000	2e+05	2.75e+05
UK	8.045e+04	1.92e+04	4.180	0.000	4.27e+04	1.18e+05
Poland	1.593e+04	1.92e+04	0.828	0.408	-2.18e+04	5.37e+04
France	3.191e+04	1.92e+04	1.658	0.098	-5849.990	6.97e+04
Hungary	1.161e+04	1.92e+04	0.603	0.547	-2.62e+04	4.94e+04
Moldova	-3.576e+04	1.92e+04	-1.858	0.063	-7.35e+04	1999.229
Slovakia	-3.566e+04	1.92e+04	-1.853	0.064	-7.34e+04	2097.694
Indonezia	1.271e+05	1.92e+04	6.602	0.000	8.93e+04	1.65e+05
Japan	5.866e+04	1.92e+04	3.048	0.002	2.09e+04	9.64e+04
China	1.136e+05	1.92e+04	5.903	0.000	7.58e+04	1.51e+05
Vietnam	3.244e+05	1.92e+04	16.855	0.000	2.87e+05	3.62e+05
Armenia	-3.277e+04	1.92e+04	-1.703	0.089	-7.05e+04	4983.837
Azerbaijan	-3.298e+04	1.92e+04	-1.714	0.087	-7.07e+04	4777.133
Iran	4.386e+04	1.92e+04	2.279	0.023	6101.491	8.16e+04
India	8.245e+04	1.92e+04	4.284	0.000	4.47e+04	1.2e+05
Zamdia	-3.369e+04	1.92e+04	-1.751	0.080	-7.15e+04	4066.157
Kenya	7.124e+04	1.92e+04	3.702	0.000	3.35e+04	1.09e+05
Zimbabwe	1869.2433	1.92e+04	0.097	0.923	-3.59e+04	3.96e+04
Egypt	4.335e+04	1.92e+04	2.252	0.024	5587.241	8.11e+04
Sudan	-2.884e+04	1.92e+04	-1.498	0.134	-6.66e+04	8923.321
Somalia	-3.958e+04	1.92e+04	-2.057	0.040	-7.73e+04	-1825.540
Tynisia	1.175e+04	1.92e+04	0.611	0.542	-2.6e+04	4.95e+04
Togo	-3.803e+04	1.92e+04	-1.976	0.048	-7.58e+04	-268.025
Canada	1.125e+04	1.92e+04	0.584	0.559	-2.65e+04	4.9e+04
Colombia	1.893e+05	1.92e+04	9.836	0.000	1.52e+05	2.27e+05
Mexico	5.526e+05	1.92e+04	28.711	0.000	5.15e+05	5.9e+05
Cuba	-3.736e+04	1.92e+04	-1.941	0.052	-7.51e+04	397.263
Chile	1.626e+04	1.92e+04	0.845	0.398	-2.15e+04	5.4e+04
Brazil	4.359e+05	1.92e+04	22.647	0.000	3.98e+05	4.74e+05
Ukraine	-2.337e+04	1.92e+04	-1.214	0.225	-6.11e+04	1.44e+04

Omnibus:	482.475	Durbin-Watson:	1.422
Prob(Omnibus):	0.000	Jarque-Bera (JB):	29755.691
Skew:	0.959	Prob(JB):	0.00
Kurtosis:	26.921	Cond. No.	2.62e+03

Рисунок 2.18 – Побудова моделі регресії випадкових ефектів

Отримана модель матиме вигляд формули (2.3):

$$\begin{aligned} \text{MAV} = 39390 + 19.47 \cdot \text{Data}_{\text{num}} - 3.204\text{e}+04 \cdot \text{Germany} + 2.373\text{t} \\ + 05 \cdot \text{Italy} + \cdots + 4.359\text{e}+05 \cdot \text{Brazil} - 2.337\text{e}+04 \\ \cdot \text{Ukraine} + \epsilon \end{aligned} \qquad (2.3)$$

Обчислена дисперсія σ²_u було оцінено як -601.47, а σ²_ϵ було оцінено як -45.28. Таким чином, частка загальної дисперсії, яка може бути віднесена до випадкового ефекту окремої одиниці, дорівнює (2.4):

$$\frac{-601.47}{-601.47+(-45.28)} = 0{,}93 \qquad (2.4)$$

Це означає, що на 93% присутній випадковий ефект у моделі, але дивлячись на показник скоригованого R, який дорівнює 0.63, можна зробити висновок, що ця модель не є кращою за модель фіксованого ефекту та об’єднану модель.

Ми можемо використовувати тест Breusch-Pagan LM для перевірки значущості випадкового ефекту [**Ошибка! Источник ссылки не найден.**].

Нульова гіпотеза тесту Бреуша-Пагана LM полягає в тому, що одинична дисперсія σ^2_u дорівнює нулю (рис. 2.19).

```
df_pooled_olsr_resid_with_unitnames = pd.concat([df_data[unit_col_name],pooled_olsr_model_results.resid], axis=1)
df_pooled_olsr_resid_group_means = df_pooled_olsr_resid_with_unitnames.groupby(unit_col_name).mean()
ssr_grouped_means=(df_pooled_olsr_resid_group_means[0]**2).sum()
ssr_pooled_olsr=pooled_olsr_model_results.ssr
LM_statistic = (n*T)/(2*(T-1))*math.pow(((T*T*ssr_grouped_means)/ssr_pooled_olsr - 1),2)
print('BP LM Statistic='+str(LM_statistic))
alpha=0.05
chi2_critical_value=st.chi2.ppf((1.0-alpha), 1)
print('chi2_critical_value='+str(chi2_critical_value))
```

```
BP LM Statistic=18.008298755186722
chi2_critical_value=3.841458820694124
```

Рисунок 2.19 -Перевірка значущості моделі випадкового ефекту

Тестова статистика тесту LM (18,0083) більша, ніж критичне значення Chi-squared = 3,84146 при α=0,05, що означає, що випадковий ефект є значущим при альфа 0,05.

Необхідно визначити залежну та незалежні змінні. Залежна змінна – це KAS, яка показує потік даних з виявлених мережевих атак. Визначаємо вхідні дані моделі (рис. 2.20):

```
y = 'KAS'
x = df_panel_with_dummies.drop(['MAV', 'KAS', 'IDS', 'Date', 'Country'], axis=1)
```

Рисунок 2.20 – Вхідні змінні моделі

Статистично незначущі фактори було усунуто з моделі та модель побудували знову (рис. 2.21).

OLS Regression Results

Dep. Variable:	KAS	R-squared:	0.923
Model:	OLS	Adj. R-squared:	0.920
Method:	Least Squares	F-statistic:	357.2
Date:	Sat, 26 Nov 2022	Prob (F-statistic):	0.00
Time:	13:38:14	Log-Likelihood:	-1211.8
No. Observations:	1240	AIC:	2586.
Df Residuals:	1199	BIC:	2716.
Df Model:	40		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	13.0191	0.037	350.683	0.000	12.946	13.092
Date_num	0.0203	0.002	9.787	0.000	0.016	0.024
Afghanistan	-3.6376	0.116	-31.369	0.000	-3.865	-3.410
Armenia	-1.9726	0.116	-17.011	0.000	-2.200	-1.745
Azerbaijan	-1.5515	0.116	-13.379	0.000	-1.779	-1.324
Brazil	3.5276	0.116	30.421	0.000	3.300	3.755
Canada	1.2488	0.116	10.769	0.000	1.021	1.476
Chile	-0.1506	0.116	-1.299	0.194	-0.378	0.077
China	5.0745	0.116	43.760	0.000	4.847	5.302
Colombia	0.6301	0.116	5.434	0.000	0.403	0.858
Costa Rica	-1.5751	0.116	-13.583	0.000	-1.803	-1.348
Cuba	-3.8480	0.116	-33.184	0.000	-4.076	-3.621
Czech Republic	1.3152	0.116	11.342	0.000	1.088	1.543
Egypt	-0.6319	0.116	-5.449	0.000	-0.859	-0.404
France	3.0072	0.116	25.933	0.000	2.780	3.235
Germany	3.1358	0.116	27.041	0.000	2.908	3.363
Hungary	0.2385	0.116	2.056	0.040	0.011	0.466
India	2.7610	0.116	23.810	0.000	2.534	2.989
Indonezia	1.3192	0.116	11.377	0.000	1.092	1.547
Iran	1.1763	0.116	10.144	0.000	0.949	1.404
Italy	2.2894	0.116	19.743	0.000	2.062	2.517
Japan	3.3639	0.116	29.009	0.000	3.136	3.591
Kenya	-0.3822	0.116	-3.296	0.001	-0.610	-0.155
Mexico	1.2818	0.116	11.054	0.000	1.054	1.509
Moldova	-0.3385	0.116	-2.919	0.004	-0.566	-0.111
Paraguay	-1.0199	0.116	-8.795	0.000	-1.247	-0.792
Poland	2.3224	0.116	20.027	0.000	2.095	2.550
Slovakia	-0.5354	0.116	-4.617	0.000	-0.763	-0.308
Somalia	-2.5877	0.116	-22.315	0.000	-2.815	-2.360
South Korea	1.6567	0.116	14.287	0.000	1.429	1.884
Sudan	-2.3620	0.116	-20.369	0.000	-2.590	-2.135
Tanzania	-2.2058	0.116	-19.022	0.000	-2.433	-1.978
Togo	-2.4489	0.116	-21.118	0.000	-2.676	-2.221
Tynisia	-0.5313	0.116	-4.582	0.000	-0.759	-0.304
UK	1.8676	0.116	16.106	0.000	1.640	2.095
Uganda	-0.0887	0.116	-0.765	0.445	-0.316	0.139
Ukraine	1.2904	0.116	11.128	0.000	1.063	1.518
United States	4.4757	0.116	38.597	0.000	4.248	4.703
Venezuala	-0.5113	0.116	-4.409	0.000	-0.739	-0.284
Vietnam	1.9832	0.116	17.102	0.000	1.756	2.211
Zandia	-2.3716	0.116	-20.452	0.000	-2.599	-2.144
Zimbabwe	-2.1955	0.116	-18.933	0.000	-2.423	-1.968

Omnibus:	87.758	Durbin-Watson:	1.297
Prob(Omnibus):	0.000	Jarque-Bera (JB):	383.362
Skew:	0.139	Prob(JB):	5.68e-84
Kurtosis:	5.710	Cond. No.	1.72e+16

Рисунок 2.21 - Об’єднана регресійна модель для змінної KAS

Після побудови об’єднаної регресійної моделі, отримуємо наступне рівняння регресії (2.5):

$$\begin{aligned} \mathrm{KAS} = {} & 13.02 + 0.0203 \cdot \mathrm{Data}_{\mathrm{num}} - 3.64 \cdot \mathrm{Afganistan} - 1.97 \cdot \mathrm{Armenia} \\ & - 1.55 \cdot \mathrm{Azerbaijan} + \cdots - 2.37 \cdot \mathrm{Zandia} - 2.2 \cdot \mathrm{Zimbabwe} \\ & + \epsilon \end{aligned} \quad (2.5)$$

Скоригований R-квадрат, який вимірює частку загальної дисперсії в y, яка пояснюється X після врахування ступенів свободи, втрачених через включення змінних регресії, становить 0,920 або близько 92,0 %. Це, безумовно гарний результат.

F - тест для регресії, який вимірює спільну значущість параметрів моделі, дав тестову статистику 357,2 із значенням p = 0,00, що дозволяє зробити висновок, що оцінки коефіцієнтів моделі є спільно значущими при p < 0,001.

Log-правдоподібність моделі становить – 1211,8, а показник AIC становить 2506.

Проаналізуємо залишкові похибки (рис. 2.22):

```
print(pooled_olsr_model_results.resid)
print('Mean value of residual errors='+str(pooled_olsr_model_results.resid.mean()))
```

```
0      -0.350885
1      -0.028983
2      -0.070020
3       0.009218
4      -0.097199
          ...
1235    0.073640
1236   -0.193884
1237   -0.388195
1238   -0.244688
1239   -0.475146
Length: 1240, dtype: float64
Mean value of residual errors=-6.661767911502407e-14
```

Рисунок 2.22 – Залишкові похибки моделі

Це говорить нам про те, що регресійна модель прогнозує KAS із середньою похибкою близько -6,66e-14.

Визначаємо залежну та незалежні змінні (рис. 2.23):

```
y_var_name = 'KAS'
X_var_names = df.drop(['MAV', 'KAS', 'IDS', 'Date', 'Country'], axis=1)
```

Рисунок 2.23 - Визначення вхідних даних моделі

Визначаємо всі країни, які будуть незалежними змінними, які впливають на залежну та будуємо рівняння регресії (2.24-2.25):

```
lsdv_expr = y_var_name + ' ~ '
i = 0
for X_var_name in X_var_names:
    if i > 0:
        lsdv_expr = lsdv_expr + ' + ' + X_var_name
    else:
        lsdv_expr = lsdv_expr + X_var_name
    i = i + 1
for dummy_name in unit_names[:-1]:
    lsdv_expr = lsdv_expr + ' + ' + dummy_name

print('Regression expression for OLS with dummies=' + lsdv_expr)
```

```
Regression expression for OLS with dummies=KAS ~ Date_num + Germany + Italy + UK + Poland + France + Hungary + Moldova + Slovak
ia + Afghanistan + Indonezia + Japan + China + Vietnam + Armenia + Azerbaijan + Iran + India + Zamdia + Kenya + Zimbabwe + Egyp
t + Sudan + Somalia + Tynisia + Togo + Uganda + Tanzania + Canada + Colombia + Mexico + Cuba + Paraguay + Chile + Brazil
```

Рисунок 2.24 – Побудова рівняння регресії

Рівняння регресії з фіксованими змінними виглядає наступним чином (2.6):

$$\mathrm{KAS} = 14.13 + 0.0203 \cdot \mathrm{Data}_{\mathrm{num}} - 2.03 \cdot \mathrm{Germany} - 1.18 \cdot \mathrm{Italy} - \\ -1.25 \cdot \mathrm{Chile} + \cdots + 2.42 \cdot \mathrm{Brazil} + \epsilon \qquad (2.6)$$

Далі розглянемо коефіцієнти для цікавих фіктивних змінних, що представляють вплив на конкретну країну. Ми спостерігаємо, що відрізок регресії, який представляє специфічний для країни ефект для України (пропущена змінна), становить 0,822 і є статистично значущим (це означає, що його значення для населення оцінюється як відмінне від нуля), при p-значенні 0,000.

Скориговане значення R-квадрат дорівнює 0,817, або 81,7% - значення показує дуже хорошу відповідність між незалежними змінними та залежною.

OLS Regression Results

Dep. Variable:	KAS	R-squared:	0.822
Model:	OLS	Adj. R-squared:	0.817
Method:	Least Squares	F-statistic:	158.6
Date:	Wed, 23 Nov 2022	Prob (F-statistic):	0.00
Time:	14:20:08	Log-Likelihood:	-1728.9
No. Observations:	1240	AIC:	3530.
Df Residuals:	1204	BIC:	3714.
Df Model:	35		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	14.1277	0.088	159.958	0.000	13.954	14.301
Date_num	0.0203	0.003	6.463	0.000	0.014	0.026
Germany	2.0272	0.192	10.554	0.000	1.650	2.404
Italy	1.1808	0.192	6.147	0.000	0.804	1.558
UK	0.7590	0.192	3.952	0.000	0.382	1.136
Poland	1.2138	0.192	6.319	0.000	0.837	1.591
France	1.8986	0.192	9.884	0.000	1.522	2.275
Hungary	-0.8701	0.192	-4.530	0.000	-1.247	-0.493
Moldova	-1.4471	0.192	-7.534	0.000	-1.824	-1.070
Slovakia	-1.6440	0.192	-8.559	0.000	-2.021	-1.267
Afghanistan	-4.7462	0.192	-24.709	0.000	-5.123	-4.369
Indonezia	0.2106	0.192	1.097	0.273	-0.166	0.587
Japan	2.2553	0.192	11.741	0.000	1.878	2.632
China	3.9659	0.192	20.647	0.000	3.589	4.343
Vietnam	0.8746	0.192	4.553	0.000	0.498	1.251
Armenia	-3.0812	0.192	-16.041	0.000	-3.458	-2.704
Azerbaijan	-2.6601	0.192	-13.849	0.000	-3.037	-2.283
Iran	0.0677	0.192	0.352	0.725	-0.309	0.445
India	1.6524	0.192	8.603	0.000	1.276	2.029
Zamdia	-3.4802	0.192	-18.118	0.000	-3.857	-3.103
Kenya	-1.4908	0.192	-7.761	0.000	-1.868	-1.114
Zimbabwe	-3.3041	0.192	-17.202	0.000	-3.681	-2.927
Egypt	-1.7405	0.192	-9.061	0.000	-2.117	-1.364
Sudan	-3.4706	0.192	-18.068	0.000	-3.847	-3.094
Somalia	-3.6962	0.192	-19.243	0.000	-4.073	-3.319
Tynisia	-1.6399	0.192	-8.537	0.000	-2.017	-1.263
Togo	-3.5575	0.192	-18.521	0.000	-3.934	-3.181
Uganda	-1.1973	0.192	-6.233	0.000	-1.574	-0.820
Tanzania	-3.3143	0.192	-17.255	0.000	-3.691	-2.937
Canada	0.1402	0.192	0.730	0.466	-0.237	0.517
Colombia	-0.4784	0.192	-2.491	0.013	-0.855	-0.102
Mexico	0.1732	0.192	0.902	0.367	-0.204	0.550
Cuba	-4.9566	0.192	-25.805	0.000	-5.333	-4.580
Paraguay	-2.1285	0.192	-11.081	0.000	-2.505	-1.752
Chile	-1.2592	0.192	-6.556	0.000	-1.636	-0.882
Brazil	2.4190	0.192	12.594	0.000	2.042	2.796

Omnibus:	150.602	Durbin-Watson:	0.605
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1164.189
Skew:	0.256	Prob(JB):	1.58e-253
Kurtosis:	7.719	Cond. No.	303.

Рисунок 2.25 – Регресія фіксованих змінних для змінної IDS

Першим кроком для побудови моделі випадкових ефектів, необхідно розрахувати σ^2_ϵ і σ^2_μ - дисперсії компонентів похибки μ і ємоделі фіксованих ефектів та об'єднаної моделі та знайти різницю між ними (рис. 2.26).

```
sigma2_epsilon = lsdv_model_results.ssr/(n*T-(n+k+1))
print('sigma2_epsilon = ' + str(sigma2_epsilon))
```

```
sigma2_epsilon = -31.901267130742646
```

```
sigma2_pooled = pooled_olsr_model_results.ssr/(n*T-(k+1))
print('sigma2_pooled = ' + str(sigma2_pooled))
```

```
sigma2_pooled = -256.29655163768166
```

```
sigma2_u = sigma2_pooled - sigma2_epsilon
print('sigma2_u = ' + str(sigma2_u))
```

```
sigma2_u = -224.39528450693902
```

Рисунок 2.26 – Розрахунок значень дисперсії компонентів похибки моделей

Обчислюємо середні значення y та X для кожної групи (тобто кожної одиниці i) на панелі даних таким чином та розраховуємо показник Тета (рис. 2.27):

```
theta = 1 - math.sqrt(c/(u))
print('theta = ' + str(theta))
```

```
theta = 0.9367805559430118
```

Рисунок 2.27 – Обчислення показника Тета

Тепер будуємо модель випадкових ефектів (рис. 2.28)

OLS Regression Results

Dep. Variable:	KAS	R-squared:	0.662
Model:	OLS	Adj. R-squared:	0.653
Method:	Least Squares	F-statistic:	73.98
Date:	Sat, 26 Nov 2022	Prob (F-statistic):	2.80e-258
Time:	13:41:40	Log-Likelihood:	-21938.
No. Observations:	1240	AIC:	4.394e+04
Df Residuals:	1207	BIC:	4.411e+04
Df Model:	32		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.168e+08	1.12e+07	10.414	0.000	9.48e+07	1.39e+08
Date_num	8.201e+04	3.76e+04	2.183	0.029	8321.468	1.56e+05
UK	-4.902e+07	3.54e+07	-1.384	0.167	-1.19e+08	2.05e+07
Germany	1.259e+08	3.54e+07	3.555	0.000	5.64e+07	1.95e+08
Italy	-9.15e+06	3.54e+07	-0.258	0.796	-7.87e+07	6.04e+07
Poland	-5.396e+06	3.54e+07	-0.152	0.879	-7.49e+07	6.41e+07
France	2.023e+08	3.54e+07	5.711	0.000	1.33e+08	2.72e+08
Hungary	-9.153e+07	3.54e+07	-2.584	0.010	-1.61e+08	-2.2e+07
Moldova	-1.096e+08	3.54e+07	-3.094	0.002	-1.79e+08	-4.01e+07
Slovakia	-1.115e+08	3.54e+07	-3.148	0.002	-1.81e+08	-4.2e+07
Indonezia	-7.33e+07	3.54e+07	-2.069	0.039	-1.43e+08	-3.8e+06
Japan	1.699e+08	3.54e+07	4.795	0.000	1e+08	2.39e+08
China	1.511e+09	3.54e+07	42.647	0.000	1.44e+09	1.58e+09
Vietnam	-3.76e+07	3.54e+07	-1.061	0.289	-1.07e+08	3.19e+07
Armenia	-1.163e+08	3.54e+07	-3.284	0.001	-1.86e+08	-4.68e+07
Azerbaijan	-1.159e+08	3.54e+07	-3.270	0.001	-1.85e+08	-4.64e+07
Iran	-7.528e+07	3.54e+07	-2.125	0.034	-1.45e+08	-5.77e+06
India	5.721e+07	3.54e+07	1.615	0.107	-1.23e+07	1.27e+08
Zamdia	-1.17e+08	3.54e+07	-3.303	0.001	-1.86e+08	-4.75e+07
Kenya	-1.094e+08	3.54e+07	-3.089	0.002	-1.79e+08	-3.99e+07
Zimbabwe	-1.168e+08	3.54e+07	-3.298	0.001	-1.86e+08	-4.73e+07
Egypt	-1.107e+08	3.54e+07	-3.126	0.002	-1.8e+08	-4.12e+07
Sudan	-1.169e+08	3.54e+07	-3.299	0.001	-1.86e+08	-4.74e+07
Somalia	-1.171e+08	3.54e+07	-3.307	0.001	-1.87e+08	-4.76e+07
Tynisia	-1.103e+08	3.54e+07	-3.114	0.002	-1.8e+08	-4.08e+07
Togo	-1.17e+08	3.54e+07	-3.303	0.001	-1.87e+08	-4.75e+07
Canada	-6.731e+07	3.54e+07	-1.900	0.058	-1.37e+08	2.19e+06
Colombia	-9.538e+07	3.54e+07	-2.693	0.007	-1.65e+08	-2.59e+07
Mexico	-5.944e+07	3.54e+07	-1.678	0.094	-1.29e+08	1.01e+07
Cuba	-1.179e+08	3.54e+07	-3.327	0.001	-1.87e+08	-4.84e+07
Chile	-1.084e+08	3.54e+07	-3.061	0.002	-1.78e+08	-3.89e+07
Brazil	2.629e+08	3.54e+07	7.421	0.000	1.93e+08	3.32e+08
Ukraine	-7.687e+07	3.54e+07	-2.170	0.030	-1.46e+08	-7.37e+06

Omnibus:	1386.189	Durbin-Watson:	0.725
Prob(Omnibus):	0.000	Jarque-Bera (JB):	118570.476
Skew:	5.516	Prob(JB):	0.00
Kurtosis:	49.618	Cond. No.	1.92e+03

Рисунок 2.28 – Побудова моделі регресії випадкових ефектів

Рівняння регресії з випадковим ефектом виглядає наступним чином (2.7):

$$\text{KAS} = 1.168\text{e} + 08 + 8.201\text{e} + 04 \cdot \text{Data}_{\text{num}} - 4.902\text{e} + 07 \cdot \text{UK1.} + 259\text{e} + 08 \cdot \text{Germany} + \cdots - 7.687 \cdot \text{Ukraine} + \epsilon \quad (2.7)$$

Обчислена дисперсія σ^2_u було оцінено як - -224,4, а σ^2_ϵ було оцінено як – 31,9. Таким чином, частка загальної дисперсії, яка може бути віднесена до випадкового ефекту окремої одиниці, дорівнює (2.8):

$$\frac{-224.4}{-224.7+(-31.9)} = 0.87 \quad (2.8)$$

Це означає, що на 87% присутній випадковий ефект у моделі, але дивлячись на показник скоригованого R, який дорівнює 0,653, можна зробити висновок, що ця модель не є кращою за об’єднану модель та модель фіксованого ефекту.

Тестова статистика тесту LM (18,0083) більша, ніж критичне значення Chi-squared = 3,84146 при α=0,05, що означає, що випадковий ефект є значущим при альфа 0,05.

Необхідно визначити залежну та незалежні змінні.

Залежна змінна – це IDS, яка показує потік даних з виявлених мережевих атак. Визначаємо вхідні дані моделі (рис. 2.29):

```
y = 'LN_IDS'
x = df_data.drop(['MAV', 'KAS', 'IDS', 'Date', 'Country','LN_IDS'], axis=1)
```

Рисунок 12.29 – Вхідні змінні моделі

Результати моделі представлені на рисунку 2.30.

OLS Regression Results

Dep. Variable:	LN_IDS	R-squared:	0.959
Model:	OLS	Adj. R-squared:	0.958
Method:	Least Squares	F-statistic:	703.9
Date:	Sat, 26 Nov 2022	Prob (F-statistic):	0.00
Time:	12:22:28	Log-Likelihood:	-787.96
No. Observations:	1240	AIC:	1658.
Df Residuals:	1199	BIC:	1868.
Df Model:	40		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Date_num	0.0022	0.001	1.479	0.139	-0.001	0.005
Afghanistan	7.1437	0.087	82.387	0.000	6.974	7.314
Armenia	8.7331	0.087	100.716	0.000	8.563	8.903
Azerbaijan	8.1421	0.087	93.901	0.000	7.972	8.312
Brazil	12.7651	0.087	147.216	0.000	12.595	12.935
Canada	11.2622	0.087	129.884	0.000	11.092	11.432
Chile	11.8809	0.087	137.019	0.000	11.711	12.051
China	13.9827	0.087	161.258	0.000	13.813	14.153
Colombia	11.7506	0.087	135.517	0.000	11.581	11.921
Costa Rica	9.7092	0.087	111.973	0.000	9.539	9.879
Cuba	6.1282	0.087	70.674	0.000	5.958	6.298
Czech Republic	10.2438	0.087	118.138	0.000	10.074	10.414
Egypt	11.3155	0.087	130.499	0.000	11.145	11.486
France	12.6110	0.087	145.440	0.000	12.441	12.781
Germany	12.4079	0.087	143.097	0.000	12.238	12.578
Hungary	9.4115	0.087	108.540	0.000	9.241	9.582
India	12.2508	0.087	141.285	0.000	12.081	12.421
Indonezia	12.7813	0.087	147.403	0.000	12.611	12.951
Iran	11.7913	0.087	135.985	0.000	11.621	11.961
Italy	12.2955	0.087	141.800	0.000	12.125	12.466
Japan	9.6814	0.087	111.653	0.000	9.511	9.851
Kenya	10.5971	0.087	122.213	0.000	10.427	10.767
Mexico	12.9486	0.087	149.332	0.000	12.778	13.119
Moldova	8.9293	0.087	102.980	0.000	8.759	9.099
Paraguay	9.1457	0.087	105.475	0.000	8.976	9.316
Poland	11.6987	0.087	134.918	0.000	11.529	11.869
Slovakia	12.0494	0.087	138.963	0.000	11.879	12.220
Somalia	4.3781	0.087	50.491	0.000	4.208	4.548
South Korea	11.2754	0.087	130.036	0.000	11.105	11.446
Sudan	10.6387	0.087	122.693	0.000	10.469	10.809
Tanzania	9.3409	0.087	107.726	0.000	9.171	9.511
Togo	5.7743	0.087	66.593	0.000	5.604	5.944
Tynisia	10.6471	0.087	122.791	0.000	10.477	10.817
UK	11.0886	0.087	127.882	0.000	10.919	11.259
Uganda	8.6322	0.087	99.553	0.000	8.462	8.802
Ukraine	9.7300	0.087	112.214	0.000	9.560	9.900
United States	13.2521	0.087	152.833	0.000	13.082	13.422
Venezuala	10.2241	0.087	117.912	0.000	10.054	10.394
Vietnam	12.8515	0.087	148.213	0.000	12.681	13.022
Zandia	6.9104	0.087	79.696	0.000	6.740	7.081
Zimbabwe	8.1591	0.087	94.097	0.000	7.989	8.329

Omnibus:	599.721	Durbin-Watson:	1.077
Prob(Omnibus):	0.000	Jarque-Bera (JB):	14763.177
Skew:	-1.692	Prob(JB):	0.00
Kurtosis:	19.562	Cond. No.	238.

Рисунок 2.30 – Об’єднана регресійна модель для змінної IDS

Після побудови об’єднаної регресійної моделі, отримуємо наступне рівняння регресії (2.9):

$$\begin{aligned} IDS = 00.22 + 7.14 \cdot Afganistan + 8.73 \cdot Armenia + 6.91 \cdot Zandia + \\ +8.16 \cdot Zimbabwe\ 6 + \epsilon \end{aligned} \quad (2.9)$$

Скоригований R-квадрат становить 0,958 або 95,8 %. Це, безумовно гарний результат.

F - тест для регресії дав тестову статистику 703,9 із значенням p 0,00, що дозволяє зробити висновок, що оцінки коефіцієнтів моделі є значущими при p < 0.001. оцінки коефіцієнтів моделі є спільно значущими при p < 0,001.

Log-правдоподібність моделі становить -787,96, а показник AIC 1658.

Проаналізуємо залишкові похибки моделі для нормальності, гетероскедастичності та кореляції - трьох властивостей, які впливають на відповідність лінійної моделі (рис. 2.31).

```
print(pooled_olsr_model_results.resid)
print("Mean value of residual errors="+str(pooled_olsr_model_results.resid.mean()))
```

```
0       0.179944
1       0.178803
2       0.216327
3       0.238730
4       0.165151
          ...
1235   -0.152933
1236   -0.247574
1237   -0.222308
1238   -0.104814
1239   -0.041814
Length: 1240, dtype: float64
Mean value of residual errors=6.778448770025915e-15
```

Рисунок 2.31 – Залишкові похибки моделі

Визначаємо залежну та незалежні змінні (рис. 2.32):

```
y_var_name = 'LN_IDS'
X_var_names = df.drop(['MAV', 'KAS', 'IDS', 'Date', 'Country', 'Country_id'], axis=1)
```

Рисунок 2.32 - Визначення вхідних даних моделі

Визначаємо всі країни, які будуть незалежними змінними, які впливають на залежну та будуємо рівняння регресії (рис. 2.33):

```
lsdv_expr = y_var_name + ' ~ '
i = 0
for X_var_name in X_var_names:
    if i > 0:
        lsdv_expr = lsdv_expr + ' + ' + X_var_name
    else:
        lsdv_expr = lsdv_expr + X_var_name
    i = i + 1
for dummy_name in unit_names[:-1]:
    lsdv_expr = lsdv_expr + ' + ' + dummy_name

print('Regression expression for OLS with dummies=' + lsdv_expr)
```

```
Regression expression for OLS with dummies=LN_IDS ~ Date_num + Germany + Italy + UK + Poland + France + Hungary + Moldova + Slo
vakia + Afghanistan + Indonezia + Japan + China + Vietnam + Armenia + Azerbaijan + Iran + India + Zamdia + Kenya + Zimbabwe + E
gypt + Sudan + Somalia + Tynisia + Togo + Uganda + Tanzania + Canada + Colombia + Mexico + Cuba + Paraguay + Chile + Brazil
```

Рисунок 2.33 – Побудова рівняння регресії

OLS Regression Results

Dep. Variable:	LN_IDS	R-squared:	0.914
Model:	OLS	Adj. R-squared:	0.912
Method:	Least Squares	F-statistic:	366.4
Date:	Sat, 26 Nov 2022	Prob (F-statistic):	0.00
Time:	12:38:52	Log-Likelihood:	-1248.3
No. Observations:	1240	AIC:	2569.
Df Residuals:	1204	BIC:	2753.
Df Model:	35		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	10.7391	0.060	179.151	0.000	10.621	10.857
Date_num	0.0022	0.002	1.023	0.307	-0.002	0.006
Germany	1.6688	0.130	12.801	0.000	1.413	1.925
Italy	1.5564	0.130	11.938	0.000	1.301	1.812
UK	0.3496	0.130	2.681	0.007	0.094	0.605
Poland	0.9596	0.130	7.361	0.000	0.704	1.215
France	1.8720	0.130	14.359	0.000	1.616	2.128
Hungary	-1.3276	0.130	-10.183	0.000	-1.583	-1.072
Moldova	-1.8097	0.130	-13.882	0.000	-2.066	-1.554
Slovakia	1.3103	0.130	10.051	0.000	1.055	1.566
Afghanistan	-3.5954	0.130	-27.579	0.000	-3.851	-3.340
Indonezia	2.0422	0.130	15.665	0.000	1.786	2.298
Japan	-1.0577	0.130	-8.113	0.000	-1.313	-0.802
China	3.2436	0.130	24.880	0.000	2.988	3.499
Vietnam	2.1125	0.130	16.204	0.000	1.857	2.368
Armenia	-2.0060	0.130	-15.387	0.000	-2.262	-1.750
Azerbaijan	-2.5970	0.130	-19.920	0.000	-2.853	-2.341
Iran	1.0522	0.130	8.071	0.000	0.796	1.308
India	1.5117	0.130	11.596	0.000	1.256	1.767
Zamdia	-3.8287	0.130	-29.368	0.000	-4.084	-3.573
Kenya	-0.1420	0.130	-1.089	0.276	-0.398	0.114
Zimbabwe	-2.5799	0.130	-19.790	0.000	-2.836	-2.324
Egypt	0.5764	0.130	4.422	0.000	0.321	0.832
Sudan	-0.1004	0.130	-0.770	0.441	-0.356	0.155
Somalia	-6.3610	0.130	-48.793	0.000	-6.617	-6.105
Tynisia	-0.0919	0.130	-0.705	0.481	-0.348	0.164
Togo	-4.9648	0.130	-38.083	0.000	-5.221	-4.709
Uganda	-2.1069	0.130	-16.161	0.000	-2.363	-1.851
Tanzania	-1.3982	0.130	-10.725	0.000	-1.654	-1.142
Canada	0.5231	0.130	4.013	0.000	0.267	0.779
Colombia	1.0116	0.130	7.759	0.000	0.756	1.267
Mexico	2.2095	0.130	16.948	0.000	1.954	2.465
Cuba	-4.6109	0.130	-35.369	0.000	-4.867	-4.355
Paraguay	-1.5934	0.130	-12.222	0.000	-1.849	-1.338
Chile	1.1418	0.130	8.758	0.000	0.886	1.398
Brazil	2.0260	0.130	15.541	0.000	1.770	2.282

Omnibus:	251.203	Durbin-Watson:	0.548
Prob(Omnibus):	0.000	Jarque-Bera (JB):	2683.788
Skew:	0.608	Prob(JB):	0.00
Kurtosis:	10.104	Cond. No.	303.

Рисунок 2.34 – Регресія фіксованих змінних для змінної IDS

Рівняння регресії з фіксованими змінними виглядає наступним чином (2.10):

$$\text{IDS} = 10.74 + 00.22 \cdot \text{Germany} + 1.53 \cdot \text{Italy} + \ldots + 1.14 \cdot \text{Chile} + 2.03 \cdot \text{Brazil} + \epsilon \quad (2.10)$$

Скориговане значення R-квадрат дорівнює 0,912, або 91,2% - значення показує дуже хорошу відповідність між незалежними змінними та залежною.

F - тест для регресійного аналізу перевіряє, чи всі коефіцієнти моделі є спільно значущими, і, отже, чи відповідність моделі FE краща, ніж у попередній моделі. Статистика F-тесту 366,4 є значною при $p < 0,00$, що означає, що відповідність моделі справді краща, ніж в об'єднаній регресійній моделі.

Першим кроком для побудови моделі випадкових ефектів, необхідно розрахувати σ^2_ϵ і σ^2_μ - дисперсії компонентів похибки μ і ϵ моделі фіксованих ефектів та об'єднаної моделі та знайти різницю між ними (рис. 2.35).

```
sigma2_epsilon = lsdv_model_results.ssr/(n*T-(n+k+1))
print('sigma2_epsilon = ' + str(sigma2_epsilon))
```

```
sigma2_epsilon = -14.695039254289002
```

```
sigma2_pooled = pooled_olsr_model_results.ssr/(n*T-(k+1))
print('sigma2_pooled = ' + str(sigma2_pooled))
```

```
sigma2_pooled = -129.37891349913338
```

```
sigma2_u = sigma2_pooled - sigma2_epsilon
print('sigma2_u = ' + str(sigma2_u))
```

```
sigma2_u = -114.68387424484438
```

Рисунок 2.35 – Розрахунок значень дисперсії компонентів похибки моделей

Обчислюємо середні значення *y* та *X* для кожної групи (тобто кожної одиниці *i*) на панелі даних таким чином та розраховуємо показник Тета (рис. 2.36):

```
theta = 1 - math.sqrt(c/(u))
print('theta = ' + str(theta))
```

```
theta = 0.939969296454986
```

Рисунок 2.36 – Обчислення показника Тета

Тепер будуємо модель випадкових ефектів (рис. 2.37):

OLS Regression Results

Dep. Variable:	IDS	R-squared:	0.741
Model:	OLS	Adj. R-squared:	0.734
Method:	Least Squares	F-statistic:	107.7
Date:	Sat, 26 Nov 2022	Prob (F-statistic):	0.00
Time:	13:37:43	Log-Likelihood:	-16342.
No. Observations:	1240	AIC:	3.275e+04
Df Residuals:	1207	BIC:	3.292e+04
Df Model:	32		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.451e+06	1.3e+05	11.194	0.000	1.2e+06	1.71e+06
Date_num	249.9978	412.073	0.607	0.544	-558.462	1058.458
Germany	2.798e+06	4.09e+05	6.835	0.000	1.99e+06	3.6e+06
Italy	2.332e+06	4.09e+05	5.698	0.000	1.53e+06	3.14e+06
UK	-3.164e+05	4.09e+05	-0.773	0.440	-1.12e+06	4.87e+05
Poland	7.927e+05	4.09e+05	1.937	0.053	-1.04e+04	1.6e+06
France	3.732e+06	4.09e+05	9.117	0.000	2.93e+06	4.53e+06
Hungary	-1.241e+06	4.09e+05	-3.032	0.002	-2.04e+06	-4.38e+05
Moldova	-1.318e+06	4.09e+05	-3.219	0.001	-2.12e+06	-5.14e+05
Slovakia	4.456e+06	4.09e+05	10.886	0.000	3.65e+06	5.26e+06
Indonezia	4.802e+06	4.09e+05	11.733	0.000	4e+06	5.61e+06
Japan	-1.174e+06	4.09e+05	-2.868	0.004	-1.98e+06	-3.71e+05
China	1.913e+07	4.09e+05	46.746	0.000	1.83e+07	1.99e+07
Vietnam	5.699e+06	4.09e+05	13.924	0.000	4.9e+06	6.5e+06
Armenia	-1.341e+06	4.09e+05	-3.275	0.001	-2.14e+06	-5.38e+05
Azerbaijan	-1.383e+06	4.09e+05	-3.379	0.001	-2.19e+06	-5.8e+05
Iran	8.704e+05	4.09e+05	2.126	0.034	6.73e+04	1.67e+06
India	2.23e+06	4.09e+05	5.449	0.000	1.43e+06	3.03e+06
Zamdia	-1.436e+06	4.09e+05	-3.508	0.000	-2.24e+06	-6.33e+05
Kenya	-5.993e+05	4.09e+05	-1.464	0.143	-1.4e+06	2.04e+05
Zimbabwe	-1.392e+06	4.09e+05	-3.402	0.001	-2.2e+06	-5.89e+05
Egypt	-6192.2912	4.09e+05	-0.015	0.988	-8.09e+05	7.97e+05
Sudan	-7.24e+05	4.09e+05	-1.769	0.077	-1.53e+06	7.9e+04
Somalia	-1.452e+06	4.09e+05	-3.548	0.000	-2.26e+06	-6.49e+05
Tynisia	-7.084e+05	4.09e+05	-1.731	0.084	-1.51e+06	9.47e+04
Togo	-1.448e+06	4.09e+05	-3.537	0.000	-2.25e+06	-6.45e+05
Canada	-1.019e+05	4.09e+05	-0.249	0.803	-9.05e+05	7.01e+05
Colombia	1.092e+06	4.09e+05	2.668	0.008	2.89e+05	1.89e+06
Mexico	6.769e+06	4.09e+05	16.538	0.000	5.97e+06	7.57e+06
Cuba	-1.446e+06	4.09e+05	-3.532	0.000	-2.25e+06	-6.43e+05
Chile	1.065e+06	4.09e+05	2.602	0.009	2.62e+05	1.87e+06
Brazil	4.64e+06	4.09e+05	11.336	0.000	3.84e+06	5.44e+06
Ukraine	-1.161e+06	4.09e+05	-2.836	0.005	-1.96e+06	-3.58e+05

Omnibus:	1540.323	Durbin-Watson:	0.882
Prob(Omnibus):	0.000	Jarque-Bera (JB):	307346.453
Skew:	6.282	Prob(JB):	0.00
Kurtosis:	79.097	Cond. No.	2.02e+03

Рисунок 2.37 – Побудова моделі регресії випадкових ефектів

Рівняння регресії з випадковим ефектом виглядає наступним чином (2.11):

$$\begin{aligned} \text{IDS} = {} & 1.451\text{e} + 06 + 249.998 \cdot \text{Data}_{\text{num}} + 2.798\text{t} + 06 \cdot \text{Germany} \\ & + 3.322\text{t} + 06 \cdot \text{Italy} + \ldots + 4.64\text{t} + 06 \cdot \text{Chile} - 1.16\text{t} + 06 \\ & \cdot \text{Ukraine} + \epsilon \end{aligned} \qquad (2.11)$$

Обчислена дисперсія σ^2_u було оцінено як - -114,68, а σ^2_ϵ було оцінено як – 14,70. Таким чином, частка загальної дисперсії, яка може бути віднесена до випадкового ефекту окремої одиниці, дорівнює (2.12):

$$\frac{-114.68}{-114.68+(-14.70)} = 0{,}89 \qquad (2.12)$$

Це означає, що на 89% присутній випадковий ефект у моделі, але дивлячись на показник скоригованого R, який дорівнює 0,734, можна зробити висновок, що ця модель не є кращою за об’єднану модель та модель фіксованого ефекту.

Проведемо оцінку отриманих результатів побудованих регресій для кожного виду кібератак. В таблиці 2.2 представлені результати для MAV.

Таблиця 2.2 – Порівняння результатів побудованих моделей для MAV

Показник	Pooled OLS Regression Model	The Fixed Effects Regression Model	The Random Effects Regression Model
Скоригований R^2	0,769	0,702	0,640
Log-Likelihood	-1785,7	-1946,1	-12222
AIC	3653	3964	24510
MRE	5,804e-15	4,410e-13	1,017e-10

Скоригований R-квадрат об’єднаної моделі (Pooled OLS) дорівнює 76.9% значно кращий порівняно з моделями фіксованого та випадкового ефекту. Pooled OLS також забезпечує невелике збільшення логарифмічної ймовірності до -1785.7 порівняно з іншими моделями, я також показник AIC теж показав кращий результат. Середня похибка залишків теж є найменшою для об’єднаної моделі. Можна зробити висновок, що об’єднана модель є найкращою для опису залежної змінної MAV.

В таблиці 2.3 представлені результати для KAS.

Таблиця 2.3 – Порівняння результатів побудованих моделей для KAS

Показник	Pooled OLS Regression Model	The Fixed Effects Regression Model	The Random Effects Regression Model
Скоригований R^2	0,920	0,817	0,653
Log-Likelihood	-1277,8	-1728,9	-21938
AIC	2506	3530	43940
MRE	-6,66e-14	8,52e-13	3,92e-08

Скоригований R-квадрат об’єднаної моделі (Pooled OLS) дорівнює 92,0% значно кращий порівняно з моделями фіксованого та випадкового ефекту. Pooled OLS також забезпечує невелике збільшення логарифмічної ймовірності до -1277.8 порівняно з іншими моделями, я також показник AIC теж показав кращий результат. Середня похибка залишків теж є найменшою для об’єднаної моделі. Можна зробити висновок, що об’єднана модель є найкращою для опису залежної змінної KAS.

В таблиці 2.4 представлені результати для IDS.

Таблиця 2.4 – Порівняння результатів побудованих моделей для IDS

Показник	Pooled OLS Regression Model	The Fixed Effects Regression Model	The Random Effects Regression Model
Скоригований R^2	0,958	0,912	0,734
Log-Likelihood	-787,96	-1248,3	-16342
AIC	1658	2569	32750
MRE	6,77e-15	6,51e-13	1,41e-08

Скоригований R-квадрат об’єднаної моделі (Pooled OLS) дорівнює 95,8% значно кращий порівняно з моделями фіксованого та випадкового ефекту. Pooled OLS також забезпечує невелике збільшення логарифмічної ймовірності до -787.96 порівняно з іншими моделями, я також показник AIC теж показав кращий результат. Середня похибка залишків теж є найменшою для об’єднаної моделі. Можна зробити висновок, що об’єднана модель є найкращою для опису залежної змінної IDS.

Визначивши найкращу модель для моделювання виду кібератаки MAV, спрогнозуємо тренд за допомогою об’єднаної моделі (Pooled model) та LSTM моделі.

Перш ніж побудувати прогноз необхідно поділити базу даних на дві частини - тестову та тренувальну. Головна проблема та складність цього процесу полягає у тому, як правильно поділити панельні дані на дві частини. Розподіл відбувається для кожної країни, а потім об’єднуємо тестову та тренувальну частини для кожної країни (рис. 2.38).

```
def train_test_split(data):
    size=int(len(data)*0.8)
    # for train data will be collected from each country's data which index is from 0-size (80%)
    x_train =data.drop(['MAV'], axis=1).iloc[0:size]
    # for test data will be collected from each country's  data which index is from size to the end (20%)
    x_test = data.drop(['MAV'], axis=1).iloc[size:]
    y_train=data['MAV'].iloc[0:size]
    y_test=data['MAV'].iloc[size:]
    return x_train, x_test,y_train,y_test
```

```
country=list(set(dt.Countries))
# loop each station and collect train and test data
X_train=[]
X_test=[]
Y_train=[]
Y_test=[]
for i in range(0,len(country)):
    data=dt[dt['Countries']==country[i]]
    x_train, x_test,y_train,y_test=train_test_split(data)
    X_train.append(x_train)
    X_test.append(x_test)
    Y_train.append(y_train)
    Y_test.append(y_test)
```

Рисунок 2.38 – Розподіл бази даних та тренувальну та тестову частини

Після розподілу бази даних на тестову та тренувальну частину, ідентифікуємо тренувальні та тестові залежні та незалежні змінні (рис. 2.39) [**Ошибка! Источник ссылки не найден.**]:

```
from sklearn.preprocessing import LabelEncoder
encoder = LabelEncoder()
#combine x train and y train as train data
train_data=pd.DataFrame()
train_data[X_train.columns]=X_train
train_data[Y_train.columns]=Y_train
train_data['Countries']= encoder.fit_transform(train_data['Countries'])
#combine x test and y test as test data
test_data=pd.DataFrame()
test_data[X_test.columns]=X_test
test_data[Y_test.columns]=Y_test
test_data['Countries']= encoder.fit_transform(test_data['Countries'])
# using the function to obtian reshaped x_train,x_test,y_train,y_test
x_train,x_test,y_train,y_test=reshape_data(train_data,test_data)
```

Рисунок 2.39 – Ідентифікація тренувальних та тестових змінних

Всі дані підготовлені до побудови прогнозної моделі на основі об’єднаної моделі та LSTM моделі [**Ошибка! Источник ссылки не найден.**].

Спочатку побудуємо об’єднану прогнозу модель з використанням тестового та тренувального розподілу бази даних (рис. 2.40).

```
y_pred = model.predict(X_test)
df_results = pd.DataFrame({'Actual': y_test, 'Predicted': y_pred})
df_results
```

	Actual	Predicted
780	1.609438	1.174482
817	8.043021	7.272133
363	9.110631	8.806825
308	6.848005	6.625916
1205	7.132498	7.771581
...	...	...
609	8.336390	8.479036
332	5.587249	4.932872
1088	4.828314	4.211598
1137	3.367296	6.138822
149	7.920810	7.571089

Рисунок 2.40 – Результати побудови прогнозної моделі на основі Pooled regression

Одним із способів оцінити, наскільки добре регресійна модель відповідає набору даних, є обчислення середньої квадратичної похибки кореня (RMSE), яка є метрикою, яка повідомляє нам середню відстань між прогнозованими значеннями від моделі та фактичними значеннями в наборі даних [**Ошибка! Источник ссылки не найден.**].

Чим нижче RMSE, тим краще дана модель здатна «підігнати» набір даних (2.13):

$$\mathrm{RMSE} = \frac{\sqrt{\sum(\mathrm{P}-\mathrm{A})^2}}{\mathrm{n}} \qquad (2.13)$$

де P – прогнозне значення;

A – значення спостереження;

n – розмір вибірки.

Тому розраховуємо середню квадратичну похибку кореня та коефіцієнт детермінації для прогнозної моделі (рис. 2.41):

```
from sklearn.metrics import r2_score, mean_squared_error
RMSE = np.sqrt(mean_squared_error(y_test, y_pred))
r2 = r2_score(y_test, y_pred)
print('RMSE:', RMSE, 'R2:',r2)
```

```
RMSE: 1.1298507283393997 R2: 0.7140165385759949
```

Рисунок 2.41 – Розрахунок показників для прогнозної моделі для змінної MAV

Середню квадратичну похибку кореня становить 1,12985 та коефіцієнт детермінації дорівнює 0,71, що являється досить гарним результатом.

Тепер необхідно побудувати нову прогнозну модель - LSTM модель для порівняння та обрати кращу модель для прогнозування явища кібератак.

Всі дані підготовлені до побудови LSTM моделі. LSTM модель буде будуватися за допомогою Sequential() задачі (рис. 2.42) [**Ошибка! Источник ссылки не найден.**].

Sequential (Класифікація послідовностей) - це задача прогностичного моделювання, де є певна послідовність входів у просторі або часі, і завдання полягає в тому, щоб передбачити категорію для послідовності.

```
model = Sequential()
model.add(LSTM(60, activation='sigmoid',input_shape=(x_train.shape[1], x_train.shape[2])))
model.add(Dense(1))
model.compile(loss='mae', optimizer='adam')
# fit network
history = model.fit(x_train, y_train, epochs=1000, batch_size=64, verbose=0, shuffle=False)
```

```
# make a prediction
y_test_pre=model.predict(x_test)
```

```
9/9 [==============================] - 0s 875us/step
```

```
# make a prediction
y_test_pre=model.predict(x_test)
y_test_pre.shape,y_test.shape
```

```
9/9 [==============================] - 0s 750us/step

((279, 1), (279,))
```

Рисунок 2.42 – Побудова прогнозної LSTM моделі

Після побудови прогнозної моделі отримуємо наступні результати (рис. 2.43).

```
pa=pd.DataFrame()
pa['Data']=X_test.reset_index().Data.iloc[1:-1]
pa['Prediction']=[i[0] for i in y_test_pre][1:]
pa['Actual Values']=y_test[:-1]
pa.head()
```

	Data	Prediction	Actual Values
1	2022-09-08	5.344671	4.553877
2	2022-09-09	5.325020	4.418841
3	2022-09-10	4.153235	2.772589
4	2022-09-11	5.082784	3.912023
5	2022-09-12	5.840456	4.867534

Рисунок 2.43 – Прогнозні дані для змінної MAV

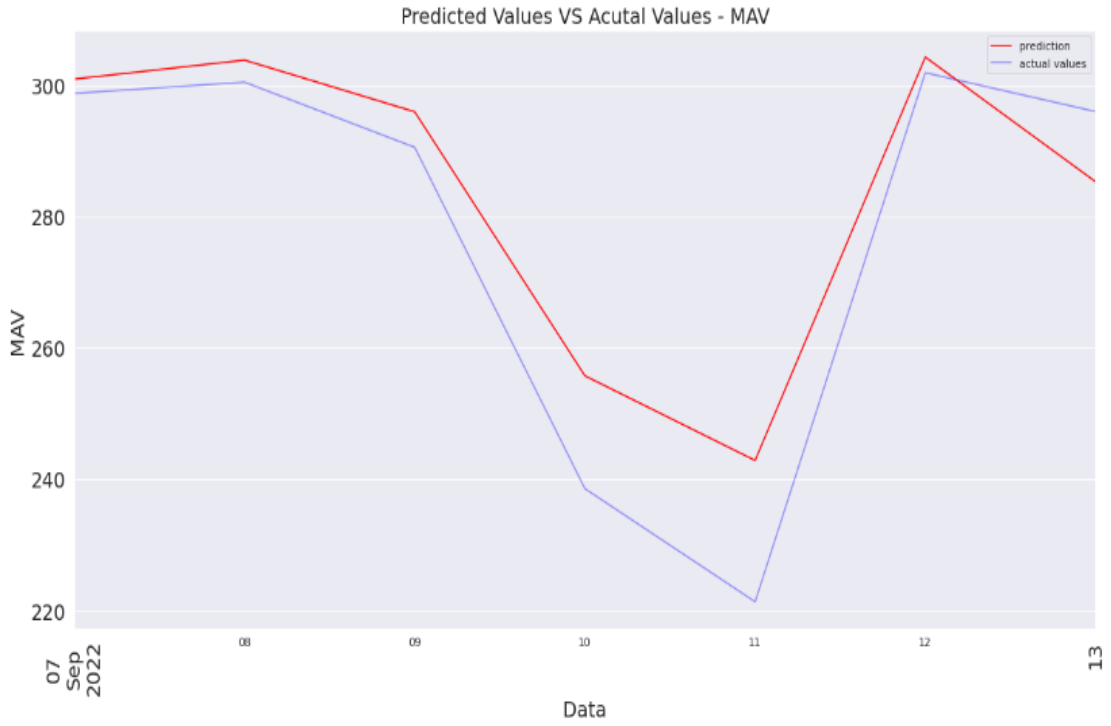


Рисунок 2.44 – Результати побудови прогнозної LSTM моделі

```
print(RMSE(y_test[:-1],[i[0] for i in y_test_pre][1:]))
```

```
0.5821494070053101
```

```
from sklearn.metrics import mean_squared_error, r2_score, mean_absolute_error
print('R2 Score: ', r2_score(y_test, y_test_pre))
```

```
R2 Score:  0.5151603123958282
```

Рисунок 2.45 – Розрахунок показників для прогнозної моделі для змінної MAV

Середню квадратичну похибку кореня становить 0,58 та коефіцієнт детермінації дорівнює 0,51, що являється теж досить гарним результатом. Але

спираючись на той факт, що коефіцієнт детермінації не є повністю надійним показником для порівняння моделей, використаємо середню квадратичну похибку, яка повідомляє нам середню відстань між прогнозованими значеннями від моделі та фактичними значеннями в наборі даних.

Тому кращою прогнозною моделлю для змінної MAV є LSTM модель. Спробуємо побудувати прогнозну LSTM модель для кожної країни (рис. 2.46).

```
def normalization_train_test_split(country):
    scaler = PowerTransformer(method='yeo-johnson', standardize=True)
    scaled = scaler.fit_transform(country.drop(columns=['Country','Date']))
    # create dateframe fir scaled data
    scaled_df=pd.DataFrame(data=scaled,columns=country.drop(columns=['Country','Date']).columns)
    scaled_df['MAV']=list(country.MAV)
    X_train, X_test,Y_train,Y_test=train_test_split(scaled_df)
    #combine x train and y train as train data
    train_data=pd.DataFrame()
    train_data[X_train.columns]=X_train
    train_data['MAV']=Y_train
    #combine x test and y test as test data
    test_data=pd.DataFrame()
    test_data[X_test.columns]=X_test
    test_data['MAV']=Y_test

    # using the function to obtian reshaped x_train,x_test,y_train,y_test
    x_train,x_test,y_train,y_test=reshape_data(train_data,test_data)
    return x_train, x_test,y_train,y_test
```

```
# Loop through top 10 countries' data
#
for i in range(len(top_10_country_names)):
     # obtian one country's data
    country=df_data[df_data.Country==top_10_country_names[i]]
    # train test split, nomalization and reshape the data
    x_train, x_test,y_train,y_test=normalization_train_test_split(country)
    # model
    model = Sequential()
    model.add(LSTM(60, activation='sigmoid',input_shape=(x_train.shape[1], x_train.shape[2])))
    model.add(Dense(1))
    model.compile(loss='mae', optimizer='adamax')
    # fit network
    history = model.fit(x_train, y_train, epochs=2000, batch_size=128,  verbose=0, shuffle=False)
    # make a prediction
    y_test_pre=model.predict(x_test)
    #RMSE
    rmse=RMSE(y_test[:-1],[i[0] for i in y_test_pre][1:])
    print('{} - RMSE: {}'.format(top_10_country_names[i],rmse))
    #create new dataframe for plot
    pa=pd.DataFrame()
    pa['Date']=list(country.Date.iloc[int(len(country)*0.8):])[1:-1]
    pa['Prediction']=[i[0] for i in y_test_pre][1:]
    pa['Actual Values']=list(y_test[:-1])

    plt.figure(figsize=(20,10))
    pa.groupby('Date')['Prediction'].sum().plot(kind='line',label='prediction',color='red',alpha=1)
    pa.groupby("Date")['Actual Values'].sum().plot(kind='line',label='actual values',color='blue',alpha=0.4)
    plt.xticks(rotation=90,size=20)
    plt.yticks(size=20)

    plt.ylabel('MAV',fontsize=20)
    plt.xlabel('Date',fontsize=20)
    plt.title('Predicted Values VS Acutal Values - MAV in {}'.format(top_10_country_names[i]),fontsize=20)
    plt.legend()
```

Рисунок 2.46 – Побудова прогнозної моделі для кожної країни

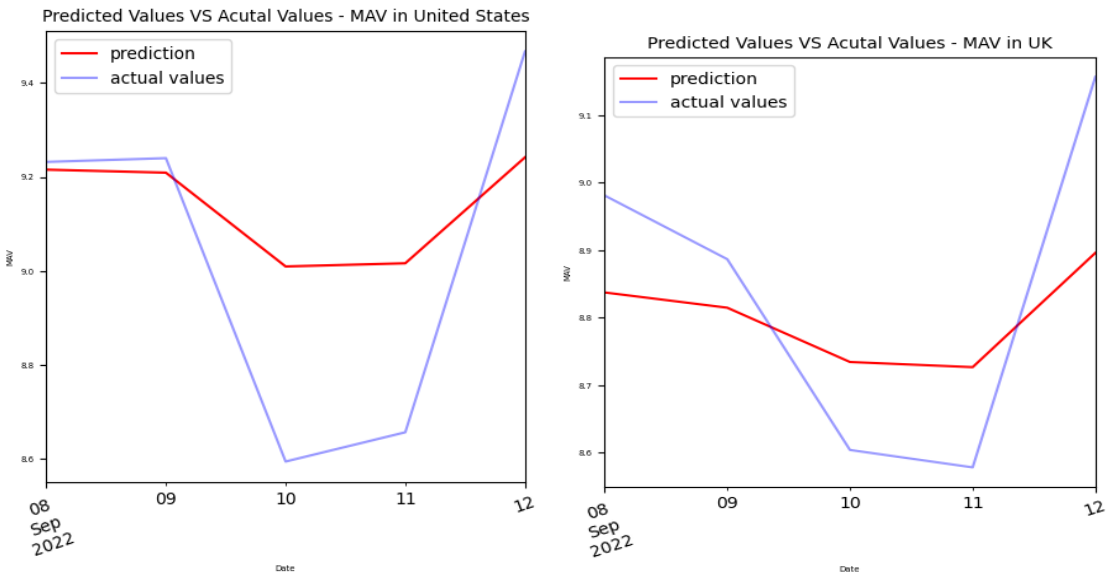


Рисунок 2.47 – Прогноз значення MAV для США та Великобританії

Отримані гарні результати для країн. Прикладом прогнозу для країн виступають США та Великобританія за значенням середньої квадратичної похибки 0,266 та 0,163 відповідно.

Визначивши найкращу модель для моделювання виду кібератаки KAS, спрогнозуємо тренд за допомогою об’єднаної моделі (Pooled model) та LSTM моделі. Спочатку побудуємо об’єднану прогнозу модель з використанням тестового та тренувального розподілу бази даних (рис. 2.48):

```
y_pred = model.predict(X_test)
df_results = pd.DataFrame({'Actual': y_test, 'Predicted': y_pred})
df_results
```

	Actual	Predicted
999	14.939141	13.788484
651	9.952278	12.723177
1023	13.138237	14.310919
250	12.398757	12.584448
860	11.362103	11.019525
...	...	...
631	11.532728	10.876397
473	15.882500	15.127464
916	11.497812	11.197719
760	10.373491	11.069807
173	16.437204	16.313346

Рисунок 2.48 – Результати побудови прогнозної моделі на основі Pooled regression

Одним із способів оцінити, наскільки добре регресійна модель відповідає набору даних, є обчислення середньої квадратичної похибки кореня (RMSE).

Тому розраховуємо середню квадратичну похибку кореня та коефіцієнт детермінації для прогнозної моделі (рис. 2.49):

```
from sklearn.metrics import r2_score, mean_squared_error
RMSE = np.sqrt(mean_squared_error(y_test, y_pred))
r2 = r2_score(y_test, y_pred)
print(RMSE, r2)
```

```
0.6147591605646721 0.9297993080835679
```

Рисунок 2.49 – Розрахунок показників для прогнозної моделі для змінної KAS

Середню квадратичну похибку кореня становить 0,6148 та коефіцієнт детермінації дорівнює 0,9298, що являється досить гарним результатом.

Тепер необхідно побудувати нову прогнозну модель - LSTM модель для порівняння та обрати кращу модель для прогнозування явища кібератак. Після побудови прогнозної моделі отримуємо наступні результати (рис. 2.50).

	Date	Prediction	Actual Values
1	2022-09-08	16.024281	15.841156
2	2022-09-09	16.052122	15.875068
3	2022-09-10	15.683013	15.471450
4	2022-09-11	15.695789	15.488205
5	2022-09-12	15.915629	15.735956

Рисунок 2.50 – Прогнозні дані для змінної KAS

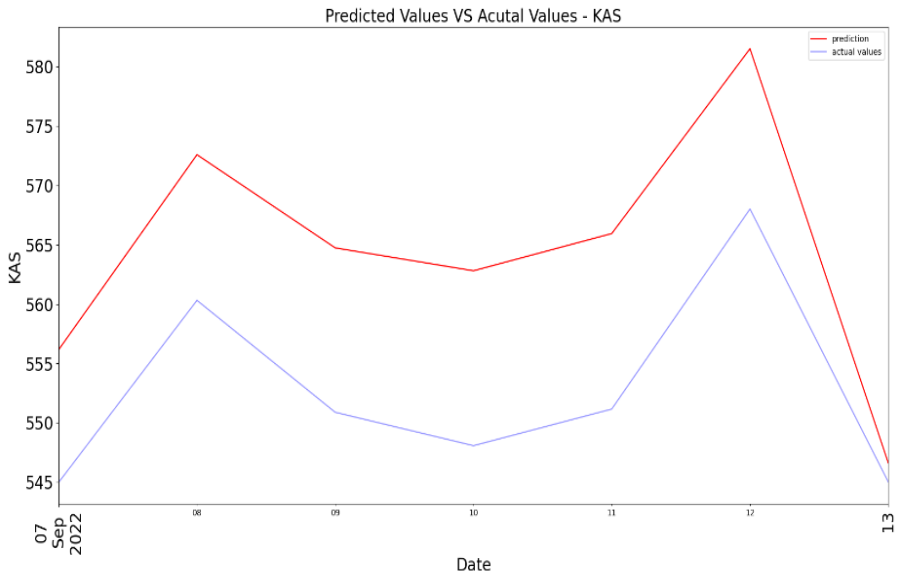


Рисунок 2.51 – Результати побудови прогнозної LSTM моделі

```
print(RMSE(y_test[:-1],[i[0] for i in y_test_pre][1:]))
```

```
0.505295611130907
```

```
from sklearn.metrics import mean_squared_error, r2_score, mean_absolute_error
print('R2 Score: ', r2_score(y_test, y_test_pre))
```

```
R2 Score:  0.7116407077885306
```

Рисунок 2.52 – Розрахунок показників для прогнозної моделі для змінної KAS

Середню квадратичну похибку кореня становить 0,51 та коефіцієнт детермінації дорівнює 0,71, що являється теж досить гарним результатом. Але спираючись на той факт, що коефіцієнт детермінації не є повністю надійним показником для порівняння моделей, використаємо середню квадратичну похибку, яка повідомляє нам середню відстань між прогнозованими значеннями від моделі та фактичними значеннями в наборі даних.

Тому кращою прогнозною моделлю для змінної KAS є LSTM модель. Спробуємо побудувати прогнозну LSTM модель для кожної країни (рис. 2.53).

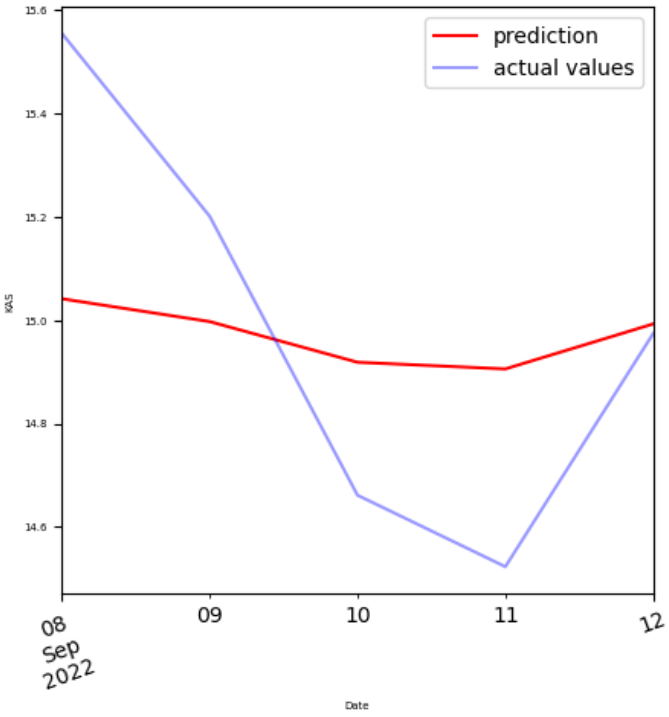


Рисунок 2.53 – Прогноз значення KAS для України

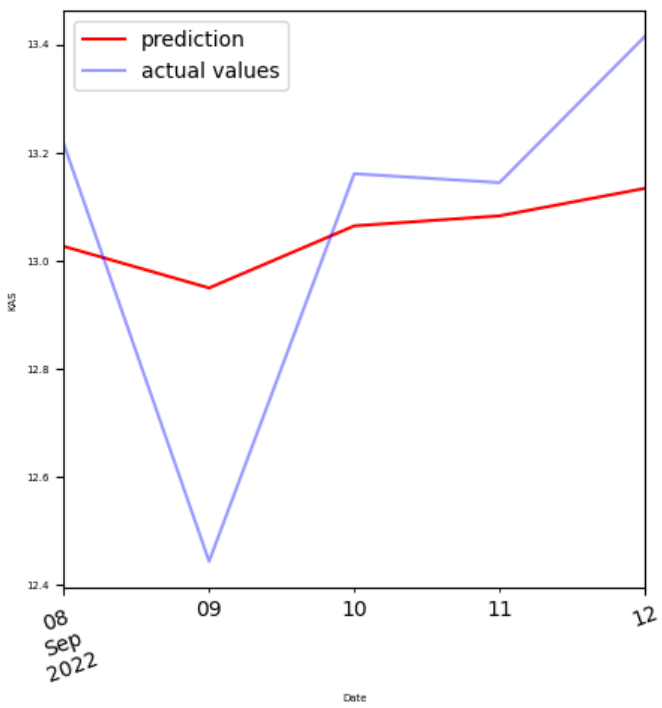


Рисунок 2.54 – Прогноз значення KAS для Тунісу

Отримані гарні результати для країн. Прикладом прогнозу для країн виступають Україна та Туніс за значенням середньої квадратичної похибки 0,329 та 0,277 відповідно.

Визначивши найкращу модель для моделювання виду кібератаки IDS, спрогнозуємо тренд за допомогою об'єднаної моделі (Pooled model) та LSTM моделі. Спочатку побудуємо об'єднану прогнозу модель з використанням тестового та тренувального розподілу бази даних (рис. 2.55).

```
y_pred = model.predict(X_test)
df_results = pd.DataFrame({'Actual': y_test, 'Predicted': y_pred})
df_results
```

	Actual	Predicted
729	11.349959	11.356542
925	9.655667	9.319099
1070	10.403202	10.244445
249	9.120963	12.228848
961	13.244149	13.253300
...	...	...
880	9.081597	8.644902
99	11.076465	11.119290
332	7.352441	7.202662
400	9.927790	9.709871
418	11.359040	11.312384

Рисунок 2.55 – Результати побудови прогнозної моделі на основі Pooled regression

Одним із способів оцінити, наскільки добре регресійна модель відповідає набору даних, є обчислення середньої квадратичної похибки кореня (RMSE).

Тому розраховуємо середню квадратичну похибку кореня та коефіцієнт детермінації для прогнозної моделі (рис. 2.56):

```
from sklearn.metrics import r2_score, mean_squared_error
RMSE = np.sqrt(mean_squared_error(y_test, y_pred))
r2 = r2_score(y_test, y_pred)
print(RMSE, r2)
```

```
0.5360467594061045 0.9429578957759858
```

Рисунок 2.56 – Розрахунок показників для прогнозної моделі для змінної IDS

Середню квадратичну похибку кореня становить 0,536 та коефіцієнт детермінації дорівнює 0,943, що являється досить гарним результатом.

Тепер необхідно побудувати нову прогнозну модель - LSTM модель для порівняння та обрати кращу модель для прогнозування явища кібератак. Після побудови прогнозної моделі отримуємо наступні результати (рис. 2.57).

	Date	Prediction	Actual Values
1	2022-09-08	9.170294	9.013839
2	2022-09-09	8.823923	8.629629
3	2022-09-10	8.770832	8.567126
4	2022-09-11	9.114446	8.941807
5	2022-09-12	9.222954	9.060099

Рисунок 2.57 – Прогнозні дані для змінної IDS

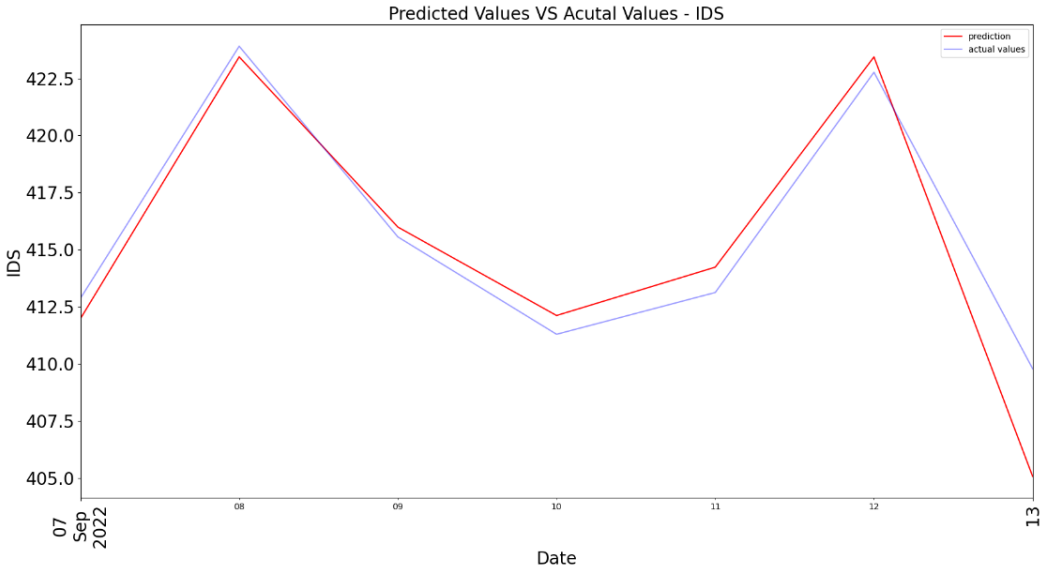


Рисунок 2.58 – Результати побудови прогнозної LSTM моделі

```
print(RMSE(y_test[:-1],[i[0] for i in y_test_pre][1:]))
```

```
0.4595955566642357
```

```
from sklearn.metrics import mean_squared_error, r2_score, mean_absolute_error
print('R2 Score: ', r2_score(y_test, y_test_pre))
```

```
R2 Score:  0.7257339291988435
```

Рисунок 2.59 – Розрахунок показників для прогнозної моделі для змінної IDS

Середню квадратичну похибку кореня становить 0,459 та коефіцієнт детермінації дорівнює 0,726, що являється теж досить гарним результатом.

Але спираючись на той факт, що коефіцієнт детермінації не є повністю надійним показником для порівняння моделей, використаємо середню квадратичну похибку, яка повідомляє нам середню відстань між прогнозованими значеннями від моделі та фактичними значеннями в наборі даних. Тому кращою прогнозною моделлю для змінної KAS є LSTM модель.

Спробуємо побудувати прогнозну LSTM модель для кожної країни (рис. 2.60-2.61).

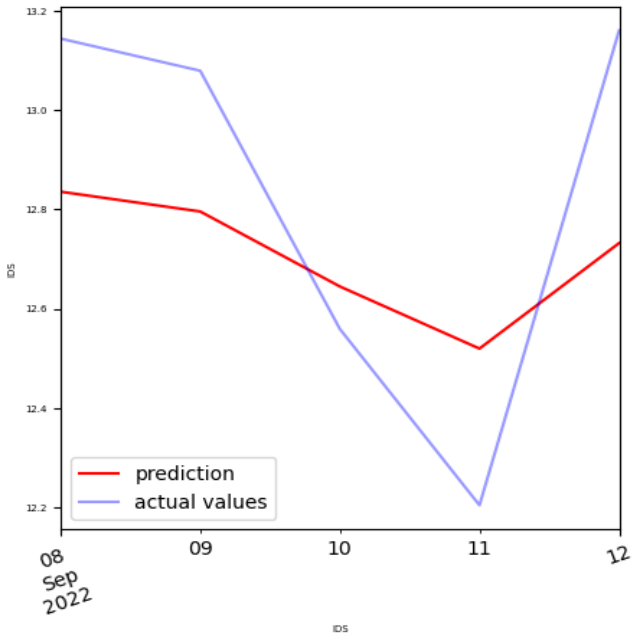


Рисунок 2.60 – Прогноз значення KAS для В’єтнаму

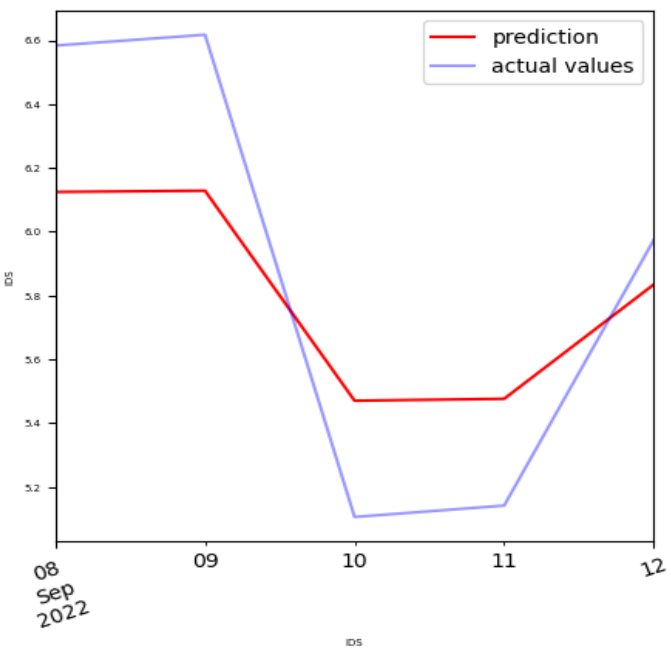


Рисунок 2.61 – Прогноз значення KAS для Того

Отримані гарні результати для країн. Прикладом прогнозу для країн виступають В'єтнам та Того за значенням середньої квадратичної похибки 0,304 та 0,377 відповідно.

У даному дослідженні було здійснено аналіз, моделювання та прогнозування трендів кібератак за допомогою побудови математичних моделей та регресії для панельних даних, а саме об'єднаної моделі, моделі фіксованих ефектів, випадкових ефектів та LSTM моделі. Відповідні розрахунки було проведено із використанням сучасної мови програмування Python. Вважаємо, що побудована об'єднана модель буде одним із найкращих методів, що дозволяє змоделювати тренди кібератак та продемонструвала гарні результати скоригованого коефіцієнту детермінації, залишкових похибок моделі та параметру Акайка (AIC). Також було побудовано прогнозну модель на основі об'єднаної та LSTM модель. На останньому етапі було проведено порівняння побудованих прогнозних моделей для кожної незалежної змінної, в результаті чого найкращі результати продемонструвала LSTM, не зважаючи на гірший результат скоригованого коефіцієнту детермінації, середня квадратна похибка кореня показала кращий результат, що означає кращу здатність «підігнати» дані

набору даних для прогнозування. Щоб мати постійне уявлення про ймовірні кібератаки, результати повинні регулярно доповнюватися, оновлюватися для використання їх у реальних умовах, що дозволить вчасно реагувати на злочинні дії та попереджати їх виникнення.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**].

2.2 Мультисервісна модель комплексної оцінки та пріоритезації ризиків легалізації кримінальних доходів та кіберризиків

2.2.1 Теоретичні засади до розуміння сутності поняття «протидія легалізації доходів отриманих незаконним шляхом» в умовах діджиталізації суспільства

Незважаючи на динамічний розвиток суспільства та активні трансформаційні зміни глобальної економіки, реальний та фінансовий сектори досі включають офіційну та тіньову складову. В офіційній економіці всі господарські операції, що здійснюються економічними агентами знаходяться під наглядом та контролем держави. Кожен рух коштів регулюється нормативними актами, чітко відбувається в межах фінансової системи та підлягає оподаткуванню. Функціонування в офіційному секторі економіки дозволяє учасникам користуватись певними привілеями, а саме відсутністю адміністративних та кримінальних покарань, формування ділової репутації, а також реалізація тактичних і стратегічних цілей організації.

Поруч з офіційним сектором економіки існує тіньова економіка, тобто та, яка знаходиться поза полем зору держави. Відповідно до даних Міністерства економіки України, у 2021 році розміри тіньового сектору економіки України були на рівні 32% від обсягу офіційного ВВП [**Ошибка! Источник ссылки не найден.**].

Тіньовий сектор економіки привабливий для недобросовісних суб’єктів господарювання, які хочуть зменшити свої витрати та для злочинців, грошові

потоки яких прямо пов’язані з незаконною діяльністю (грабежі, крадіжки, обіг наркотиків, шахрайство, фінансування тероризму тощо).

Через тісну сплетеність зазначених секторів, у суб’єктів тіньової економіки постійно виникає необхідність надати фінансовим ресурсам, накопиченим у тіньовій економіці вигляду «чистих», офіційних. Тільки в офіційному секторі економіки є можливість отримувати товари і послуги найширшого спектру, оскільки в тіньовому секторі за нелегальні кошти можна отримати тільки нелегальний товар.

Значний обсяг тіньової економіки носить деструктивний характер для економічної безпеки національної економіки: від деформації податкової та бюджетної сфер, до неможливості побудувати адекватну макроекономічну політику, інвестиційний клімат, стабільну грошово-кредитну систему [**Ошибка! Источник ссылки не найден.**].

Легалізація доходів, отриманих незаконним шляхом спричиняє підвищення інфляції, зниження рівня валютної безпеки держави, підрив довіри до банківського сектору, зростання обсягу тіньової економіки [**Ошибка! Источник ссылки не найден.**] та зниження рівня довіри до країни на міжнародному ринку (через ризик потрапляння до чорного списку ФАТФ [**Ошибка! Источник ссылки не найден.**]). Через це держава прикладає значних зусиль для боротьби з легалізацією незаконних доходів.

Процес легалізації доходів, отриманих незаконним шляхом складається з трьох етапів: розміщення, розшарування та інтеграція.

Розміщенням прийнято вважати введення коштів, отриманих незаконним шляхом у офіційну систему фінансових операцій. Розміщення може бути реалізоване через придбання акцій чи інших цінних паперів, оформлення депозиту, здійснення банківського переказу тощо. Основна мета розміщення – вивести кошти від прямого зв’язку з безпосереднім злочином, внаслідок якого вони були отримані.

Розшарування коштів, раніше розміщених у офіційну систему фінансових операцій направлене на подальше приховування зв’язку з коштів та початковим

злочином. Реалізується через перепродаж цінних паперів, придбання чи продаж нерухомості чи інших товарів, переведення коштів за кордон чи на інші рахунки. Під час розшарування незаконні кошти можуть переплітатись із законними.

Коли незаконно отримані кошти настільки зливаються з офіційно отриманими, що вся сума виглядає законно отриманою говорять про настання етапу інтеграції незаконних доходів. Протягом цього етапу злочинці отримують на перший погляд законні підстави володіння активами. В такому випадку тільки ретельне розслідування дозволяє встановити зв’язок між інтегрованими коштами після розміщення та розшарування і незаконно отриманими внаслідок злочину.

Цифровізація фінансово-економічних відносин відкрила нові можливості для злочинців у сфері легалізації доходів, отриманих незаконним шляхом. З одного боку, цифровий слід грошей стало простіше відслідковувати, саме тому одним із напрямків протидії легалізації доходів є контроль та орієнтування на поступову відмову від готівки. З іншого боку, зловмисники можуть створювати десятки акаунтів у платіжних системах та робити сотні транзакцій не виходячи з дому. Генерація великих масивів даних в такому випадку логічно веде до застосування сучасних методів аналізу даних, заснованих на глибокому та машинному навчанні.

Методи легалізації незаконних доходів можуть бути різні: застосування недосконалості платіжних систем, створення та використання фіктивних підприємств, контрабанда, використання банківських переказів, укладання договорів псевдострахування, використання ломбардів та кредитних спілок, використання криптовалюти.

Проте, з огляду нашого дослідження, їх доцільно поділити на дві групи: ті які не зазнали значних змін від цифровізації відносин, та ті, які виникли внаслідок цифровізації або ж зазнали значної модернізації внаслідок неї.

До першої групи можна віднести:

1. Використання контрабандного способу легалізації доходів, отриманих незаконним шляхом полягає у махінаціях при декларуванні готівки,

дорогоцінних банківських металів. Вчиняючи злочинні дії, порушуючи митні правила, шляхом декларування у іншій країні готівки в іноземній валюті як особистих заощаджень та ухилення від декларування їх при проходженні митниці в Україні реалізується розшарування доходів. Заплутаність митних правил в різних країнах, помилки, халатність чи злочинний умисел під час проходження особою митного контролю, використання «обхідних шляхів» на державному кордоні підвищує ризик легалізації доходів, отриманих незаконним шляхом [**Ошибка! Источник ссылки не найден.**]. Безумовно митні органи значно збільшили власні інструменти контролю з розвитком цифровізації, так зараз сформовані значні бази даних перевірки митної вартості та інші системи контролю за товарами, проте в межах схеми легалізації ці зміни не здійснюють суттєвий вплив.

2. Окремим прикладом легалізації доходів, отриманих незаконним шляхом є використання страхових компаній. Зловмисниками здійснюється підробка страхових випадків, оформлення договорів псевдострахування, ухилення від оподаткування [**Ошибка! Источник ссылки не найден.**]. З точки зору діджиталізаційних процесів, які хоч і впливають на страхову сферу, проте сутнісно схеми легалізації не зазнали змін.

3. Використання ломбардів для обміну предметів розкоші та інших цінних активів на готівку та навпаки. Відповідно суттю схем є готівковий обіг, якого майже не стосується цифровізація.

До другої групи варто віднести:

1. Платіжні системи. За допомогою платіжних систем зловмисники здійснюють перерахунок коштів з рахунків на інші рахунки, в тому числі закордон, реалізуючи розшарування незаконних доходів. Використання мережі підставних осіб дозволяє зменшити розміри транзакції, уникаючи їх підозрілості, що ускладнює контроль та можливості виявлення фактів легалізації. Розвиток цифровізації спричинив зростання кількості платіжних систем та поширення доступу до них з будь-якої країни. Реєстрація гаманців у платіжних системах здійснюється за простішою ніж у банках процедурою.

2. Конвертаційні центри. Створення та використання конвертаційних центрів дозволяє зловмисникам перетворювати безготівкові кошти в готівкові, що спричиняє втрату їх цифрового сліду. Афілійовані з підприємствами особи, які мають право прийняття рішень, в разі їх залучення до процесів легалізації доходів, отриманих незаконним шляхом допомагають заплутувати походження коштів шляхом переуступки права вимоги по боргам, надання чи отримання благодійної допомоги. Цим самим замінюючи реальне походження коштів. Іншим способом використання конвертаційного центру – фіктивна господарська діяльність. Шляхом оформлення фіктивних платіжних доручень, отримання кредитів на діяльність, отримання коштів за державними цільовими програмами підтримки бізнесу здійснюється як розміщення, так і розшарування і інтеграція незаконних доходів [**Ошибка! Источник ссылки не найден.**].

Наразі, зареєструвати підприємницьку діяльність в Україні можна в режимі онлайн. Додатково до цього, цифровізація зумовила виникнення цілого ряду видів економічної діяльності, пов’язаної з інформаційними технологіями. Наприклад, розробка програмного забезпечення є по суті видом інтелектуальної діяльності, для підтримки якої не потрібно значного статутного капіталу чи основних фондів. Підприємства з орієнтацією на розробку програмного забезпечення підходять для легалізації незаконних доходів, оскільки легко імітувати факт надання послуг чи виконання робіт.

3. Банки. Банківські перекази мають бути під пильним контролем з точки зору протидії легалізації. Через банківські перекази найчастіше відбувається заплутування джерел походження коштів. Значна кількість транзакцій між фізичними особами з подальшим їх переказом чи виведення у готівку, перерахунок за роботи чи послуги фіктивного підприємства, придбання цінних паперів через банки є шляхами легалізації незаконних доходів через банківські установи. Розвиток мобільного банкінгу та фінтех прискорили процес розшарювання доходів, отриманих незаконним шляхом.

4. Віртуальні активи. Останнім часом серед злочинців часто стала використовуватись криптовалюта. Електронні кошти слугують засобом

розшарування, оскільки можливість створення багатьох криптогаманців за короткий період часу з мінімальною, на відміну від банку, індентифікацією особи. Здійснення тисяч транзакцій за допомогою сотень криптогаманців розмиває походження коштів і стає складно пов’язати кошти з первинним злочином [**Ошибка! Источник ссылки не найден.**].

Окремим підвидом можливих маніпуляцій за допомогою блокчейн-середовища є NFT. Унікальний в мережі, невзаємозамінний блокчейн токен (NFT) дозволяє надати цінності та закріпити авторське право за будь-яким цифровим об’єктом. Відтак, це дозволяє завищувати ціну на цифровий актив за допомогою аукціону та надавати злочинним коштам ознак легальності. Обсяг продажів NFT у 2020 році складав 250 млн дол. США, а у 2021 році уже понад 2 млрд дол. США [**Ошибка! Источник ссылки не найден.**].

Розуміючи це, криптобіржі вдаються до заходів протидії легалізації незаконних доходів, які базуються на встановленні ризиковості транзакції на основі аналізу співпраці особи з даркнетом чи сумнівними біржами.

Ускладнює роботу з протидії легалізації доходів, отриманих незаконним шляхом через цифрові активи недосконаліть регулюючого законодавства у цій сфері, що дає змогу вільно користуватись різними способами обігу криптовалюти.

У користувачів криптовалюти є декілька шляхів конвертувати цифрові гроші у гривні на банківський рахунок чи за допомогою готівки.

Обмінники криптовалют надають сервіс для швидкого обміну криптовалюти на фіатні кошти чи на інші криптовалюти. Діяльність таких обмінників не заборонена, проте ЗУ «Про віртуальні активи» ще не набрав чинності.

Крипто біржі дозволяють купувати та продавати різні види криптоактивів. В Україні діє декілька криптобірж, найбільші з них: KUNA, WhiteBIT, BTC TRADE UA, QMALL. Проте, через середовище інтернет та недосконаліть регуляції на міжнародному ринку крипто бірж поняття кордонів для віртуальних активів розмивається. Відтак, є можливість переводити грошові суми кому-

завгодно за кордоном. Криптобіржі дозволяють P2P обмін, за курсом, обговореним у онлайн-чаті, що дає ширші можливості для легалізації незаконних доходів.

Криптовалюти можна обмінювати через електронні платіжні системи, такі як PayPal, Perfect Money, Payeer, Portmone, Rupay. З березня 2022 року PayPal почав в Україні повноцінну роботу [**Ошибка! Источник ссылки не найден.**]. З одного боку це дозволило швидко переказувати кошти фізичним особам з-за кордону, оминаючи тривалий час очікування банківського переказу. А з іншого боку – надало можливість зловмисникам заплутувати шляхи переказу незаконних доходів: криптовалюта – електронна платіжна система – банківський рахунок.

Використання криптоматів дозволяє переводити готівку одразу в криптовалюту [**Ошибка! Источник ссылки не найден.**]. Таким способом готівка, яка найчастіше використовується злочинцями обмінюється в криптовалюту, з чого починається процес легалізації незаконних доходів.

Питання впливу цифровізації на легалізацію незаконних доходів та її протидію все частіше стає предметом наукових досліджень.

Таким чином, справедливо зробити висновок, що тіньовий сектор економіки активно розвивається паралельно із розвитком суспільства, а цифровізація значно вплинула на трансформаційні процеси легалізації кримінальних доходів збільшивши їх різноманітність та механізми реалізації. Виходячи з цього, державні органи влади повинні випереджаючими темпами розвивати власну систему протидії легалізації доходів одержаних незаконним шляхом, а також вивчати випереджаючи міжнародні практики та проводити їх імплементацію у національне законодавство.

Переходячи до дослідження теоретичної сутності поняття «протидія легалізації доходів отриманих незаконним шляхом» зауважимо, що даний процес запропоновано реалізувати з використанням програмного продукту VOSviewer. Отже, провівши бібліометричний аналіз наукових публікацій, що індексуються наукометричною базою даних Scopus за допомогою програмного

продукту VOSviewer розглянемо зв’язки між категоріями, які досліджують провідні науковці світу. На рисунку 2.62 представлена мапа зі взаємозв’язками поняття «протидія легалізації доходів, отриманих незаконним шляхом (anti-money laundering) з іншими категоріями, відповідно до публікацій за 2012-2022 роки. Дослідження дозволило виділити 6 основних кластерів, які на рисунку 2.62 відповідають зеленому, червоному, фіолетовому, блакитному, рожевому та коричневому кольорам, та 2 другорядних кластери: сірий та помаранчевий. Варто зауважити, що чим більше публікацій, які містять ключові слова з терміном «anti-money laundering», тим більший прямокутник за площею на рисунку 2.62 вони займають.

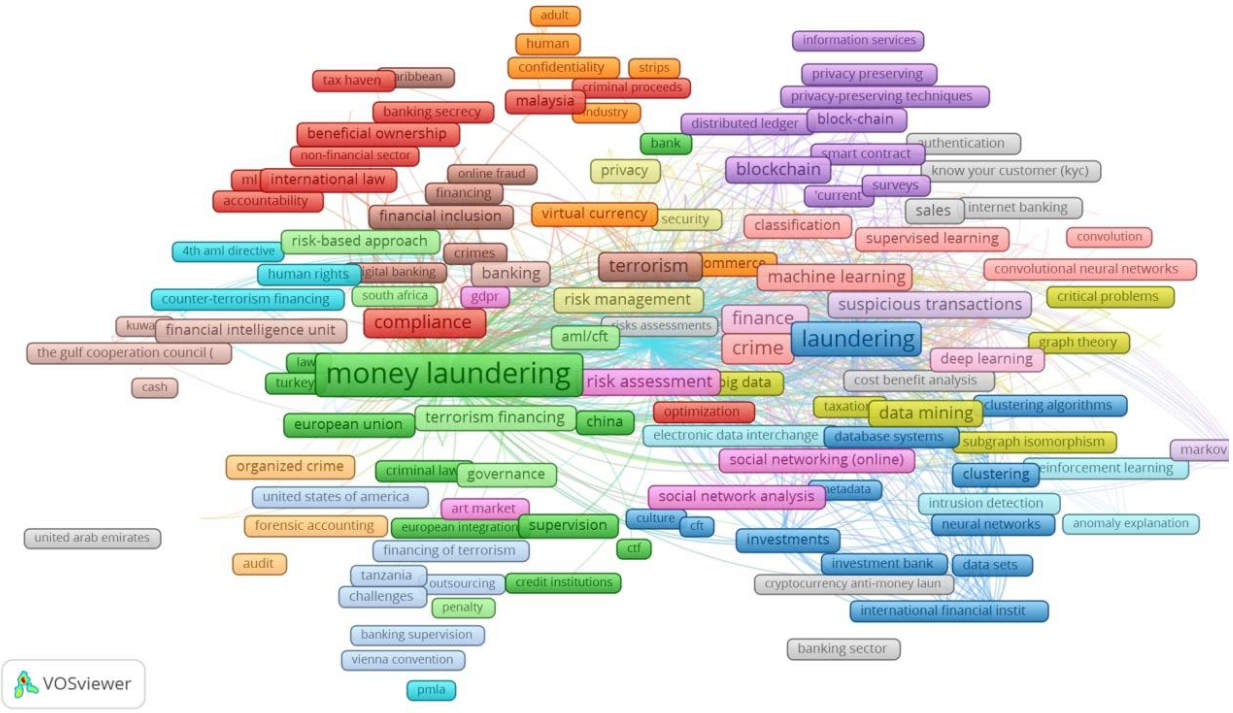


Рисунок 2.62 – Наукова бібліографія поняття «anti-money laundering» (протидія легалізації доходів, отриманих незаконним шляхом) за 2012-2022 роках

Відповідно до рисунку 2.62, можна простежити зв’язок поняття «протидія легалізації доходів, отриманих незаконним шляхом» з такими категоріями як легалізація грошей, легалізація, тероризм та фінансування тероризму, ризик менеджмент, банківський нагляд, злочин, комплаєнс. Значна частина зв’язків

присвячена економіко-математичним методам та машинному навчанню: нейронні мережі, кластеризація, ряди Маркова, бази даних, машинне навчання, великі дані.

Окремо, по фіолетовому кластеру можемо прослідкувати направленість досліджень, пов’язаних із сучасними інформаційними технологіями: блокчейн, розумний контракт, розподілений реєстр, криптовалюта та інформаційні сервіси.

Продовжуючи дослідження у часовому контексті, що відображений на рисунку 2.63, робимо висновок, що використання економетричних методів та моделей досліджувалось науковцями поряд з протидією легалізації доходів, отриманих незаконним шляхом, починаючи з 2012 року. Цьому свідчать фіолетовий колір блоків «clustering», «data sets», «neural networks», «database systems». А дослідження останніх років саме присвячені зв’язку криптовалюти та протидії легалізації незаконних доходів «cryptocurrency anti-money laundering», «block-chain», «distributed ledger», «smart contract», «art market». Зелено-жовтий відтінок свідчить, що особливу увагу цій тематиці почали приділяти у 2020-2022 роках. Окремо варто виділити наявність зв’язку в публікаціях присвячених протидії легалізації незаконних доходів та методами збереження конфіденційності «privacy-preserving techniques». Оскільки розвиток цифровізації з одного боку генерує велику кількість даних, які стають публічні в інтернеті, що порушує принципи конфіденційності, а з іншого боку, дотримання конфіденційності допомагає злочинцям реалізовувати схеми легалізації незаконних доходів. Тому значна частина науковців приділяють увагу питанням конфіденційності в мережі та протидії легалізації незаконних доходів.

Аналіз візуалізаційних карт, зображених на рисунку 2.62 та 2.63 підтверджує необхідність продовжувати дослідження у сфері взаємозв’язків між цифровими технологіями та протидією легалізації доходів, отриманих незаконним шляхом.

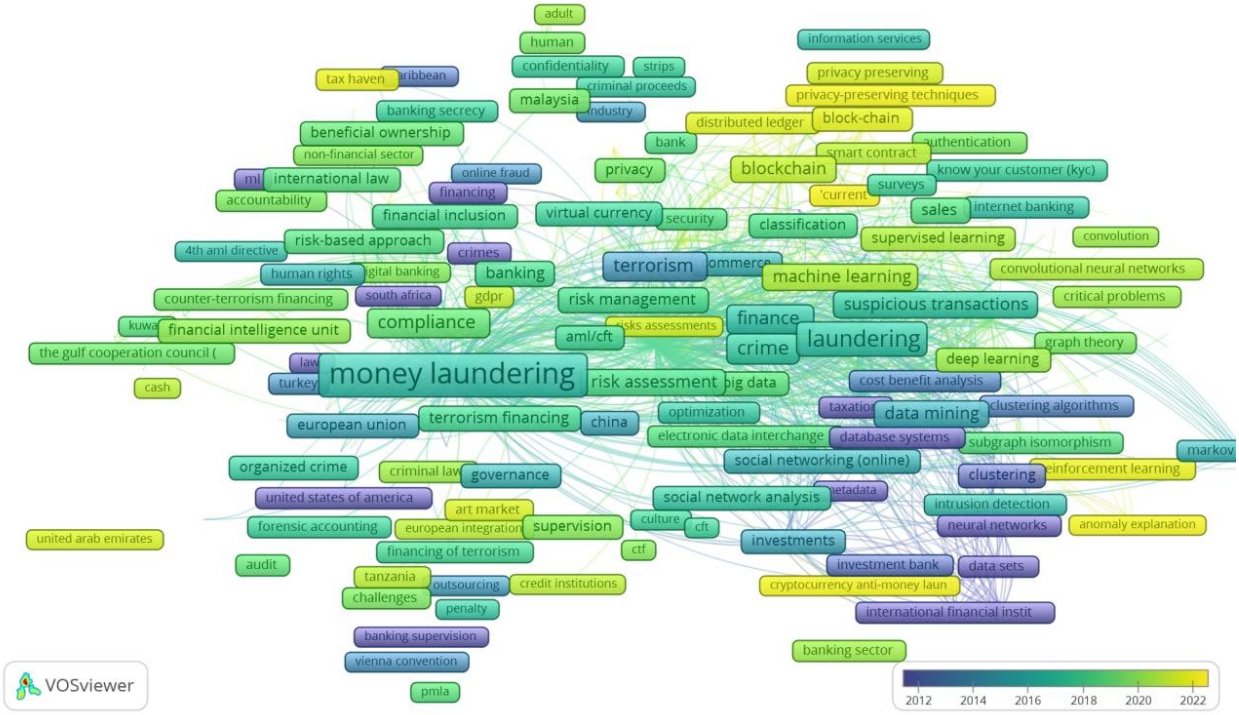


Рисунок 2.63 – Візуалізація часового виміру досліджень стосовно протидії легалізації доходів, отриманих незаконним шляхом, опублікованих у виданнях, що індексуються наукометричною базою даних Scopus у 2012-2022 роках

Зосереджуючись на безпосередньому взаємозв’язку цифровізації та протидії легалізації доходів, отриманих незаконним шляхом, можна виділити наступну залежність (рисунок 2.64): «anti-money laundering» – «digitalization» – «money laundering» – «cryptocurrency».

Рисунок 2.64 – Візуалізація взаємозв’язку протидії легалізації доходів, отриманих незаконним шляхом з цифровізацією відповідно до досліджень, опублікованих у виданнях, що індексуються наукометричною базою даних Scopus у 2012-2022 роках

Цифровізація найчастіше в наукових працях зустрічається у взаємозв'язку з протидією легалізації доходів, отриманих незаконним шляхом. Тоді як у контексті легалізації доходів, отриманих незаконним шляхом частіше зустрічається термін криптовалюта.

Переходячи до аналізу наукового доробку в сфері протидії легалізації доходів, отриманих незаконним шляхом варто зазначити, що дана тема є популярною як серед науковців-економістів, так і серед науковців у галузі юриспруденції.

Вплив легалізації доходів, отриманих незаконним шляхом на сталий розвиток досліджували Z. Dobrowolski та L. Sulkowski [**Ошибка! Источник ссылки не найден.**] та запропонували стійку модель боротьби з легалізацією доходів, отриманих незаконним шляхом через посилення аудиторського потенціалу органів фінансового контролю, і як наслідок покращення слідчих функцій парламентських наглядових органів. Визначення ризику легалізації доходів, отриманих незаконним шляхом для бізнес-сектору стало основою для дослідження J. Ferwerda та E.R. Kleemans [**Ошибка! Источник ссылки не найден.**], автори довели, що ризик легалізації доходів, отриманих незаконним шляхом відрізняється для різних секторів бізнес-діяльності та встановили, що для європейського простору найвищий ризик мають компанії, діяльність яких пов'язана з казино, готельним бізнесом та реалізацією об'єктів мистецтва.

Застосування методів data mining для протидії легалізації доходів, отриманих незаконним шляхом досліджували A. Salehi, M. Ghazanfari та M. Fathian [**Ошибка! Источник ссылки не найден.**]. Автори провели порівняльну характеристику побудованих моделей: багатонейронної мережі персептрона, імовірнісної нейронної мережі, радіально-базисної функції та лінійної нейронної моделі в якості класифікації транзакцій банку та обрали найкращу. Отримана модель дозволяє пришвидшити процес виявлення транзакцій, які мають ризик легалізації доходів, отриманих незаконним шляхом. Натомість A.I. Canhoto [**Ошибка! Источник ссылки не найден.**] досліджувала можливості використання методів машинного навчання для протидії легалізації доходів,

отриманих незаконним шляхом. У дослідженні автор справедливо зауважує, що основна проблема застосування методів математичного моделювання для протидії легалізації незаконних доходів це відсутність якісних наборів даних. У доступі до науковців є дані, які дозволяють виявити нетипову фінансову поведінку особи. Але дані, які напряму пов'язують транзакції, здійснені фінансовими інституціями та факти легалізації грошей відсутні. Оскільки фактичні дані про легалізацію знаходяться у органів слідства різних країн, які є конфіденційними.

Переходячи до дослідження наукових робіт присвячених впливу цифровізації на роботу суб'єктів системи протидії легалізації кримінальним доходам, зазначимо, що безумовно діджиталізація підвищує якість фінансового моніторингу в банках та сприяє ефективнішому впровадженню рекомендацій FATF у сфері протидії легалізації. Автори [**Ошибка! Источник ссылки не найден.**] запропонували схему вдосконалення процесу фінансового моніторингу в банку, а також сформували пропозиції щодо проведення фізичних заходів для підвищення рівня дотримання законодавства у сфері протидії легалізації доходів, отриманих незаконним шляхом. До схожого висновку дійшли і K. Said та D.Karimi [**Ошибка! Источник ссылки не найден.**], зазначаючи що автоматизація банківських процесів підвищує фінансову безпеку банку та покращує фінансовий моніторинг. Особливо вагомий вплив діджиталізації на легалізацію незаконних доходів прослідковується через зменшення кількості готівки в обігу. Автори [**Ошибка! Источник ссылки не найден.**] доходять висновку, що готівка залишається основним інструментом прискорення легалізації незаконних доходів, а впровадження технологій безготівкового розрахунку не дає поки що очікуваного ефекту у вигляді зниження обсягів легалізації незаконних доходів використовуючи готівкові потоки. Автори встановили пряму залежність попиту на готівку та тіньової економіки.

Останні дослідження свідчать що фінтех широко впроваджується в банках в країнах що розвиваються [**Ошибка! Источник ссылки не найден.**], що дозволяє набагато швидше приєднатись фінансовим установам даних країн до

глобальної системи протидії легалізації кримінальним доходам. Значна група науковців – K. Djalilov та J. Hölscher [**Ошибка! Источник ссылки не найден.**]; K. Djalilov та C. Hartwell [**Ошибка! Источник ссылки не найден.**] дійшла висновку про те, що впровадження цифрових технологій в комплексі покращує можливості банківського середовища та безумовно впливає на можливість оперативного виявлення незаконних операцій. Окрім зазначеного, не можна нехтувати тим, що цифровізація пришвидшує роботу співробітника відповідного органу системи протидії легілації: інтерактивний пошук замінює довге перелистування папірців, тобто економляться людино-години роботи [**Ошибка! Источник ссылки не найден., Ошибка! Источник ссылки не найден.**].

A. Addo та PK. Senyo [**Ошибка! Источник ссылки не найден.**] у своїх дослідженнях зупиняються на вивченні впливу цифровізації на правоохоронну діяльність, а саме боротьбу з корупцією. Так, вони визначають, що цифровізація є потужним антикорупційним інструментом. Приклади застосування цифрових технологій у прокурорській діяльності досліджував Denis de Castro Halis [**Ошибка! Источник ссылки не найден.**], автор зазначає, що засоби цифровізації одночасно з підвищенням швидкості обробки документів формують й засади до їх безпосереднього контролю, забезпечуючи незалежність слідства та відповідність його законам. Окремо варто розглянути застосування сучасних технологій у судовій діяльності. Автор Yu. Mulyana [**Ошибка! Источник ссылки не найден.**] зазначає, що цифровізація судів починається з електронного документообігу, але не обмежується ним: всі процеси пов’язані зі справами можуть обслуговуватись в електронному суді, або наприклад суди можуть проводитись у онлайн-форматі, що пришвидшує процес.

Легалізацію доходів, отриманих незаконним шляхом за допомогою криптовалюти досліджував C. Wronka, зосереджуючись не тільки на аналізі сутності легалізації, а й можливих превентивних заходів їй протидії [**Ошибка! Источник ссылки не найден.**]. Автор визначив, що віртуальні активи становлять значну загрозу легалізації незаконних доходів. А як заходи протидії легалізації пропонується введення на законодавчому рівні ліцензування

діяльності з видачі криптогаманців, цим самим покладаючи відповідальність за легалізацію незаконних доходів на постачальника криптогаманців. За такого підходу з’явиться можливість регулювати створення та використання криптогаманців. А з іншого боку, постачальники криптогаманців будуть прискіпливіше ставитись до процесу ідентифікації та аутентифікації клієнтів, що однозначно позитивно вплине на протидію легалізації незаконних доходів за допомогою криптовалюти. D. Dupuis та K. Gleason займались питанням необхідності регулювання обігу криптовалюти, в контексті протидії легалізації незаконних доходів [**Ошибка! Источник ссылки не найден.**], визначаючи наявні шляхи обміну криптовалюти на фіатні кошти, що надає змогу злочинцям легалізувати кошти, отримані незаконним шляхом.

Отже, справедливо зробити висновок, що в сучасному науковому середовищі протидії легалізації доходів отриманих незаконним шляхом в умовах діджиталізації суспільства відводиться значна роль та розглядають в своїй більшості як комплекс інформаційно-технологічних заходів з попередження, виявлення та подальшого покарання злочинних дій, спрямованих на маскування незаконного походження коштів або іншого майна чи володіння ними, а також набуття, або використання коштів чи іншого майна з метою надання правомірного вигляду їх використанню або поширенню чи дій, спрямованих на приховування джерел їх походження, а також вчинення з такими коштами або іншим майном фінансової операції за умови усвідомлення особою того, що вони були одержані злочинним шляхом.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден., Ошибка! Источник ссылки не найден.**].

2.2.2 Оцінка ризику конвергенції системи протидії відмивання грошей та кібербезпеки

З кожним роком роль цифрових технологій у різноманітних сферах життя зростає швидкими темпами. Їх запровадження та вдосконалення відбувається не лише на рівні конкретного підприємства, фізичної особи чи організації, але й на рівні держави. В цифровому просторі розміщено велику кількість даних, такі як особисті документи громадян, їх персональна інформація, що знаходиться, як у відкритих джерелах (соціальні мережі), так і закритих (дані міграційних служб, реєстрів відділів соціального захисту, податкових служб, банківських установ і т.д.), фінансова звітність підприємств, їх установчі документи, тощо. Наявність такого простору вимагає досить високого рівня захисту даних. І в ідеальному середовищі цей захист має відбуватися не лише окремо в обмеженій установі, але й охоплювати більш широкий спектр – рівень країн та міжнародних спільнот. Кібербезпека є не лише однією із найважливіших складових національної безпеки країни, але й закладає основи міжнародних відносин між державами по всьому світу.

Оцінка ризиків, пов’язаних з кіберзагрозами, на сьогодні є одним із пріоритетних напрямків діяльності теоретиків та практиків. Проведення аналітики щодо їх виявлення та усунення є надзвичайно складним завданням, оскільки дана група ризиків є динамічною. Ризики такого типу здатні модифікуватись та адаптуватися до змін системи ще задовго до виникнення можливості їх передбачення та усунення. А методи, що діють для одного кіберризику, можуть бути зовсім недієвими для іншого, тому навіть найменший ризик може спричинити значні втрати і призвести до краху цілої системи. За останні роки рівень та масштаби кіберзлочинів невпинно зростають, а збитки від них значно переважають збитки від торгівлі наркотиками та зброєю, чого не спостерігалося ще п’ять років тому. Банківські та фінансові установи останнім часом найбільше потерпають саме від кіберзлочинців, ніж, наприклад, від зміни на фондових біржах чи боргових криз. Це все зумовлює потребу в удосконаленні системи боротьби проти різного виду кібернетичних шахрайств. Тому методи боротьби із кіберзлочинністю розвиваються на законодавчому рівні багатьох країн, що дозволяє підтримувати загальний рівень безпеки країни.

В сучасному світі темі вивчення систем протидії кібернетичним злочинам та фінансовим махінаціям присвячено багато праць, як вітчизняних, так і закордонних вчених. Так, однією із найбільш ґрунтовних робіт вважається робота М. Еллінга, який досліджував кіберризики та можливості їх протидії [**Ошибка! Источник ссылки не найден.**]. Важливу роль у дослідженні кібербезпеки та її взаємодії із фінансовою системою вивчали Ю. Кожедуб, К. Семенова та інші [**Ошибка! Источник ссылки не найден.**]. Значний внесок у дослідження кіберризиків мають приватні консалтингові компанії, які мають на меті практичне застосування напрацювань. Серед них варто виділити Deloitte, AON, IBM, тощо [**Ошибка! Источник ссылки не найден.**]. Саме ці компанії в останні роки сформували велику базу знань, що допомагає не лише науковцям, але й підприємствам та урядам у власній діяльності.

Національний індекс кібербезпеки (NCSI)– один із збірних показників, який є одним із основних характеристик стану інформаційної безпеки [**Ошибка! Источник ссылки не найден.**]. Для розрахування даного індексу враховується значна кількість показників. Найважливішим серед них є стан законодавства, яке пов’язане з охороною даних та кібербезпекою, оскільки саме юридичне забезпечення дозволяє підготувати основу для реалізації стабільних заходів щодо протидії злочинам. Враховується також частота виникнення кіберінцидентів, рівень освіти громадян в сфері кібербезпеки, види та ефективність заходів щодо захисту персональних даних громадян, щодо реагування на кібератаки та можливості зниження кіберризиків. Даний індекс включає в себе результативність та кількісне виявлення загального рівня боротьби з кіберзлочинністю. Саме цей індекс враховується при оцінці інвестиційної привабливості країни в цілому та є відображенням стабільності та надійності її інформаційної системи.

Поряд із забезпеченням сталого розвитку кібербезпеки та підтримки її на належному рівні значну роль в утворенні та формуванні національної безпеки є фінансова. Фінансова безпека є підґрунтям для побудови міцної економічної безпеки та конкурентоспроможної економіки в цілому. Основним завданням

фінансової безпеки є побудова сприятливого середовища, як правового, так і економічного, а також підтримка інституційної інфраструктури, яка дозволить стимулювати розвиток потенційно життєздатних підприємств та запустить інвестиційні процеси, що допоможуть підтримувати загальний фінансовий рівень держави.

Фінансова система є однією із найбільш схильною до негативних впливів внутрішнього та зовнішнього середовищ, а також є вразливою до більшості видів ризиків. Великим дестабілізуючим фактором, що знижує рівень фінансової безпеки є дестабілізація бюджету України, особливо в нинішніх умовах, коли він розподіляється нерівномірно і має значний дефіцит. Великою проблемою є збільшення грошового обігу поза банківською системою країни, особливо в операціях з валютою. Це свідчить про недовіру до внутрішнього ринку країни та сприяє збільшенню рівня тінізації економіки та відмиванню коштів. Відмивання коштів та подальша легалізація таких доходів є великою проблемою не лише для України, але й для багатьох країн, що розвиваються. Розвиток технологій та вдосконалення кібератак дозволило злочинцям знаходити нові методи у відмиванні коштів та отримувати прибуток, використовуючи не лише традиційні гроші, але й криптовалюту та інші види електронних грошей. Це значно підвищило потенційні ризики та можливі збитки від їх настання. Тому оцінка фінансової стабільності та методів протидії відмиванню коштів є важливою складовою національної безпеки Так, основним індексом для оцінки є індекс протидії відмиванню коштів (AML–індекс) [**Ошибка! Источник ссылки не найден.**]. Вперше даний показник був розрахований та застосований на базі Базельського інституту управління у 2012 році. Даний інститут є незалежним асоційованим інститутом Базельського університету та діє під керівництвом Програми Організації Об’єднаних Націй в сфері попередження злочинності та кримінального правосуддя [**Ошибка! Источник ссылки не найден.**]. Він застосовується для виміру та оцінки ризиків, що пов’язані з відмиванням коштів та спонсоруванням тероризму. Результати всіх досліджень є незалежними та використовуються різноманітними державними, міжнародними, комерційними,

фінансовими установами для оцінки можливостей настання ризиків та способів їх уникнення. Даний індекс є корисним, оскільки він враховує не рівень корупції та кримінальної діяльності, що пов’язана з відмиванням коштів та фінансуванням тероризму, а більше ризики їх виникнення та розвитку. Даний індекс включає в себе різні показники, що мають різну спрямованість та значимість в оцінці. Тому показники, що застосовуються при визначенні індексу, розподіляються по основним категоріям та мають свою вагу. Це дозволяє отримати всебічну картину та впроваджувати подальші заходи для мінімізації чи усунення зазначених ризиків.

Якість системи протидії відмиванню коштів та фінансування тероризму складає 65 % від загального індексу. Тут враховуються політичні, правові, соціально-економічні методи, спрямовані на забезпечення та протидію даному виду злочинності. Ризик корупції у державі складає 10 %, фінансова прозорість також 10%, а загальна прозорість та підзвітність має 5 % . Ще 10 % становлять правові та політичні ризики, оскільки саме від них залежить стан фінансової системи щодо тероризму та усі можливі наслідки такої діяльності. Зазначений індекс охоплює усю економічну, політичну та фінансову систему, що дозволяє отримати чіткі та показові результати, необхідні для подальшого розвитку країни. Отже, можна зазначити, що для кожної країни Національний індекс кібербезпеки та AML – індекс мають власне значення і базуються на власних показниках [**Ошибка! Источник ссылки не найден.**]. Але в загальному підсумку деякі країни мають схожі тенденції до розвитку і мають загальні риси в підтримці сукупної національної безпеки. Необхідно проаналізувати декілька країн і визначити, чи дійсно існують великі розбіжності між схожими країнами і на чому саме ґрунтується можлива подібність.

Для аналізу в даному випадку використовуються методи кластеризації, оскільки саме вони дозволять отримати групи країн, які матимуть подібні дані та схожі риси. Silhouette analysis використовується для визначення відстані поділу між отриманими кластерами. Він відображає, наскільки близько розташована кожна точка в одному кластері відповідно до інших кластерів. Показники даного

аналізу мають діапазон від -1 до 1. Коефіцієнти біля 1 вказують на те, що вибірка знаходиться далеко від сусідніх кластерів. Тобто розподіл є найбільш чітким і придатним для аналізу. Значення 0 відображає, що вибірка знаходиться на межі прийняття рішення між двома сусідніми кластерами або дуже близько до неї, а від'ємні значення вказують на те, що ці вибірки могли бути призначені неправильному кластеру [**Ошибка! Источник ссылки не найден.**].

Після проведення оцінки Silhouette для здійснення кластеризації індексом AML (рис. 2.65), можна зробити висновок, що найбільш оптимальною кількістю кластерів є вісім, оскільки оцінка в даному випадку є найвищою. Тому для проведення кластеризації методом K-means необхідно використати саме таку кількість кластерів.

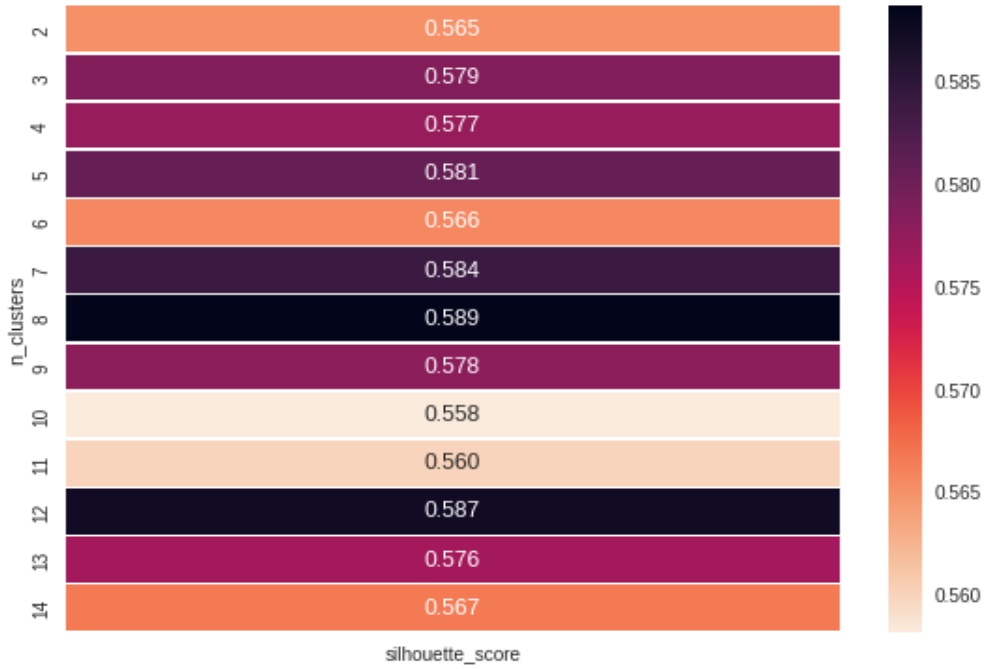


Рисунок 2.65 – Оцінка «Silhouette» щодо оптимального вибору кластерів країн за індексом AML

Рисунок 2.66 показує розподіл даних по кластерам в результаті Silhouette-оцінювання. Результати розподілені за кластерами із відсутністю вийнятків, то можна сказати, що всі дані кластерів є однорідними.

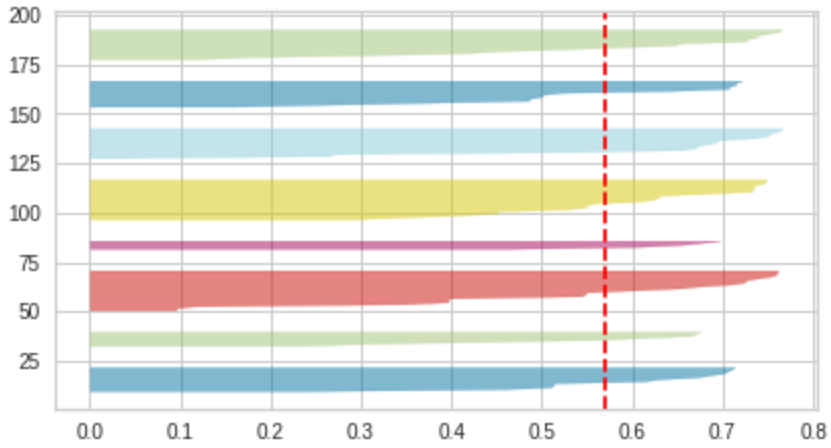


Рисунок 2.66 – Результати Silhouette-analysis

На рисунку 2.67 представлений результат проведеної кластеризації країн за індексом AML. Розподіл на кластери дозволив отримати досить показові групи, які надають змогу зробити певні оцінки. Як можна побачити, країни, які є достатньо розвиненими та мають високі показники в забезпеченні фінансової безпеки, віднесені до однієї групи. Так, найкращі показники мають такі країни, як Австралія, Велика Британія, Данія, Норвегія, Фінляндія, Нова Зеландія Греція, Ізраїль, Франція, Литва, Словенія, Ісландія та Швеція. Вони мають високий рівень протидії відмиванню коштів і фінансування тероризму та впроваджують ефективні рішення у боротьбі з легалізацією кримінальних доходів. Вони не є привабливими для злочинців, оскільки мають потужні системи захисту у фінансових установах.

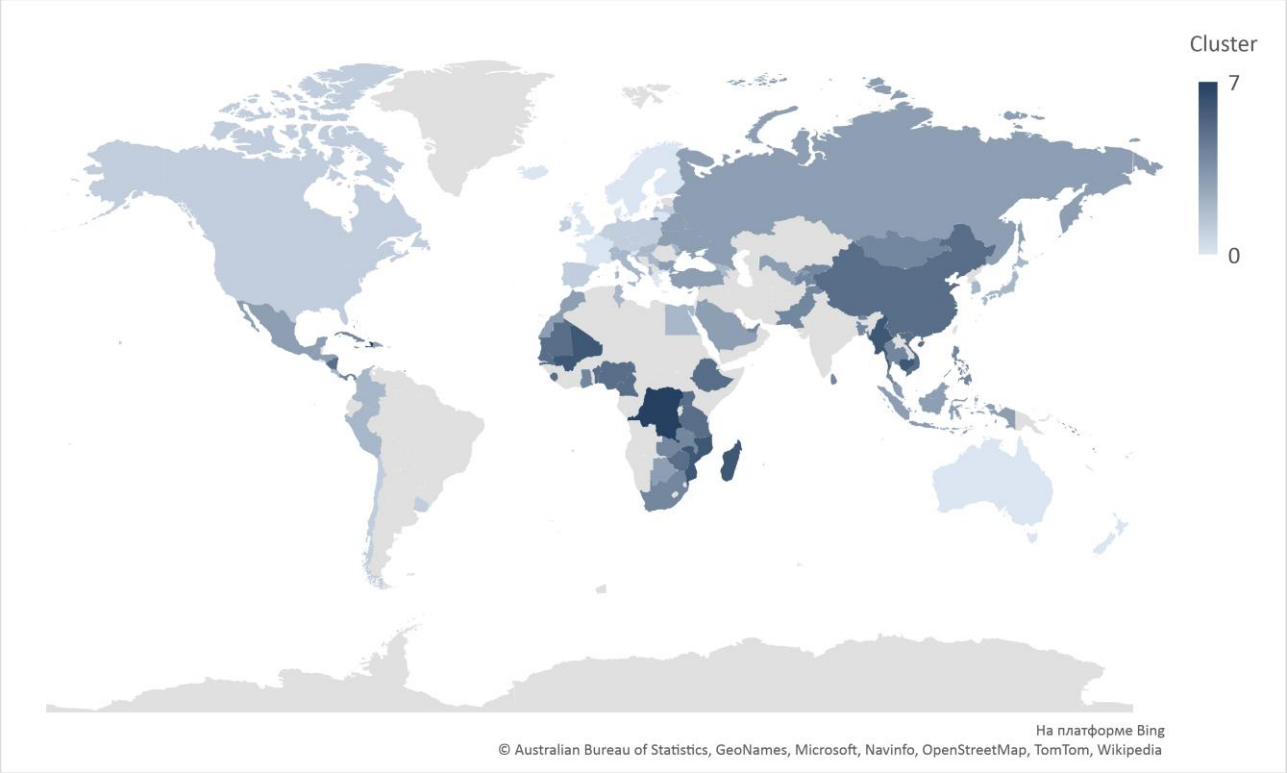


Рисунок 2.67 – Карта кластерів країн за індексом AML

Україна входить до третьої групи за показником ризиковості щодо відмивання кримінальних доходів. Це обумовлено тим, що наша країна має досить високий рівень тінізації та корумпізації економіки, що в сукупності з військовими діями створює сприятливе підґрунтя для зменшення ризиків для легалізації нелегальних коштів. Ці процеси гальмують розвиток в економіці та соціальній сфері. До даної групи увійшли ще 21 країна, серед яких слід зазначити Гондурас, Туреччину, Сейшельські острови, Барбадос, Ямайку та інші.

Країни, які є сприятливими для легалізації кримінальних доходів є країни 6 та 7 кластерів. Сюди відносяться Малі, Камбоджі, Мадагаскар, Мозамбік, М’янма, Гаїті та Демократична республіка Конго. Перелічені країни відносяться до найменш розвинених. Реформи, що проводяться в них, спрямовані на внутрішні сфери і не приносять відповідних позитивних результатів для їх розвитку. Режими цих країн є досить обмежувальними, а рівень розвитку економіки нестабільним.

Розглянемо результати кластеризації щодо «Національного показника кібербезпеки» (рис. 2.68). Найвище значення оцінки відповідає кількості

кластерів, яка дорівнює двом. Але дана кількість кластерів не дозволяє виявити більш детальні групи країн, тому для проведення Silhouette-analysis оберемо кластери з найвищими оцінками – 4, 5, 6, 7.

Рисунок 2.68 – Оцінка «Silhouette» щодо оптимального вибору кластерів країн за індексом NSCI

Результати Silhouette-analysis (рис. 2.69) показують, що кластеризація країн із використанням 4, 5, 6 або 7 кластерів дозволить отримати групи з однорідними даними. Тому, оберемо кількість кластерів 6, оскільки цьому значенню відповідає найвища оцінка Silhouette.

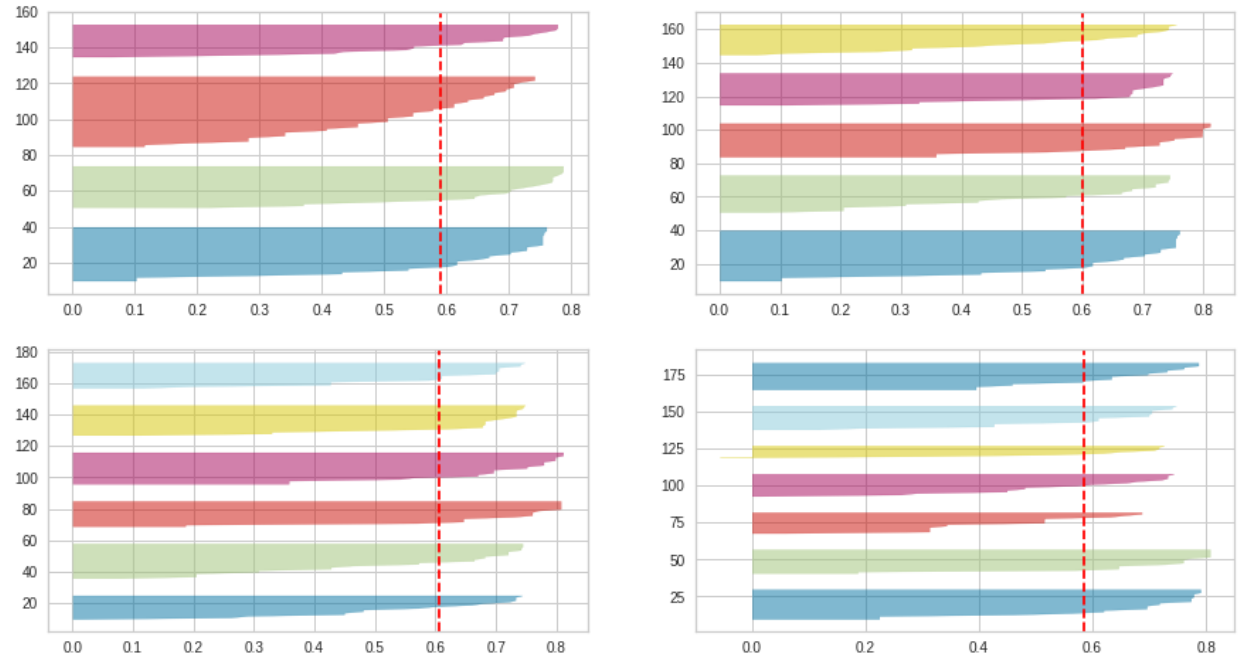


Рисунок 2.69 – Результати Silhouette-analysis

На рисунку 2.70 представлений результат проведеної кластеризації країн за індексом NSCI після проведеного Silhouette-analysis.

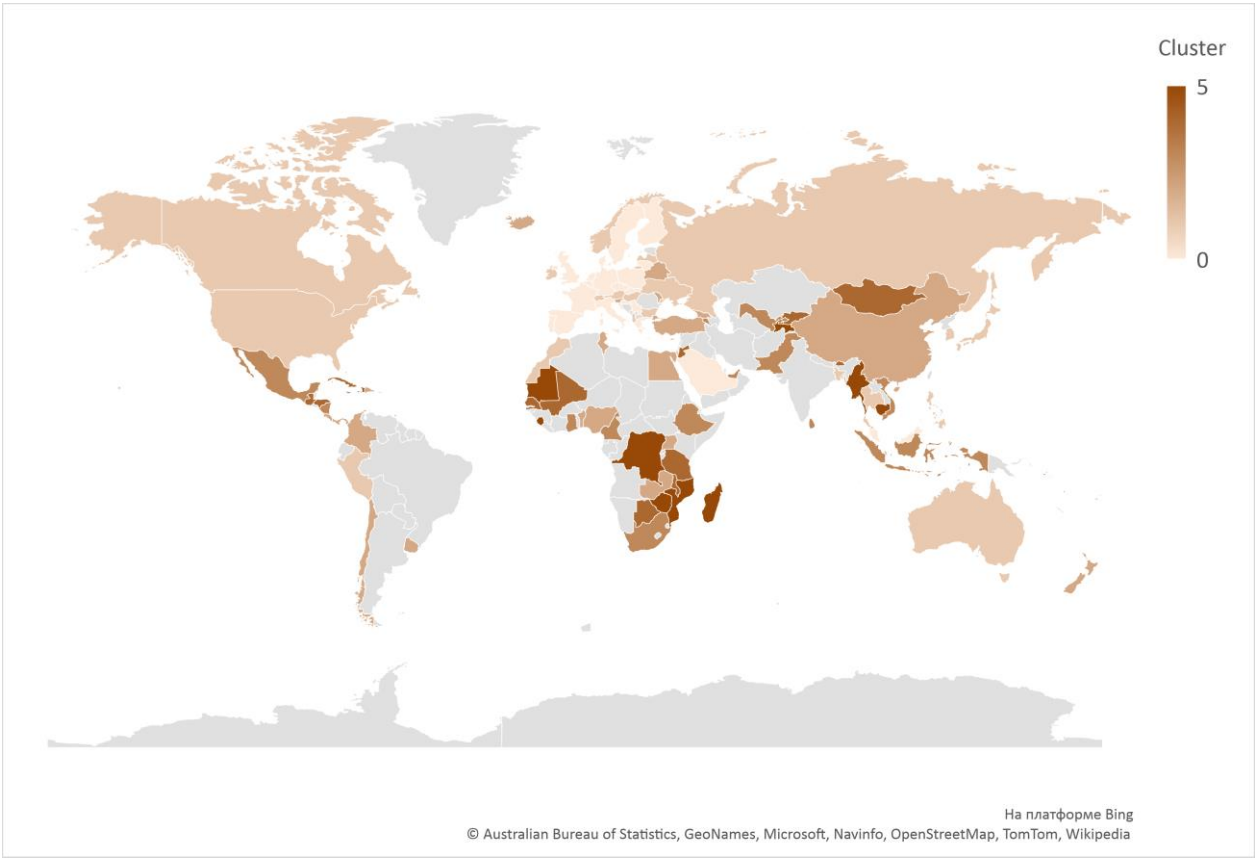


Рисунок 2.70 – Карта кластерів країн за індексом NSCI

Як можна побачити, до групи з найбільш високо розвиненим рівнем кібербезпеки відносяться країни нульового кластеру: Греція, Бельгія, Нідерланди, Німеччина, Іспанія, Малайзія, Саудівська Аравія, Сербія, Хорватія, Італія, Польща, Словаччина, Португалія, Чехія, Велика Британія, Данія, Франція, Литва, Швеція та Фінляндія. Такі результати свідчать, що ці країни мають високий рівень національної кібербезпеки, який передбачає організацію потужного комплексу інформаційного, програмного, технічного та організаційного забезпечення з питань кіберзахисту. Так само, можна побачити, що найнижчі показники мають Конго, Мадагаскар, Мозамбік, Гаїті, Камбоджі, Зімбабве, Таджикистан, Вануату, тощо. Тобто ці країни є менш розвиненими і потребують вкладень і модифікації не лише окремої системи кіберзахисту, але прогресивних державних заходів в цілому щодо розвитку її політичної та соціально-економічної сфер.

Для визначення ризиків конвергенції системи протидії фінансовим та кібершахрайствам доцільно впровадити комплексну оцінку, яка б в собі поєднувала можливості країн щодо кіберзахисту від різного роду загроз та їх потенціал щодо зниження ризиків відмивання кримінальних доходів та фінансування тероризму. Пропонуємо визначити інтегральний індекс конвергенції наступним чином:

– здійснюється нормалізація AML – індекса за критерієм Севіджа, як дестимулятора (формула 2.14):

$$x_{ij}^{*} = \frac{x_j^{max} - x_{ij}}{x_j^{max} - x_j^{min}} \tag{2.14}$$

– здійснюється нормалізація NSCI за природньою нормалізацією як стимулятора (формула 2.15):

$$x_{ij}^{*} = \frac{x_{ij} - x_j^{min}}{x_j^{max} - x_j^{min}} \tag{2.15}$$

– інтегральний індекс конвергенції утворюється як середньогеометричне (формула 2.16):

$$G_m = \left(\prod_{i=1}^{n} \widetilde{x_{ik}} \right)^{1/n} \tag{2.16}$$

Розглянемо результати кластеризації щодо «Інтегрального показника конвергенції» (рис. 2.71). Найвище значення оцінки відповідає кількості кластерів, яка дорівнює двом. Але така кількість кластерів не дозволить отримати уявлення про ризик конвергенції. Тому для проведення Silhouette-analysis оберемо кластер з наступною найвищою оцінкою – 9.

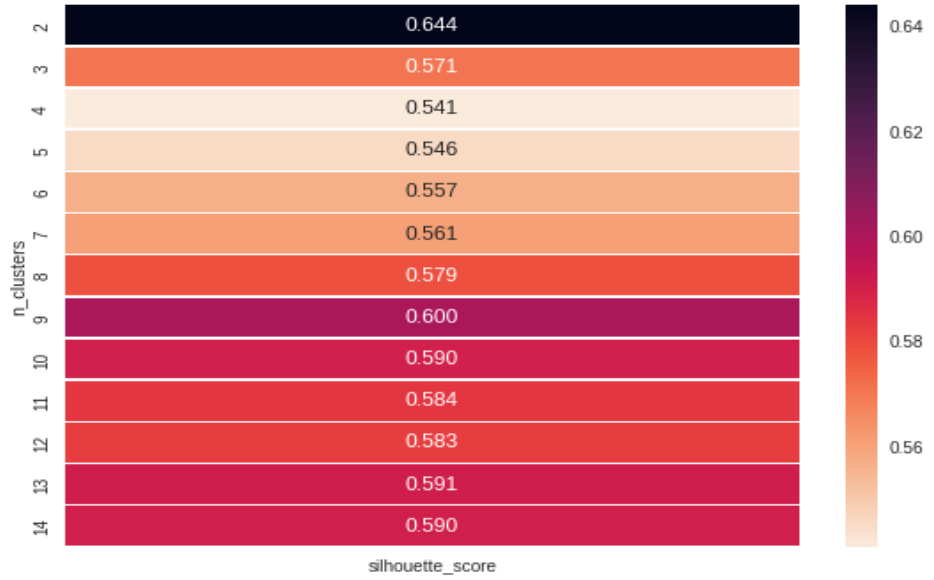


Рисунок 2.71 – Оцінка «Silhouette» щодо оптимального вибору кластерів країн за умови конвергенції системи протидії відмивання грошей та кібербезпеки

Результати Silhouette-analysis (рис. 2.72) показують, що кластеризація країн із використанням 9 кластерів дозволить отримати групи з однорідними даними, що дозволяє провести кластерний аналіз за методом k-means.

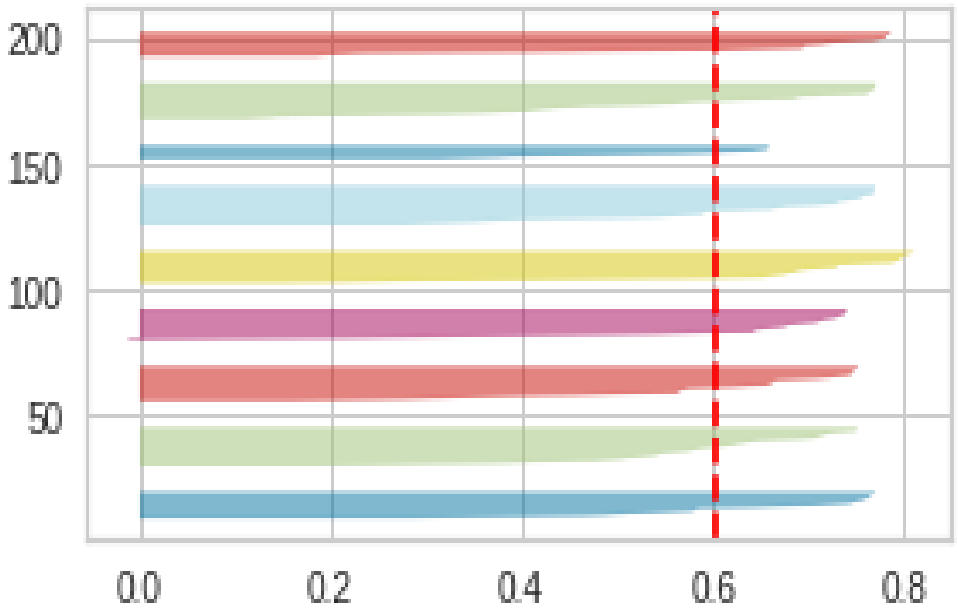


Рисунок 2.72 – Результати Silhouette-analysis

На рисунку 2.73 представлений результат проведеної кластеризації країн за «Інтегральним індексом конвергенції» після проведеного Silhouette-analysis.

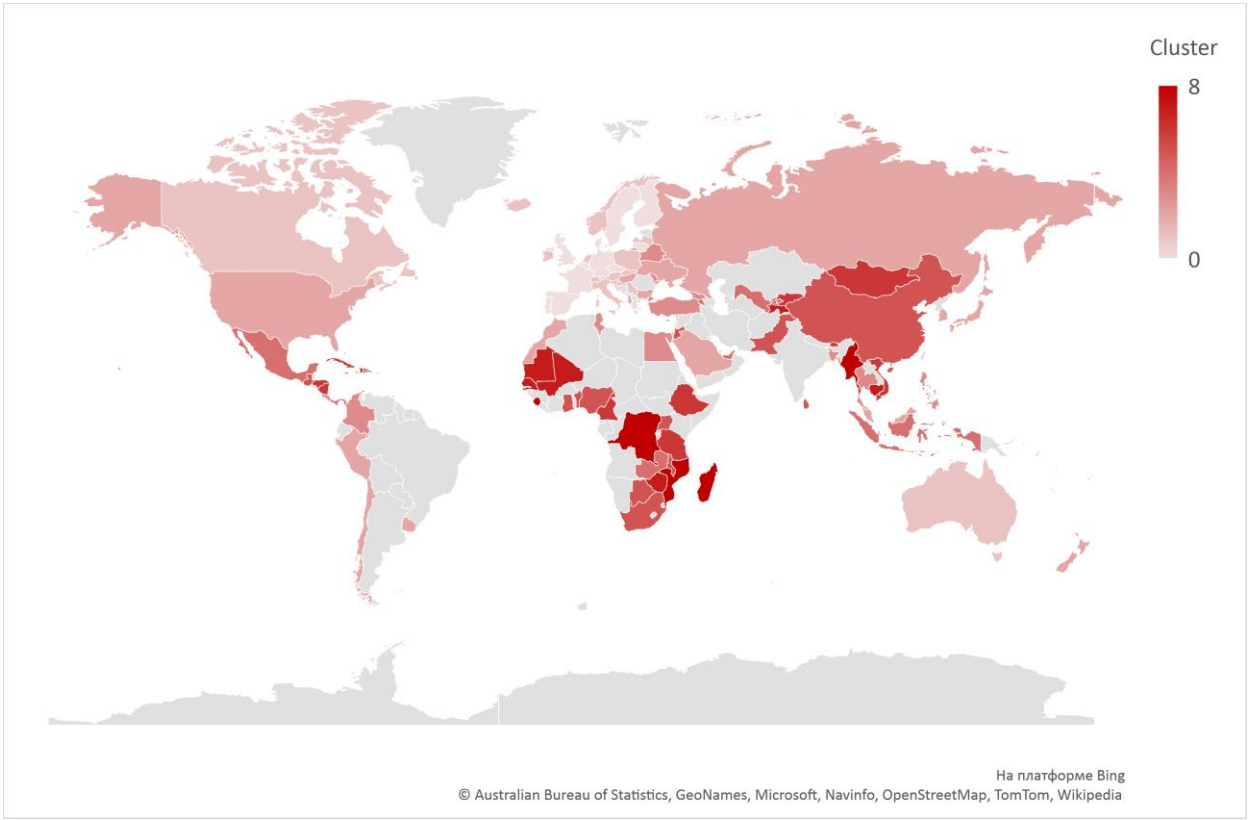


Рисунок 2.73 – Карта кластерів країн за рівнем конвергенції систем протидії відмиванню кримінальних доходів та кібербезпеки

Проведення кластеризації на основі інтегрального індексу конвергенції систем показує розподіл країн за можливостями протидіяти кіберзагрозам та фінансовим злочинам. Чим ближче його значення до 1, тим нижчий рівень ризику конвергенції, тобто в країнах створені сприятливі умови, які дозволяють інтегрувати систему кіберзахисту та систему фінансового моніторингу. Якщо значення даного показника наближається до 0, то це свідчить про неготовність країни до конвергенції двох систем, що може бути викликано сформованими сприятливими умовами для розвитку фінансової та кіберзлочинності.

Виходячи з отриманих даних сформуємо критерії ризику в залежності від отриманих кластерів для індексу конвергенції. Результати представлені в таблиці 2.5. Таблиця містить усереднені значення AML та NSCI для відповідного кластеру. Країни, що відносяться до 0-го кластеру генерують найнижчий ризик конвергенції. Відповідно, країни 8-го кластеру генерують найвищий ризик. Застосування даних карти 2.70 та таблиці 2.5 дозволить сформувати висновки щодо потенційних ризиків конвергенції систем протидії фінансовим та кіберзлочинам.

Таблиця 2.5 – Ідентифікація ризику в залежності від кластерів індексу конвергенції

Низький ризик		Помірний ризик		Високий ризик	
Кластери	AML / NSCI	Кластери	AML / NSCI	Кластери	AML / NSCI
0	3,65 / 88,42	3	5,09 / 55,95	6	6,07 / 22,94
1	4,09 / 72,96	4	5,30 / 41,43	7	6,30 / 12,51
2	4,72 / 67,61	5	5,83 / 37,29	8	7,75 / 9,31

Карта кластерів 2.70 показує, що найбільш сприятливі умови для конвергенції сформовані в 12 країнах, таких як Німеччина, Бельгія, Португалія, Іспанія, Чехія, Греція, Велика Британія, Данія, Франція, Литва, Швеція, Фінляндія. Ці країни генерують найнижчий ризик. Що стосується України, то її було віднесено до 2-го кластеру. У даному випадку вона має високий ризик легалізації кримінальних доходів, що компенсується розвиненим рівнем кібербезпеки. Це дозволяє зменшувати ризики відмивання коштів за рахунок

можливостей системи кіберзахисту. Тобто Україна має значний потенціал для створення взаємодії фінансової та інформаційної систем та підтримки їх безпеки. Рівень її системи відмивання коштів та легалізації доходів значно може скорочуватися за рахунок безпекового потенціалу.

В країнах, що віднесені до 6-8 кластерів, сформовані найбільш несприятливі умови, оскільки вони знаходяться на нижчій стадії економічного і соціального розвитку. Сюди відносять найменш розвинені країни Африки, Азії та острівні держави.

На основі отриманих даних побудуємо прогнозну модель ризику конвергенції, яка дозволить визначити відповідний його рівень за рахунок зміни умов конвергенції системи кібербезпеки та фінансового моніторингу. Для побудови класифікаційного дерева рішень було використано мову програмування Python. Модель було визначено на основі коефіцієнту Джині.

Результат класифікаційної моделі представлений на рисунку 2.74. Оцінка якості побудованого дерева представлена на рисунку 2.75.

```
Confusion Matrix:
[[4 0 0 0 0 0 0 0 0]
 [2 3 0 0 0 0 0 0 0]
 [0 2 0 0 0 0 0 0 0]
 [0 0 0 2 0 0 0 0 0]
 [0 0 0 0 7 0 0 0 0]
 [0 0 0 0 0 3 0 0 0]
 [0 0 0 0 0 1 3 0 0]
 [0 0 0 0 0 0 0 2 1]
 [0 0 0 0 0 0 0 0 1]]
Classification Report:
              precision    recall  f1-score   support

           0       0.67      1.00      0.80         4
           1       0.60      0.60      0.60         5
           2       0.00      0.00      0.00         2
           3       1.00      1.00      1.00         2
           4       1.00      1.00      1.00         7
           5       0.75      1.00      0.86         3
           6       1.00      0.75      0.86         4
           7       1.00      0.67      0.80         3
           8       0.50      1.00      0.67         1

    accuracy                           0.81        31
   macro avg       0.72      0.78      0.73        31
weighted avg       0.79      0.81      0.78        31

Accuracy: 0.8064516129032258
```

Рисунок 2.75 – Оцінка якості класифікаційної моделі прогнозування ризиків конвергенції системи протидії фінансовим та кіберзлочинам

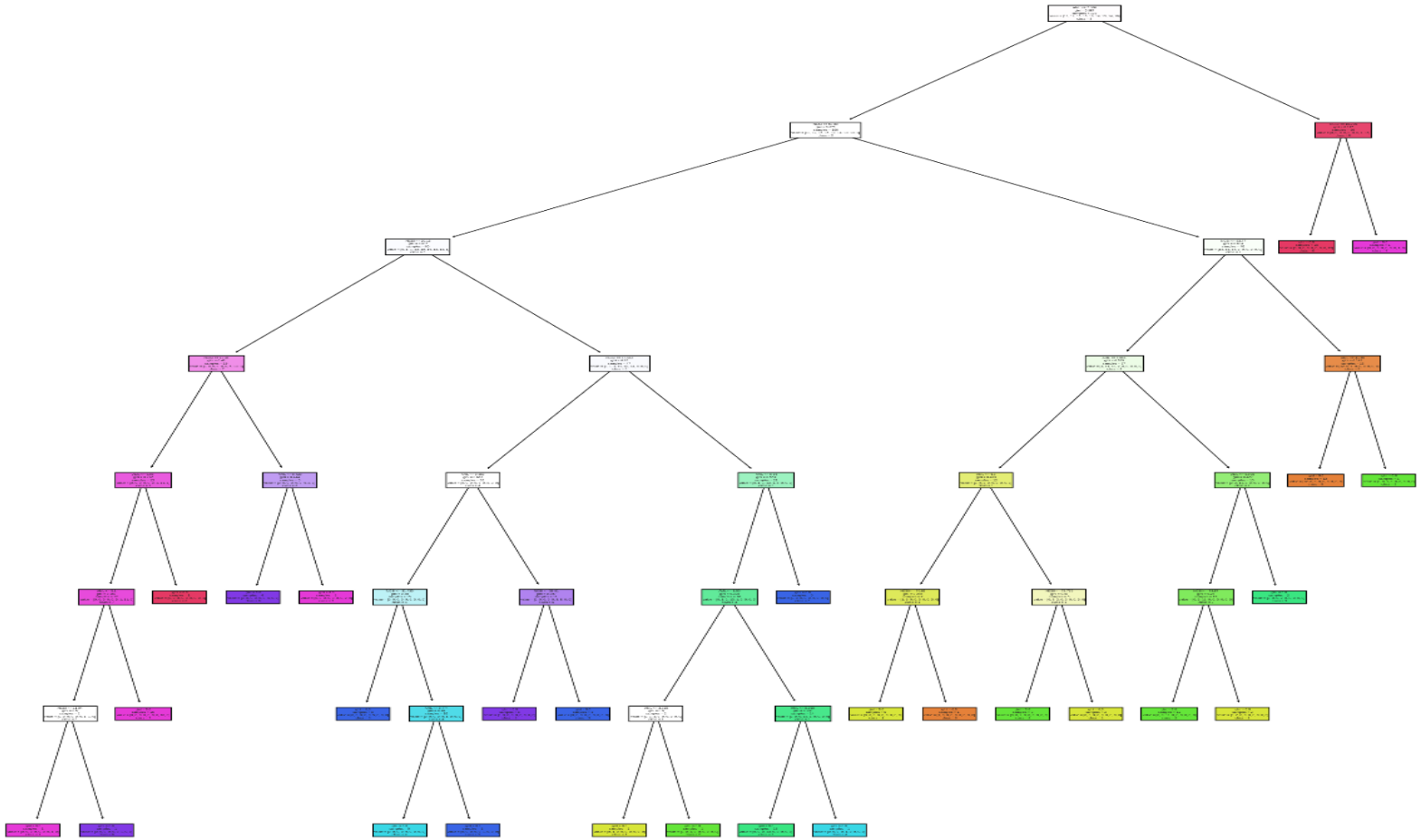


Рисунок 2.74 – Класифікаційна модель прогнозування ризиків конвергенції системи протидії фінансовим та кіберзлочинам

250

В цілому загальна точність моделі є високою і відповідає приблизно 81%. Хоча даний показник не є гарним для моделей такого рівня, але це пов’язано із тим, що вона передбачає класифікацію значної кількості груп. Наприклад, для другого кластеру модель не зможе зробити жодного передбачення, хоча інші рівні ризику вона передбачатиме на рівні вище середнього.

Отже, як можна побачити стабільність систем захисту як проти кіберризиків, так і проти відмивання фінансових коштів, залежить від рівня розвитку країни та комплексу заходів щодо мінімізації та попередження можливих загроз. Неможливо створити єдину модель, яка б задовольняла потреби усіх країн та організацій, оскільки не дивлячись на можливі загальні тенденції, кожна система має власну основу і власні вразливі місця, до яких можуть адаптуватися різні види атак. Тому для підвищення рівня захищеності даних та зниження рівня відмивання коштів необхідно застосовувати комплексні заходи. В першу чергу необхідно будувати міцну законодавчу базу, яка дозволить правоохоронним органам та організаціям вільно ділитися інформацією та швидко протидіяти можливим злочинам. Необхідно застосовувати всі можливі заходи безпеки в інформаційному просторі, такі як аутентифікація користувачів, використання цифрового підпису тощо. Саме комплексні дії дозволять пристосуватися до мінливого середовища та забезпечити надійний захист і підвищення рівня національної безпеки.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**].

2.2.3 Побудова нейромережевої моделі потенційної конвергенції системи кібербезпеки та протидії легалізації кримінальним доходам

Статистичний аналіз потенційної конвергенції системи кібербезпеки та протидії фінансовим злочинам.

Аналіз базових статистик було проведено з допомогою статистичних методів обробки даних, їх систематизації, наочного представлення як таблиць і

графіків, а також кількісний опис даних з допомогою системи статистичних показників. Статистичний аналіз проводився з допомогою мови програмування Python. Python - високорівнева мова програмування загального призначення з динамічною строгою типізацією та автоматичним управлінням пам'яттю, орієнтована на підвищення продуктивності розробника, читання коду та його якості, а також на забезпечення переносимості написаних на ньому програм [**Ошибка! Источник ссылки не найден.**]. У даній роботі мова програмування Python була використана для розрахунку базових статистик і візуалізації даних конвергенції системи кібербезпеки та протидії фінансовим злочинам [**Ошибка! Источник ссылки не найден.**].

Першим кроком був імпорт необхідних бібліотек, таких як: Pandas, Numpy, Preprocessing, Matplotlib.Pyplot і деякі інші [**Ошибка! Источник ссылки не найден.**, **Ошибка! Источник ссылки не найден.**, **Ошибка! Источник ссылки не найден.**].

Для відображення даних було використано функцію .head() (рис. 2.76).

```
df.head()
```

	Country	GCI	ICTDI	NRI	NCSI	DDL	PSI	GEI	EDB	CI	CPI	GTI	FCI
0	Australia	89	82	79	59.74	80.49	1.00	1.60	80.14	42.55	77	2.827	244.358302
1	Austria	83	80	77	68.83	78.67	0.91	1.45	78.54	20.41	76	1.852	310.412705
2	Bahrain	59	76	73	25.97	74.43	-0.84	0.18	68.03	36.96	36	3.883	490.706709
3	Barbados	17	73	0	15.58	73.10	0.92	0.43	56.78	51.31	68	0.000	230.952985
4	Belgium	81	78	77	85.71	77.62	0.41	1.17	71.71	42.17	75	4.060	212.965184

Рисунок 2.76 – Відображення вхідних даних

На рисунку 2.76 представлено дані, які характеризують потенційний процес конвергенції системи кібербезпеки та протидії фінансовим злочинам. Представлено 14 стовпців, перший стовпець вказує на порядковий номер, стовпець під назвою «Country» містить перелік країн, з 3 по 14 стовпці

представлено такі статистичні дані, як: GCI – Глобальний індекс кібербезпеки; ICTDI – Індекс розвитку інформаційно-комунікаційних технологій; NRI – Індекс мережевої готовності; NCSI – Національний індекс кібербезпеки; DDL – Рівень цифрової трансформації; PSI – Індекс політичної стабільності; GEI – Індекс ефективності уряду; EDB – Індекс легкості ведення бізнесу; CI – Індекс злочинності; GTI – Глобальний індекс тероризму; CPI – Індекс споживчих цін; FCI – Індекс фінансової таємниці.

Наступним етапом був розрахунок базових статистик для кожного з вказаних показників. До основних статистик відноситься: загальна кількість спостережень, середнє значення, стандартне відхилення, мінімальне значення, і максимальне значення (рис. 2.77).

```
df.describe()
```

	GCI	ICTDI	NRI	NCSI	DDL	PSI	GEI	EDB	CI	CPI	GTI	FCI
count	76.000000	76.000000	76.000000	76.000000	76.000000	76.000000	76.000000	76.000000	76.000000	76.000000	76.000000	76.000000
mean	66.078947	65.078947	61.894737	54.255000	65.576184	0.322763	0.633684	70.199342	42.055000	55.342105	2.143276	284.696287
std	24.289786	18.071534	19.904826	23.195725	13.973113	0.779067	0.848850	10.234283	14.360367	18.745704	2.317043	279.032311
min	2.000000	0.000000	0.000000	3.900000	28.100000	-1.860000	-1.580000	30.850000	13.100000	18.000000	0.000000	27.860721
25%	56.000000	56.000000	57.000000	35.060000	57.725000	-0.227500	0.040000	64.265000	33.897500	40.500000	0.052250	127.230995
50%	75.000000	69.500000	63.500000	57.140000	66.815000	0.465000	0.495000	71.825000	40.170000	55.000000	1.011500	208.255223
75%	85.000000	79.000000	77.000000	71.755000	78.145000	0.950000	1.272500	78.095000	49.292500	72.250000	3.958000	355.705963
max	93.000000	90.000000	86.000000	96.100000	85.130000	1.540000	2.230000	86.590000	83.600000	88.000000	7.568000	1589.573888

Рисунок 2.77 – Базова статистика

Середнє значення вибірки характеризує розташування значень випадкової величини та вказує на центр розсіювання даних. Стандартне відхилення – це найпоширеніший показник розсіювання значень випадкової величини щодо її математичного очікування. Мінімальне значення - вказує на найменше значення вибірки на вказаному інтервалі даних. Максимальне значення, відповідно, вказує на найбільше значення вибірки.

З малюнку 2.77 видно, що загальна кількість спостережень для кожного показника складає 76. Середнє значення, стандартне відхилення, максимальне і мінімальне значення різняться.

Візуалізація даних - це представлення даних у вигляді, який забезпечує найефективнішу роботу людини, яка їх вивчає. Візуалізація даних знаходить широке застосування у багатьох сферах, таких як: наукових та статистичних дослідженнях, у педагогічному дизайні для навчання та тестування, у новинних зведеннях та аналітичних оглядах [**Ошибка! Источник ссылки не найден.**]. Візуалізація даних допомагає досягати результату, оцінювати значення інформації чи даних. Під візуалізацією даних мається на увазі представлення інформації у графічній формі, наприклад, у вигляді кругової діаграми, графіка або візуального представлення іншого типу [**Ошибка! Источник ссылки не найден.**]. Графіки зручно використовувати, якщо потрібно зобразити характер чи загальну тенденцію розвитку явища чи явищ. Лінії зручні і за зображенні кількох динамічних рядів їхнього порівняння, коли потрібно порівняння темпи зростання [**Ошибка! Источник ссылки не найден.**]. Гістограми є одним із найважливіших інструментів аналізу даних. Подання результатів спостережень з допомогою дозволяє оцінити ряд статистичних показників, зробити висновки про функції розподілу і визначити можливі відхилення, і навіть порівняти набори даних [**Ошибка! Источник ссылки не найден.**].

Для візуалізації даних Pandas було використано бібліотеку Matplotlib. З її допомогою можна з легкістю будувати діаграми [**Ошибка! Источник ссылки не найден.**]. За допомогою вже імпортованого модуля Matplotlib.Pyplot та метода Plot() були побудовані графіки 2.78-2.79.

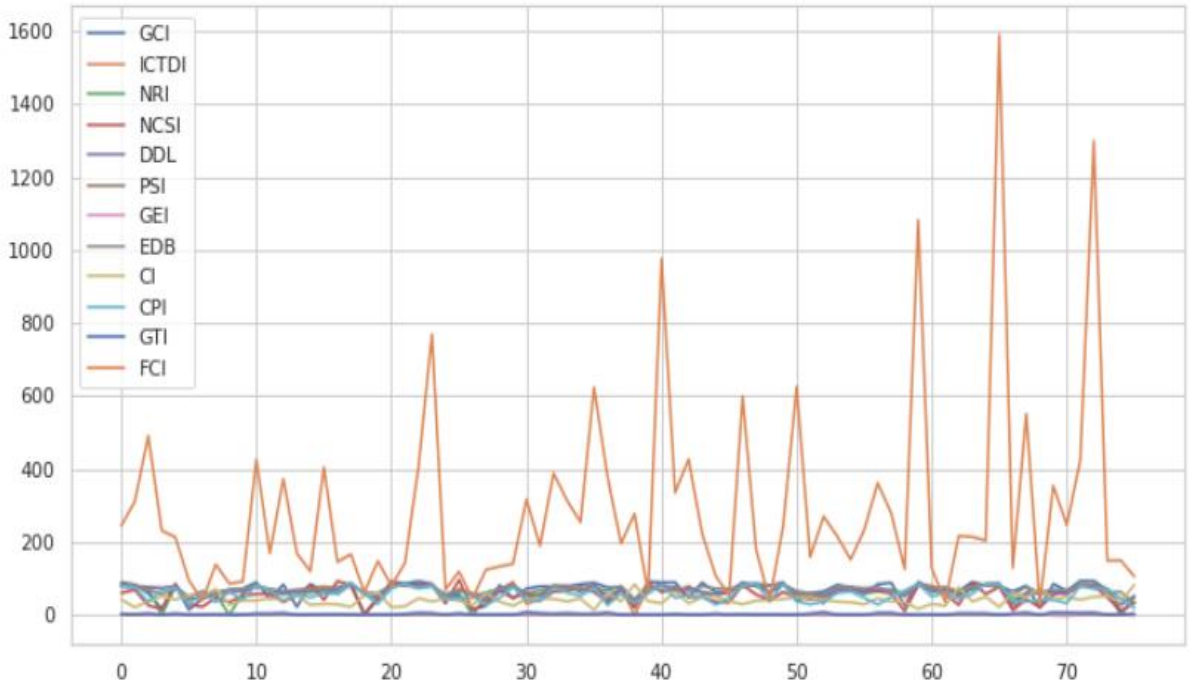


Рисунок 2.78 – Графік розподілу даних

На рисунку 2.78 зображено розподіл всіх даних. Як видно, показники, окрім FCI, знаходяться приблизно в одних межах. Показник FCI відрізняється, значення цієї характеристики значно більше, ніж значення інших. Максимальне і мінімальне значення цього показника дорівнює 1589,573888, 27,860721, відповідно. З метою кращого розуміння розподілу інших показників був зроблений ще один графік, він описує всі вхідні дані, окрім FCI (рис. 2.79).

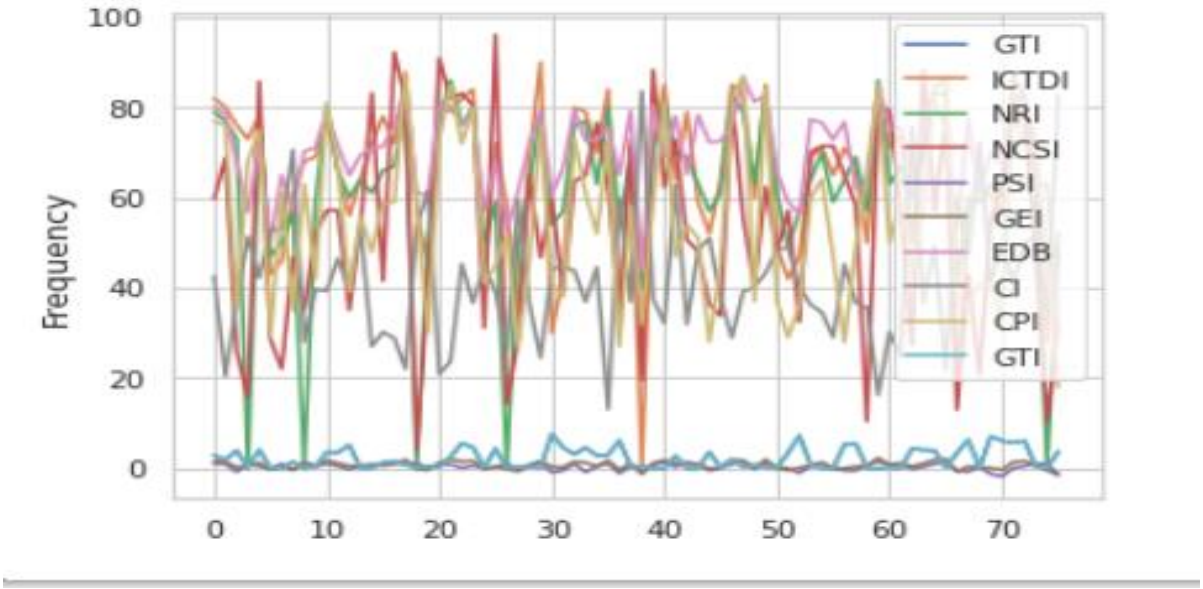


Рисунок 2.79 – Графік розподілу даних

Рисунок 2.79 описує розподіл даних. Загальна кількість спостережень описаних на рисунку – 2.80, значення показників варіюються від 0 до 100.

Наступник кроком є побудова гістограми для кожного показника окремо (рис. 2.80).

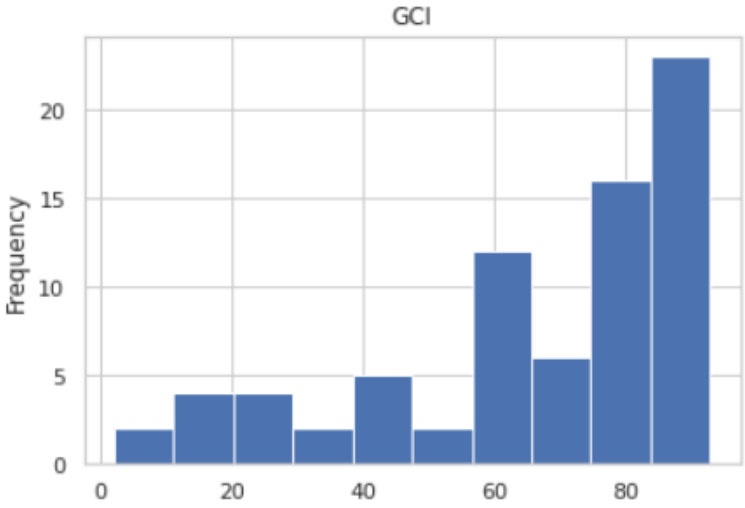


Рисунок 2.80 – Гістограма розподілу показника «GCI»

Як видно з рисунку 2.80, більшість значень знаходяться на проміжку 60-93. Кількість спостережень від 0 до 60 повторюється набагато менше.

На рисунку 2.81 показано, що найбільша кількість спостережень зосереджена на проміжку від 70 до 80. Тобто, найчастіше у вибірці повторюється саме ці значення.

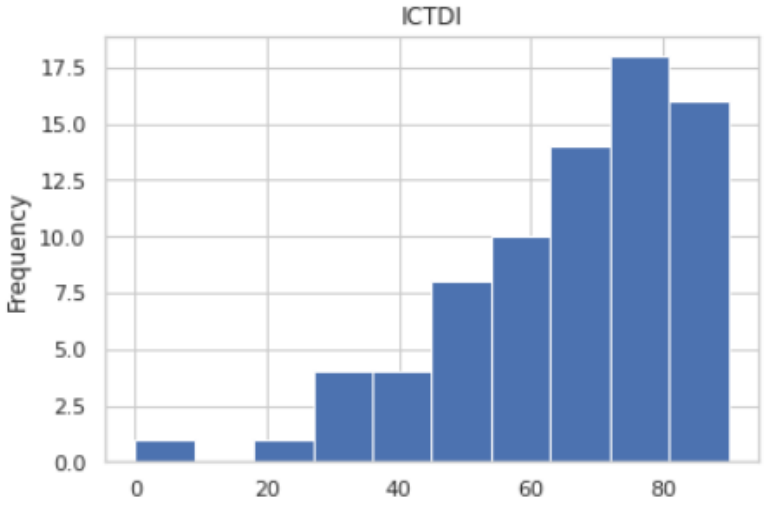


Рисунок 2.81 – Гістограма розподілу показника «ICTDI»

Рисунок 2.82 описує характеристики «NRI». Найбільші значення зосереджені на проміжку від 50 до 90, і повторюються більшу кількість разів, ніж значення від 0 до 50.

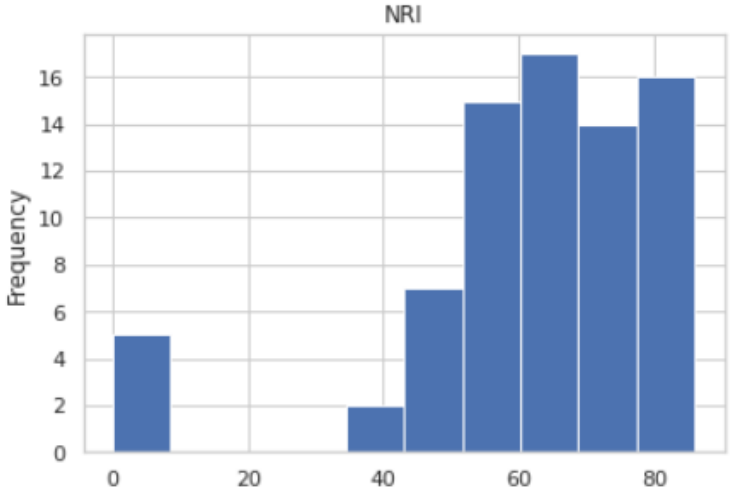


Рисунок 2.82 – Гістограма розподілу показника «NRI»

На рисунку 2.83, порівнюючи з попередніми, дані розподіляються більш рівномірно, найчастіше повторюються значення починаючи від 40 і до 90. Інші значення зустрічаються не так часто.

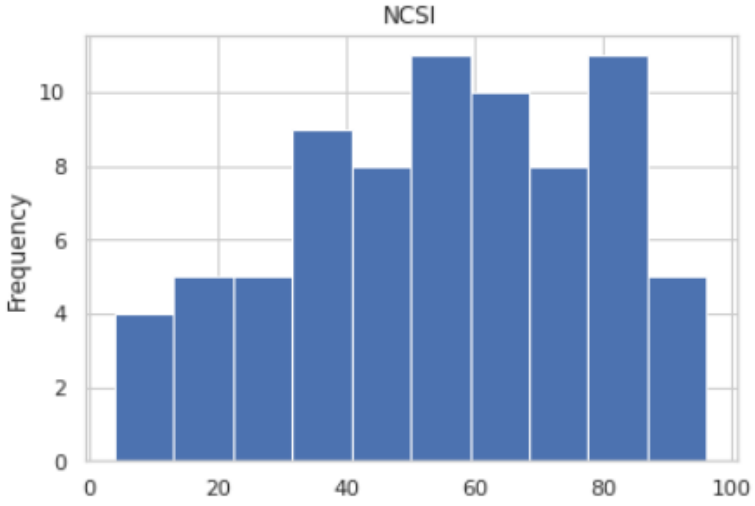


Рисунок 2.83 – Гістограма розподілу показника «NCSI»

На рисунках додатку Б зображено особливості розподілу факторів «PSI», «EDB», «CI» , «CPI», «GTI», відповідно, які також демонструють нерівномірність розподілу змінних.

Канонічний аналіз - це багатовимірний метод аналізу, що стосується визначення взаємозв'язків між групами змінних у наборі даних. Його основна мета – пошук максимальних кореляційних зв'язків між групами вихідних змінних [**Ошибка! Источник ссылки не найден.**]. Канонічний аналіз даних був проведений за допомогою програми STATISTICA (рис. 2.84).

N=76	Canonical Analysis Summary (Convergensy.sta) Canonical R: .96635 Chi²(35)=297.75 p=0.0000 Left Set	Right Set
No. of variables	5	7
Variance extracted	100.000%	85.9457%
Total redundancy	58.7705%	58.7748%
Variables: 1	GCI	PSI
2	ICTDI	GEI
3	NRI	EDB
4	NCSI	CI
5	DDL	CPI
6		GTI
7		FCI

Рисунок 2.84 – Результати канонічного аналізу по всім змінним

Отримане значення канонічного R достатньо велике і дорівнює 0.96635. Це свідчить про те, що наявний сильний кореляційний зв'язок між факторами, які характеризують системи кібербезпеки та запобігання фінансовим злочинам. Критерій Пірсона, який у даному випадку дорівнює 297,75, підтверджує статистичну значимість коефіцієнта кореляції, і рівень значущості даного коефіцієнта не перевищує 0,05 (p=0,0000). Значення лівої множини, яка була сформована з індексів системи кібербезпеки, дорівнює 58,7705%. Дане значення говорить про те, що фактори, описані у правій множині, які характеризують рівень протидії фінансовим злочинам, пояснюють на 58,7705% мінливість факторів системи кібербезпеки. Система протидії фінансовим злочинам певною мірою залежить від системи кібербезпеки в країні. Фактори системи кібербезпеки на 58,7748% пояснюють мінливість факторів, які характеризують

рівень протидії фінансовим шахрайствам. З проведеного аналізу показників видно, що фактори системи кібербезпеки мають вплив на процес протидії фінансовим злочинам.

Наступним етапом був аналіз впливу кожного фактору системи кібербезпеки на протидію фінансовим шахрайствам. На рисунку 2.85 показані результати канонічного аналізу фактору системи кібербезпеки GCI та факторів протидії фінансовим шахрайствам.

N=76	Canonical Analysis Summary (Convergensy.sta) Canonical R: .62413 Chi²(7)=34.794 p=.00001 **Left Set**	Right Set
No. of variables	1	7
Variance extracted	100.000%	18.2672%
Total redundancy	38.9537%	7.11577%
Variables: 1	GCI	PSI
2		GEI
3		EDB
4		CI
5		CPI
6		GTI
7		FCI

Рисунок 2.85 – Результати канонічного аналізу.

Як видно, отримане значення канонічного R не є достатньо великим(R = 0,62413). Це говорить про низьким кореляційний зв'язок між глобальним індексом кібербезпеки та запобігання фінансовим злочинам. Критерій Пірсона, який дорівнює 34,794, також підтверджує статистичну незначущість коефіцієнта кореляції. Значення лівої множини дорівнює 38,9537% і говорить про те, що фактори, які описують протидію фінансовим злочинам, на 38,9537% пояснюють мінливість глобального індексу кібербезпеки.

Як видно з рисунка 2.85, фактори протидії фінансовим злочинам у дуже незначному відсотку залежать від обраного фактору системи кібербезпеки. Фактор глобального індексу кібербезпеки лише на 7,11577% пояснює мінливість факторів протидії фінансовим злочинам. Отримане значення говорить по те, що вплив глобального індексу кібербезпеки на протидію фінансовим шахрайствам низький і не має значущого впливу на систему протидії фінансовим шахрайствам.

На рисунку В.1 додатку В наведені результати аналізу впливу індексу розвитку інформаційно-комунікаційних технологій на протидію фінансовим шахрайствам. Отримане значення R є низьким, це говорить про низький кореляційний зв'язок між обраними факторами. Коефіцієнт Пірсона дорівнює 60,519, що підтверджує статистичну незначущість коефіцієнта кореляції. Значення факторів лівої множини дорівнює 57,6170%, тобто фактори протидії фінансовим шахрайствам на 57,6170% описують мінливість фактору системи кібербезпеки. Значення факторів правої множини дорівнює 12,6134%, тобто індекс розвитку інформаційно-комунікаційних технологій на 12,6134% описує мінливість факторів протидії фінансовим злочинам. Отримані значення вказують на те, що вплив індексу розвитку інформаційно-комунікаційних технологій на протидію фінансовим злочинам є, але він не настільки великий.

Рисунок В.2 додатку В описує вплив індексу мережевої готовності на фактори протидії фінансовим злочинам. Значення отриманого R є низьким, і це говорить про низький кореляційний зв'язок між факторами. Коефіцієнт Пірсона підтверджує припущення статистичної незначущість коефіцієнта кореляції. Значення факторів лівої множини складає 50,5428%. Це говорить про те, що фактори протидії фінансовим шахрайствам на 50,5428% описують мінливість індексу мережевої готовності. Значення факторів правої множини дорівнює 7,99112%. Це говорить про те, що індекс мережевої готовності на 7,99112% описує мінливість факторів протидії фінансовим злочинам. Отримані результати вказують на невисокий вплив фактору мережевої готовності на фактори протидії фінансовим злочинам.

Кореляційний зв'язок між факторами протидії фінансовим злочинам та національним індексом кібербезпеки є слабким, про це каже коефіцієнт кореляції, який дорівнює 0,74559 (рис. В.3). Критерій Пірсона дорівнює 57,225 і підтверджує, що коефіцієнт кореляції не є статистичну значущим. Значення надмірності для лівої множини, яка складається з фактору системи кібербезпеки, а саме фактору «Національний індекс кібербезпеки» дорівнює 55,5897%. Це означає, що фактори правої множини, які складаються з індексів протидії фінансовим злочинам, на 55,5897% пояснюють мінливість системи кібербезпеки. Протидія фінансовим

шахрайствам частково залежить від національного індексу кібербезпеки, оскільки фактор системи кібербезпеки на 23,4440% описує мінливість системи протидії фінансовим злочинам. Хоча отримані значення є помірними, але цього є достатньо для доказу невеликого впливу показника «Національний індекс кібербезпеки» на протидію фінансовим шахрайствам в країнах.

На рисунку В.4 зображено результати аналізу впливу фактору «Рівень цифрової трансформації» на протидію фінансовим злочинам в країнах. Як видно, отримане значення канонічного R=0,95472. Це говорить про те, що існує сильний кореляційний зв'язок між факторами, які характеризують рівень цифрової трансформації та протидію фінансовим злочинам. Критерій Пірсона, який дорівнює 170,94, і рівень значимості якого не перевищує 0,05 (p=0,0000), підтверджує статистичну значущість коефіцієнта кореляції. Значення надмірності для лівої множини, яка складається з фактору системи кібербезпеки, а саме «Рівень цифрової трансформації», дорівнює 91,1491%. Цей свідчить про те, що фактори правої множини, які описують протидію фінансовим злочинам, на 91,1491% пояснюють мінливість індексу рівня цифрової трансформації, що свідчить про високе значення впливу. Процес протидії фінансовим шахрайствам в країнах залежить від кіберзахисту фінансових систем, так як індекс рівня цифрової трансформації на 39,7615% описує мінливість факторів, які характеризують протидію фінансовим шахрайствам у країнах. Отримане значення є високим і це говорить про те, що система кібербезпеки (індекс рівня цифрової трансформації) має сильний вплив на протидію фінансовим махінаціям. Отже, в процесі аналізу впливу факторів кібербезпеки було виявлено один індекс, який має сильний вплив на фактори, які характеризують протидії фінансовим злочинам, цим фактор є «Рівень цифрової трансформації».

Усунення мультиколінеарності факторів із використанням методу головних компонентів.

Метод головних компонентів — це техніка для зменшення розмірності наборів даних, підвищення інтерпретації, але водночас мінімізації втрати інформації. Це

робиться шляхом створення нових некорельованих змінних, які послідовно максимізують дисперсію [**Ошибка! Источник ссылки не найден.**].

Метод головних компонентів виконаний на мові програмування Python. Його необхідність викликана високим рівнем кореляції між змінними, які було відібрано в результаті канонічного аналізу (рис. 2.86).

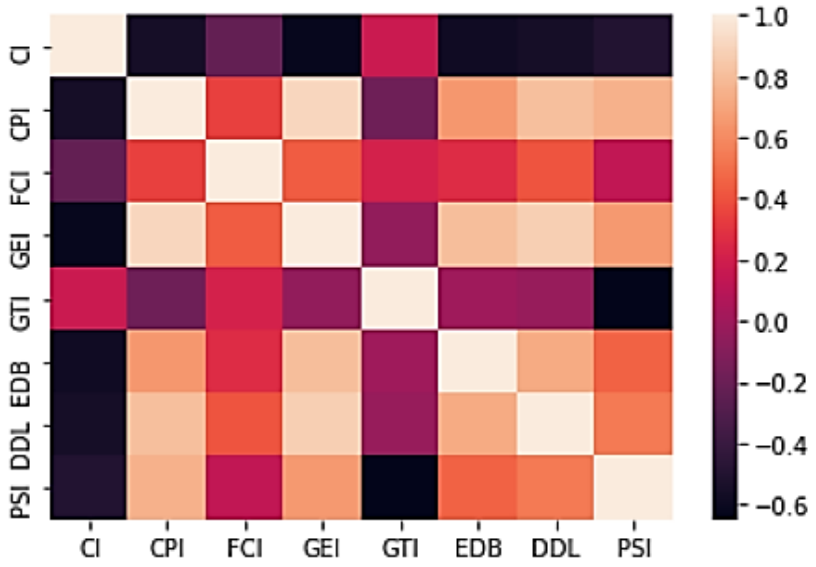


Рисунок 2.86 – Кореляційна матриця

Показники, які були вибрані, мають високий рівень мультиколлінеарності. Оскільки показники мультиколінеарні, то подальша робота з цими даними неможлива, тому й необхідно використати метод головних компонентів. Після побудови його графіка та виконання розрахунків видно, що отримано всього 4 статистично значущі компоненти. Як видно з рисунків 2.87-2.88, 4 компоненти накопичують варіацію 93% і рівень значущості кожного з них не повинен бути меншим за 0,05.

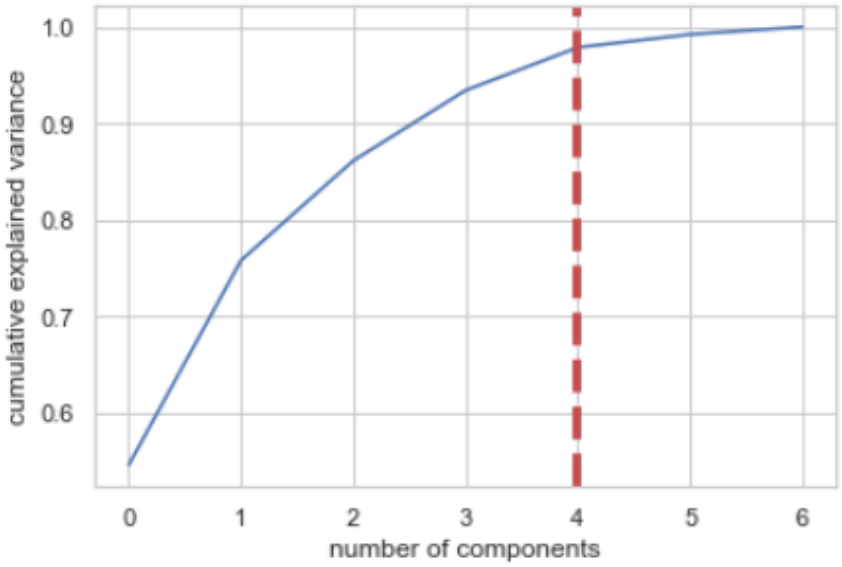


Рисунок 2.87 – Графік нових компонентів

	Cumulative Variance Ratio	Explained Variance Ratio
0	0.545733	0.545733
1	0.757906	0.212173
2	0.861269	0.103363
3	0.934204	0.072935

Рисунок 2.88 – Результати розрахунку

Результат отриманих значень головних компонент представлений на рисунку 2.89, які будуть використані для побудови нейронної мережі.

Побудова нейромережевої моделі.

Нейронна мережа — це серія алгоритмів, які створені для розпізнавання основних зв’язків в наборі даних. Нейронні мережі можуть адаптуватися до зміни вхідних даних; таким чином мережа генерує найкращий можливий результат без необхідності перепроектувати критерії виводу. Результати розрахованих коефіцієнтів нейронної мережі представлені на рисунку 2.89.

```
array([[ 8.75068112e-01,  1.14594642e+00,  9.77763578e-01,
         3.46989096e-02,  1.16302930e+00,  2.97045324e-01,
        -1.45524418e-01],
       [ 7.58777731e-01,  9.68062577e-01,  8.20387503e-01,
        -1.51728868e+00,  1.10932929e+00, -1.26545664e-01,
         9.27752513e-02],
       [-1.50242411e+00, -5.38020594e-01, -2.13376592e-01,
        -3.57153423e-01, -1.03867137e+00,  7.55826948e-01,
         7.43208829e-01],
       [ 7.71698884e-01, -2.41547529e-01, -1.31992712e+00,
         6.48764461e-01,  6.79729155e-01, -9.31151315e-01,
        -1.93885806e-01],
       [ 1.12720062e-01,  6.36012744e-01,  1.48588382e-01,
         8.06136284e-03,  1.05562927e+00,  8.32725004e-01,
        -2.58779088e-01],
       [-7.78839522e-01, -1.13096672e+00, -1.97008703e+00,
         7.59520577e-01, -1.41457149e+00, -9.31151315e-01,
        -6.84996207e-01],
```

Рисунок 2.89 – Фрагмент розрахованих значень компонент

Для побудови нейромережевої сітки була використана активаційна функція ReLU. ReLU - це нелінійна функція активації. Ця функція є найчастіше використовуваною функцією. Її використовують для згорткових нейронних мереж та глибокого навчання для всіх шарів, крім вихідного.

На рисунку 2.90 зображено результати нейромережевої сітки, яка показує фактичні значення та прогнозовані значення, а також оцінку. Виходить, що критерій детермінації дорівнює за результатами тесту 0,799 та оцінка тренувального коефіцієнта детермінації дорівнює 0,821 (рис. 2.91). Модель містить три шари, в кожному шарі міститься по 45 вузлів. Її характеристики наведені у додатку Г.

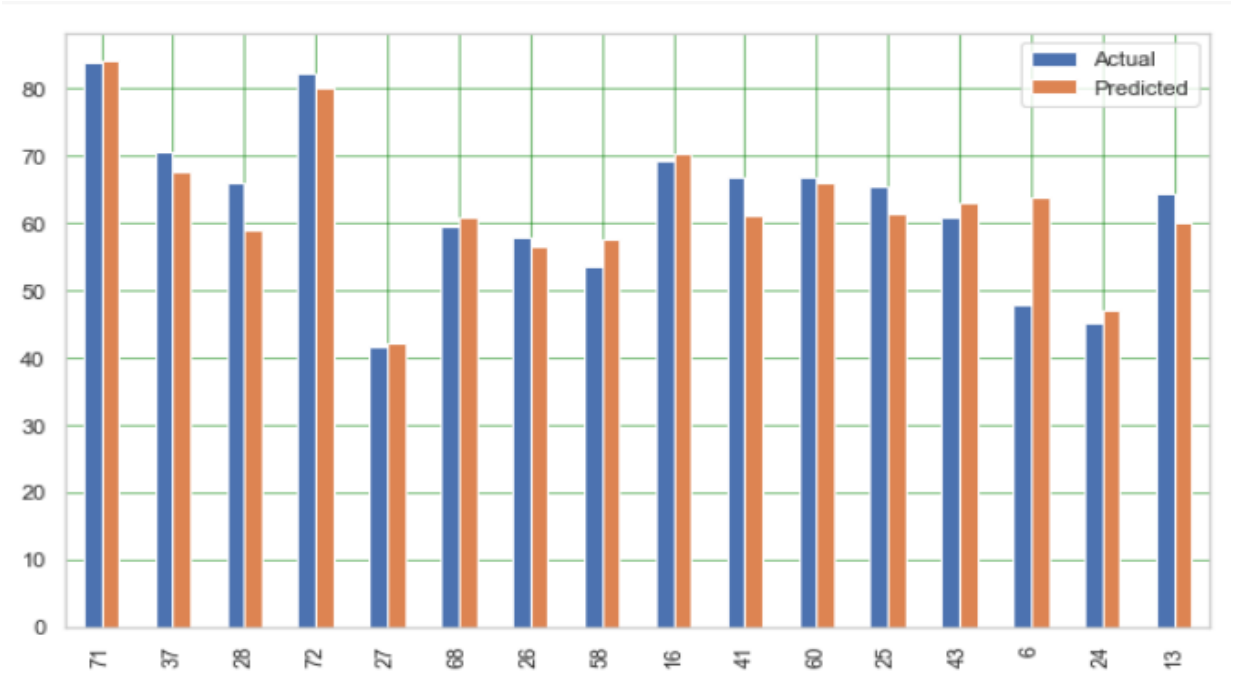


Рисунок 2.90 – Нейромережева сітка

```
Mean Absolute Error: 3.47495173501873
Mean Squared Error: 26.07682730734608
Root Mean Squared Error: 5.106547493889201
Test R^2 Score : 0.799
Training R^2 Score : 0.821
```

Рисунок 2.91 – Результати розрахунку

Такі оцінки як, середня абсолютна помилка (яка дорівнює 3,47495173501873) і середня квадратична помилка (яка дорівнює 26,07682730734608) описують порівняння якість прогнозованих даних і фактичних. Як видно з рисунка 2.27, помилки є незначними і це говорить про високу якість прогнозу.

Наступний етап – побудова кривої втрат. Одним з найбільш часто використовуваних графіків для налагодження нейронної мережі є крива втрат під час навчання (рис. 2.92).

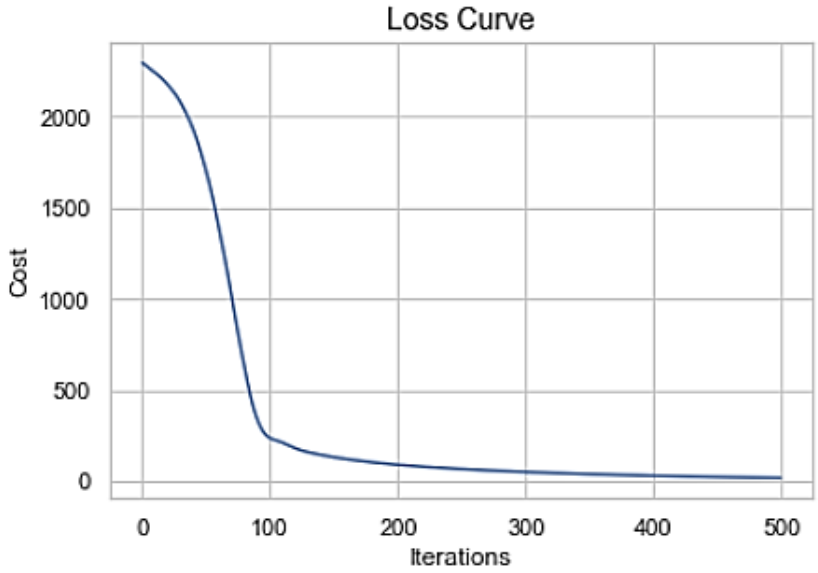


Рисунок 2.92 – Графік втрат

Дана крива вимірює помилку моделі і показує «наскільки погано працює модель». Усі ці показники фіксують продуктивність моделі, тому чим вони вищі, тим кращою стає модель. В даному випадку кількість втрат знижується, крива з часом зменшується і вирівнюється до певного рівня.

Коефіцієнти, які були отримані для нейромережевої сітки зображені на рисунку 2.93-2.94.

```
[(4, 45), (45, 45), (45, 45), (45, 1)]
[array([[-0.26004826, -0.12029681,  0.20590884, -0.06945696, -0.3113571 ,
         0.21721447, -0.30771875, -0.10858792, -0.2631856 , -0.21179765,
         0.03922925,  0.34470119,  0.00092399, -0.03584113, -0.21201222,
         0.39837075,  0.21987484, -0.09763449, -0.34011139, -0.13194794,
        -0.32285526,  0.35387933,  0.0824541 , -0.32409041, -0.31574998,
         0.01327073, -0.374209  , -0.2635563 ,  0.23914328, -0.16130783,
        -0.08460061, -0.0878057 , -0.27154981, -0.22982154,  0.18132776,
         0.0032163 , -0.04726616, -0.38168097, -0.05618342, -0.0240545 ,
        -0.01425243,  0.09723332,  0.25665729, -0.27056029,  0.04194362],
       [-0.37137422,  0.29138754,  0.4493272 , -0.09557417,  0.32288134,
        -0.40760883,  0.12893772, -0.281831  , -0.24050574,  0.45293821,
        -0.12016031,  0.40809908,  0.20684486, -0.14231219, -0.36514533,
        -0.4131254 ,  0.36199315, -0.07692699, -0.33685647,  0.35394341,
        -0.43567026,  0.46121382,  0.26172665,  0.32925829,  0.08731418,
        -0.21769305,  0.37403948, -0.31338637,  0.16617805,  0.32513892,
         0.27063393, -0.14837558, -0.05335017, -0.26485246,  0.47923464,
        -0.23381631,  0.43610618, -0.272976  , -0.34756817, -0.39744876,
         0.23763593, -0.327351  , -0.09516424, -0.16217758,  0.4214515 ],
       [-0.34914899,  0.03479934, -0.38569757, -0.05618054, -0.32536316,
         0.3083773 , -0.50888427,  0.03729243,  0.12824579, -0.29934257,
         0.18971017,  0.03014801, -0.32204607, -0.06726018,  0.35458359,
         0.03199527, -0.16551574,  0.27497358, -0.08863406, -0.07647509,
        -0.34039611,  0.16389043,  0.15351833,  0.11692195, -0.07506765,
        -0.0310238 ,  0.02180154,  0.3428187 , -0.37137128, -0.03728706,
        -0.46024447,  0.12702506, -0.55350416,  0.06861506,  0.0110601 ,
         0.08716832, -0.06640432, -0.26295906, -0.23210827, -0.11435545,
        -0.39570294, -0.34034274,  0.08491891, -0.00243018,  0.32409415],
       [ 0.28956792,  0.6455377 ,  0.28963336, -0.00493871, -0.44261102,
        -0.53542088, -0.18523631,  0.47829689,  0.0411324 ,  0.1874174 ,
         0.42292073, -0.36328156, -0.1570596 , -0.23801606, -0.36803248,
        -0.15363543, -0.09741371,  0.60318768,  0.2947905 , -0.05998706,
        -0.11174835,  0.19493713,  0.42810767,  0.1572119 , -0.05909069,
        -0.21803694,  0.13179591, -0.24389973, -0.12475218,  0.03903671,
        -0.3136397 , -0.33920583,  0.0564947 , -0.28340583, -0.40138855,
         0.41286151, -0.02299765, -0.15389607, -0.32725482, -0.36204329,
        -0.24209945, -0.63771505,  0.04305627,  0.13866181,  0.3861812 ]]),
```

```
array([[-2.09553329e-03, -2.83383686e-01, -1.89348098e-01, ...,
         4.10836329e-09,  1.78167763e-01, -2.12532525e-02],
       [-1.56179435e-07,  1.68390463e-01, -4.12112578e-02, ...,
        -7.84874525e-03,  3.62979784e-01,  1.52828725e-01],
       [-5.30306052e-13, -3.59982154e-01, -4.34660216e-01, ...,
        -8.13580897e-03,  2.33873914e-01,  1.51226849e-01],
       ...,
       [-1.00980314e-02,  2.62556370e-01,  1.98985453e-02, ...,
        -4.36060544e-09,  3.31775193e-01, -1.40769837e-01],
       [-7.81491532e-03, -1.47737230e-01, -4.39706741e-01, ...,
        -1.85035924e-03, -1.51867423e-01,  2.72673594e-01],
       [-1.27720020e-03,  1.45879762e-01,  1.30810459e-01, ...,
        -5.32661953e-05,  3.22773706e-01,  2.96754380e-01]]),
array([[ 6.67087614e-03,  2.57824102e-04, -2.10369387e-04, ...,
         1.30100327e-05,  3.10330738e-10, -3.63647439e-03],
       [-2.22895452e-02,  3.96579367e-02, -2.10459127e-01, ...,
         2.29254822e-12,  1.34435617e-01,  1.47489465e-01],
       [-7.98080545e-02,  6.39805649e-02, -2.07827714e-01, ...,
         6.69372477e-02, -2.34963646e-01,  4.64588216e-01],
       ...,
       [-5.49662702e-04,  9.97029747e-04,  2.77062508e-04, ...,
         3.03950320e-13, -3.10835098e-10, -6.57735393e-05],
       [-1.54747232e-01, -6.46594534e-02,  7.36793926e-02, ...,
        -1.13950265e-01, -9.87774397e-02,  3.27221960e-01],
       [ 1.27878535e-02, -8.27450917e-02, -7.72087166e-02, ...,
        -2.36712608e-01,  2.10399132e-01,  5.84909485e-02]]),
```

Рисунок 2.93 – Коефіцієнти нейромережевої сітки

```
array([[-2.25687092e-01],
       [-4.25735261e-02],
       [-3.20274899e-02],
       [-2.52926024e-01],
       [ 4.52188309e-01],
       [ 3.49553936e-01],
       [-8.57636211e-03],
       [-3.17339902e-01],
       [-1.04082447e-01],
       [-2.69530392e-02],
       [-2.28642096e-01],
       [-2.27235498e-01],
       [-2.59030798e-01],
       [-2.94106680e-01],
       [ 4.95360488e-01],
       [ 2.99785542e-01],
       [-3.01470028e-01],
       [ 4.66513231e-01],
       [-2.97192536e-01],
       [-1.10176387e-01],
       [ 2.18239491e-01],
       [ 3.80657163e-01],
       [-1.49110418e-01],
       [-3.89642835e-02],
       [ 4.35854289e-01],
       [ 1.95112094e-01],
       [ 4.82352862e-01],
       [-9.72168594e-02],
       [ 3.83305037e-01],
       [-1.74542005e-02],
       [-2.59554935e-01],
       [-5.32480436e-02],
       [-3.13696354e-01],
       [-3.34995432e-01],
       [-3.39534231e-01],
       [ 4.27805082e-01],
       [ 4.69683121e-01],
       [ 6.01660508e-05],
       [ 4.53446623e-02],
       [-1.43142808e-01],
       [-1.58161829e-01],
       [ 1.62127858e-01],
       [-2.67300605e-01],
       [-1.48890230e-01],
       [ 4.75176427e-01]])]
```

Рисунок 2.94 – Коефіцієнти нейромережевої сітки

На малюнку 2.95 зображено графік нейронної мережі, параметри моделі були підібрані вручну, як видно, мережа вийшла досить гарною, і передбачувані значення дуже близькі до реальних значень. Існує так само і метод пошуку параметрів моделі по решітці. Цей метод передбачає, що підбір гіперпараметрів задаються вручну, потім виконується повний перебір. Популярною реалізацією цього методу є Grid Search із sklearn. У цій роботі був виконаний метод підбору параметрів, три варіації нейронної мережі були випробувані і був обраний найкращий варіант мережі (рис. 2.95). Результати розрахунку критерію детермінації, середньої абсолютної помилки, середньої квадратичної помилки представлені на рисунку 2.96.

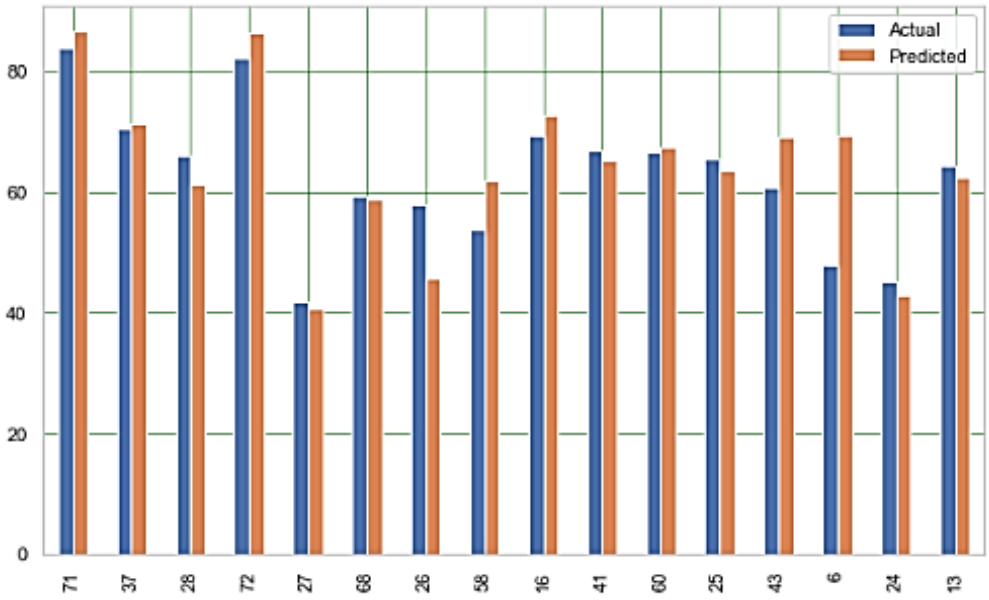


Рисунок 2.95 – Нейромережева сітка

```
Mean Absolute Error: 4.766978809140717
Mean Squared Error: 51.81089511376816
Root Mean Squared Error: 7.197978543575145
Test R^2 Score : 0.601
Training R^2 Score : 0.889
```

Рисунок 2.96 – Результати розрахунку

Як видно з рисунка 2.96, передбачувані значення мають доволі високі відхилення від реальних значень. Отриманий коефіцієнт детермінації невисокий, середня абсолютна помилка (4,766978809140717) і середня квадратична помилка (51,81089511376816). Отже, перша нейромережева модель є кращою для подальшого використовування і прогнозування.

Регресійний аналіз.

Регресійний аналіз — це набір статистичних процесів для оцінки зв'язків між залежною змінною та однією або кількома незалежними змінними [25]. Регресійний аналіз в основному використовується для двох концептуально різних цілей. По-перше, регресійний аналіз широко використовується для передбачення та прогнозування, де його використання суттєво перекривається сферою машинного навчання [24]. По-друге, у деяких ситуаціях регресійний аналіз можна використовувати для висновку про причинно-наслідкові зв’язки між незалежними та залежними змінними [23].

Для аналізу та прогнозування впливу факторів системи кібербезпеки на протидію фінансовим шахрайствам було проведено регресійний аналіз (рис. 2.97). Проведений регресійний аналіз показав, що більшість параметрів має значення P, яке перевищую 0,05. Відповідно, ці параметри не мають впливу на протидію фінансовим шахрайствам.

Наступний етап – проведення відбору параметрів, значення P яких нижче 0,05. Відібраними параметрами для регресії є : «EDB», «CPI», «FCI». FCI-цей параметр був залишений тому що індекс фінансової таємності все ж таки важливий для системи протидії фінансовим злочинам. Відхилення значення 0,06 від 0,05 незначне, ґрунтуючись на цьому, було прийнято рішення залишити цей

показник і провести регресійний аналіз, ґрунтуючись на трьох факторах (рис. 2.98).

OLS Regression Results

Dep. Variable:	DDL	R-squared (uncentered):	0.989
Model:	OLS	Adj. R-squared (uncentered):	0.988
Method:	Least Squares	F-statistic:	917.1
Date:	Thu, 20 Jan 2022	Prob (F-statistic):	1.94e-65
Time:	17:50:40	Log-Likelihood:	-254.77
No. Observations:	76	AIC:	523.5
Df Residuals:	69	BIC:	539.9
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
PSI	-4.2121	2.741	-1.537	0.129	-9.680	1.255
GEI	3.3332	2.472	1.349	0.182	-1.598	8.264
EDB	0.5366	0.069	7.748	0.000	0.398	0.675
CI	0.0723	0.071	1.016	0.313	-0.070	0.214
CPI	0.4296	0.103	4.177	0.000	0.224	0.635
GTI	-0.5404	0.650	-0.831	0.409	-1.838	0.757
FCI	0.0047	0.003	1.376	0.173	-0.002	0.012

Omnibus:	8.057	Durbin-Watson:	2.162
Prob(Omnibus):	0.018	Jarque-Bera (JB):	13.461
Skew:	-0.286	Prob(JB):	0.00119
Kurtosis:	4.981	Cond. No.	1.48e+03

Рисунок 2.97 – Результати регресійного аналізу

OLS Regression Results

Dep. Variable:	DDL	R-squared (uncentered):	0.989
Model:	OLS	Adj. R-squared (uncentered):	0.988
Method:	Least Squares	F-statistic:	2126.
Date:	Thu, 20 Jan 2022	Prob (F-statistic):	6.24e-71
Time:	17:53:51	Log-Likelihood:	-257.13
No. Observations:	76	AIC:	520.3
Df Residuals:	73	BIC:	527.3
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
EDB	0.6010	0.047	12.902	0.000	0.508	0.694
CPI	0.3881	0.059	6.634	0.000	0.272	0.505
FCI	0.0061	0.003	1.908	0.060	-0.000	0.013

Omnibus:	16.234	Durbin-Watson:	2.211
Prob(Omnibus):	0.000	Jarque-Bera (JB):	49.537
Skew:	-0.490	Prob(JB):	1.75e-11
Kurtosis:	6.832	Cond. No.	35.5

Рисунок 2.98 – Результати регресійного аналізу

У результаті проведення регресійного аналізу маємо прогнозовані значення для тестового та тренувального набору даних (рис. 2.99-2.102).

```
38    40.289868
13    64.451372
61    68.930772
14    62.065520
45    61.376066
9     59.607550
11    69.969857
25    59.177610
49    83.789158
62    56.873089
34    65.434820
32    78.437056
73    65.492145
46    81.183942
29    78.009303
24    50.685918
dtype: float64
```

Рисунок 2.99 – Прогнозовані значення для тестового набору даних

```
Mean Absolute Error: 2.7086619105116534
Mean Squared Error: 11.779761215591133
Root Mean Squared Error: 3.4321656742632825
```

Рисунок 2.100 – Результати розрахунку

Середня абсолютна помилка для тестового набору даних складає 2,7086619105116534, що є допустимим. Середня квадратична помилка прогнозу складає 11,779761215591133, середньоквадратична помилка складає 3,432165674263282.

```
                  1     78.597861
71    83.116014  25     59.177610
7     48.714471  28     61.899997
5     41.986723  3      61.929266
24    50.685918  57     57.911556
38    40.289868  47     86.897676
9     59.607550  46     81.183942
14    62.065520  58     63.895372
37    71.341739  39     71.660891
53    71.067008  0      79.543090
64    83.077241  43     68.212335
16    69.491275  74     56.676619
62    56.873089  59     90.679383
67    63.852640  6      62.943810
55    63.432625  63     70.461803
68    52.700596  34     65.434820
30    54.270308  18     59.179764
```

Рисунок 2.101 – Прогнозовані значення для тренувального набору даних

```
Mean Absolute Error: 4.424836546812158
Mean Squared Error: 43.873275551329534
Root Mean Squared Error: 6.623690478225076
```

Рисунок 2.102 – Результати розрахунку

Середня абсолютна помилка для тренувального набору даних складає 4,424836546812158. Середня квадратична помилка прогнозу складає 43,873275551329534, середньоквадратична помилка складає 6,623690478225076.

Проведений регресійний аналіз допоміг виявити фактори, які в поєднанні з індексом рівня цифрової трансформації є важливими в процесі протидії фінансовим шахрайствам у країнах. Цими факторами є індекс легкості ведення бізнесу, індекс споживчих цін, індекс фінансової таємниці.

На сьогодні проблема впливу розвитку технологій та цифровізації економіки на зростання кількість кібершахрайств у сфері фінансів у всьому світі постає перед людством. Темпи розвитку технологій і якості кіберсистем підвищується, і, звичайно, зростає кількість шахрайств у різних сферах, а особливо у сфері фінансів. На даний момент системи фінансів не мають достатнього кіберзахисту, і є вразливими у час інформаційних технологій. Інформація, гроші у різних валютах, цінні папери можуть бути викраденими хакерами з різних куточків світу. З метою покращення якості функціонування фінансової сфери, а також для зменшення кількості кіберзлочинів і протидії фінансовим махінаціям був проведений аналіз потенційної конвергенції системи кібербезпеки та протидії фінансовим шахрайствам. У ході виконання роботи статистичний та візуальний аналіз описали фактори, які можуть впливати на рівень захищеності фінансової сфери. За допомогою канонічного аналізу був виявлений фактор «DDL», який характеризує рівень цифрової трансформації у сфері кібербезпеки. Процес протидії фінансовим шахрайствам в країнах залежить від кіберзахисту фінансових систем, а саме від індексу рівня цифрової трансформації. Цей висновок був зроблений на основі розрахунків, зроблених за допомогою канонічного аналізу. У роботі було використано два методу аналізу – нейромережева сітка і регресійний аналіз. Це було зроблено з метою порівняння двох видів аналізу та виявити, який з них краще описує потенційний процес конвергенції системи кібербезпеки та протидію фінансовим шахрайствам. Основуючись на цих видах аналізу, було виявлено, що важливими

показниками в сфері кібербезпеки та протидії фінансовим шахрайствам є фактори рівня цифрової трансформації і три показника, які характеризують індекс легкості ведення бізнесу, індекс споживчих цін, індекс фінансової таємниці, відповідно.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**].

2.3 Алгоритми розпізнавання поведінки кібершахраїв

2.3.1 Формування кіберпрофілю жертви: гендерний аналіз

За останні двадцять років у світі спостерігається бурхливий розвиток науково-технічного прогресу завдяки Четвертій промисловій революції. Його результати призвели до повної інтеграції та впровадження різноманітних технологій у всі сфери життя суспільства. На державному рівні набувають поширення інструменти, починаючи з електронного уряду та демократії і закінчуючи технологіями розумного міста. Це сприяє сталому економічному та соціальному розвитку цілих регіонів і веде до підвищення рівня життя населення. Прогрес торкнувся і бізнес-сфери. Неможливо уявити діяльність будь-якого суб'єкта господарювання без використання інформаційних систем. Наприклад, компанії використовують Системи планування ресурсів підприємства для повної автоматизації бізнес-процесів, Customer Relationship Management Systems для інтеграції відносин з клієнтами в діяльність компанії, Business Intelligent Systems для аналізу діяльності та прийняття рішень тощо. Крім того, комп’ютери та цифрові технології вплинули на життя багатьох людей у світі. Завдяки їм з’явилися можливості дистанційної роботи, заробітку в Інтернеті, здійснення платежів через мобільні додатки, отримання інформації з різних куточків світу, не виходячи з дому, навчання в престижних світових університетах, не відвідуючи їх фізично тощо.

Усі ці приклади свідчать про позитивний вплив техніки на існування та розвиток суспільства та держави. Але водночас набуло поширення таке явище,

як кіберзлочинність, яка передбачає здійснення протиправних дій щодо фізичної чи юридичної особи, і навіть держави, із застосуванням комп'ютерних технологій. У результаті відбувається незаконне привласнення або порушення персональних даних, що призводить до втрати фінансових ресурсів особами, компаніями та державними установами. Статистичний аналіз даних щодо вартості витоку даних через кіберзлочинність показує її поступове зростання, тобто у 2022 році вона склала 4,35 млн доларів, що на 24,29% більше, ніж у 2014 році [**Ошибка! Источник ссылки не найден.**]. При цьому найбільш значні обсяги збитків характерні для таких галузей, як охорона здоров'я (10,1 млн. дол. США), фінансова (5,97 млн. дол. США), фармацевтична (5,01 млн. дол. США), технологічна (4,97 млн. дол. США) та ін. енергетики (4,72 млн дол. США) [**Ошибка! Источник ссылки не найден.**]. Ця інформація відображає не тільки фінансові втрати компаній, але й ідентифікує їх клієнтів, оскільки втрата персональних даних призводить до збільшення кількості жертв кібершахраїв після таких витоків даних. За даними дослідження Comparitech, 71,1 мільйона людей щороку страждають від кіберзлочинів; тобто близько 1% населення планети коли-небудь стикалося з цим видом злочинності [**Ошибка! Источник ссылки не найден.**]. Ці статистичні дані свідчать про те, що кіберзлочинність є глобальною проблемою, яка потребує відповідних заходів боротьби з нею та протидії. Тому в даному контексті постає необхідність визначити аспекти, критично важливі для формування портрета потенційної жертви кібершахрая. Поведінка жінок і чоловіків може відрізнятися в ситуаціях, коли вони стикаються з подібними злочинами. Крім того, ймовірний вплив може сформувати ставлення чоловіків і жінок до використання засобів особистого кіберзахисту. Такі аспекти потребують відповідного аналізу та дослідження, проведеного в даній науковій роботі.

Кіберзлочини набули актуальності в науковому середовищі з кінця 90-х років 20 століття, коли персональні комп'ютери стали доступні багатьом користувачам світу, а явище кібершахрайства почало набувати масового характеру. На сьогоднішній день сформувалося багато наукових напрямків, які

досліджують різні аспекти цієї проблеми. Найбільша кількість досліджень відноситься до галузі інформатики. Варто згадати публікації, які пропонують методи виявлення кіберзлочинів, такі як машинне навчання [**Ошибка! Источник ссылки не найден.**], нечітка логіка та аналіз даних [**Ошибка! Источник ссылки не найден.**], нейронні мережі [**Ошибка! Источник ссылки не найден.**], блокчейни [**Ошибка! Источник ссылки не найден.**], тощо.

Також виділено напрямок дослідження психологічних та соціальних аспектів кіберзлочинності. Дюпон і Холт обґрунтували критичну роль людського фактору в кіберзлочинах [**Ошибка! Источник ссылки не найден.**]. Лазар та ін. досліджували різницю у сприйнятті кібершахрайства жінками та чоловіками. Вони прийшли до висновку, що психосоціальні кіберзлочини є більш гендерними, тоді як соціально-економічні шахрайства жодним чином не залежать від статі [**Ошибка! Источник ссылки не найден.**]. Юрина Коннолі та Борріон висунули гіпотезу та перевірили свідомий вибір жертви заплатити викуп шахраю [**Ошибка! Источник ссылки не найден.**]. Witsenboer та ін. представила результати анкетування школярів, які дали змогу зробити висновок про необхідність набуття відповідних знань щодо використання засобів індивідуального захисту, починаючи зі шкільного віку [**Ошибка! Источник ссылки не найден.**]. Бретт Друрі та ін. доведено, що популярність соціальних мереж призвела до їх використання для кібершахрайства, оскільки вони можуть охопити велику аудиторію, створюючи широкі можливості для злочинців [**Ошибка! Источник ссылки не найден.**]. Лі та ін. досліджував ризики жертв фішингу та запропонував теоретичну перспективу розуміння впливу звичайних дій в Інтернеті на можливості кіберзлочинців [**Ошибка! Источник ссылки не найден.**].

Незважаючи на значний внесок науковців у дослідження проблем кіберзлочинності, питання їх гендерного аналізу та визначення критичних характеристик для формування портрета жертв кібершахрайства потребує додаткового дослідження.

Для цього дослідження автор використав результати опитування «Спеціальний Євробарометр 499: ставлення європейців до кібербезпеки (кіберзлочинності)», проведеного Європейською комісією у 2019 році [**Ошибка! Источник ссылки не найден.**]. Дані охоплюють респондентів з 28 європейських країн, таких як Бельгія, Данія, Греція, Іспанія, Фінляндія, Франція, Ірландія, Італія, Люксембург, Нідерланди, Австрія, Португалія, Швеція, Німеччина, Велика Британія, Болгарія, Кіпр, Чехія, Естонія, Угорщина, Латвія, Литва, Мальта, Польща, Румунія, Словаччина, Словенія та Хорватія. Загалом було опитано 27607 осіб, з них 48,44% чоловіків та 51,56% жінок. Найбільша кількість респондентів у Німеччині (16,00%), Великій Британії (13,00%), Франції (13,00%), Італії (12,00%), Іспанії (9,00%) та Польщі (8,00%).

Вибрані дані стосуються питань про типи пристроїв, якими користуються респонденти, типи їхньої онлайн-діяльності, обізнаність про ризики кіберзлочинності, особистий досвід щодо можливостей і результатів впливу на різні види кібершахрайства та інструменти для боротьби з потенційним шахрайством. Ця інформація була обрана для проведення гендерного аналізу та формування портрету потенційної жертви кіберзлочину, щоб виділити можливі категорії осіб з урахуванням їх гендерного розподілу за типом пристроїв, обізнаності про можливі ризики, впливу методів захисту тощо.

Основним методом цього дослідження є гендерний аналіз. Організація з безпеки та співробітництва в Європі визначає цю концепцію як «збір та аналіз даних, дезагрегованих за статтю, щоб виявити будь-який різний вплив дії на жінок і чоловіків, а також наслідки гендерних ролей і обов’язків. Це також передбачає якісний аналіз які допомагають з’ясувати, як і чому виникли ці різні ролі, відповідальність і вплив» [**Ошибка! Источник ссылки не найден.**]. Його реалізація також передбачає порівняння та оцінку поточних і майбутніх подій з урахуванням гендерних особливостей [**Ошибка! Источник ссылки не найден.**].

Застосування цього аналізу вимагає наступних етапів. На першому етапі формується та розподіляється за статтю база даних емпіричних даних. Далі

вибираються найбільш значущі характеристики, за якими буде проходити порівняння. На наступному етапі дані візуалізуються за допомогою діаграм і графіків. Ці результати аналізуються та формуються відповідні висновки щодо гендерних відмінностей у досліджуваних явищах.

Результати опитування показали, що у повсякденному житті найбільша кількість респондентів переважно використовують смартфони та персональні комп'ютери для виходу в Інтернет (рис. 2.103). І чоловіки, і жінки більше віддають перевагу смартфонам, ніж комп'ютерам через їх мобільність для виконання багатьох повсякденних завдань, хоча відсоток жінок у цьому випадку вищий. Чоловіки активніше за жінок користуються комп'ютерами, телевізорами та ігровими приставками.

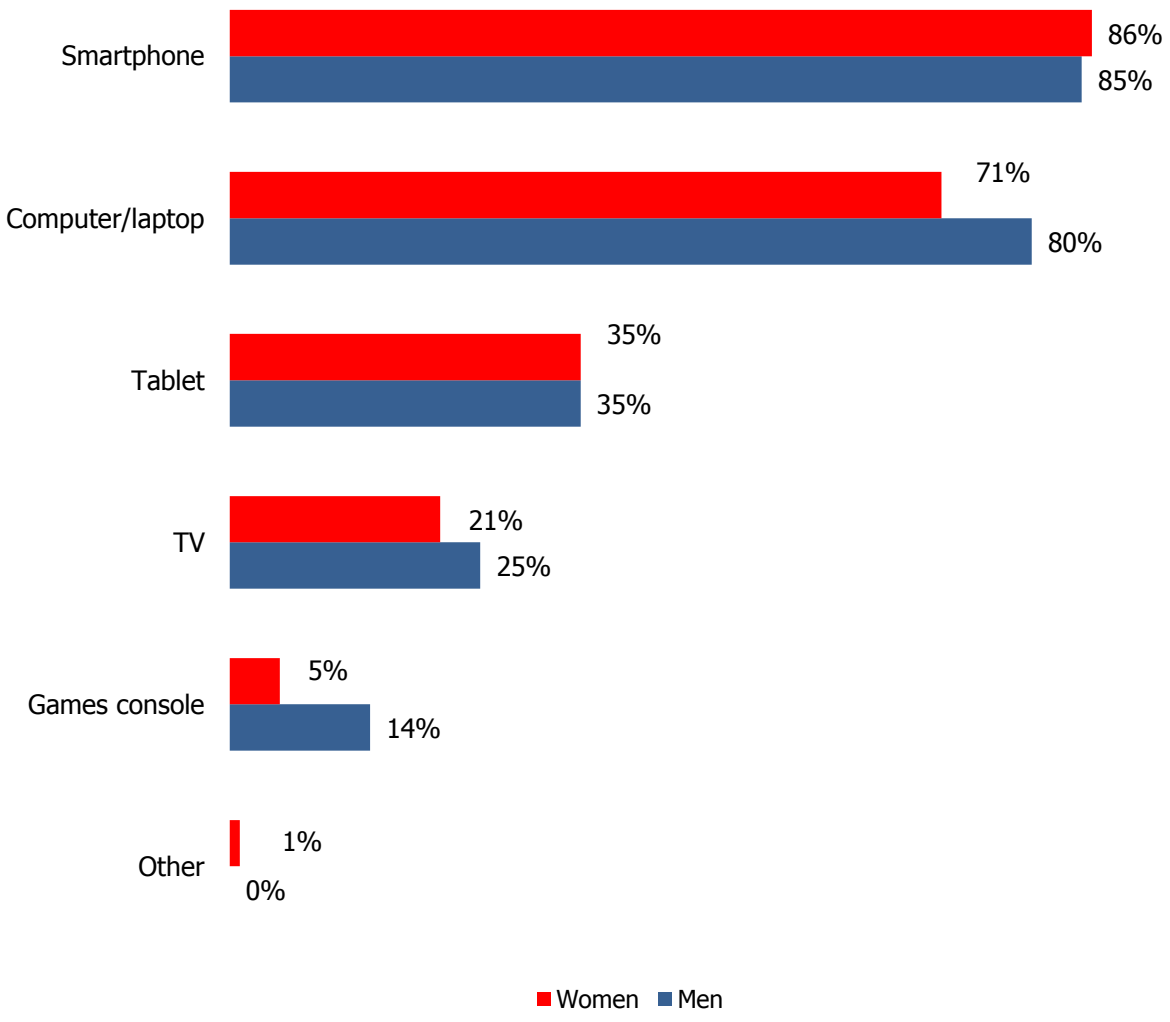


Рисунок 2.103 – Гендерний розподіл респондентів щодо використання пристроїв для доступу до Інтернету (створено за результатами опитування [**Ошибка! Источник ссылки не найден.**])

Для виконання яких завдань респонденти використовують смартфони, комп’ютери та інші пристрої? Аналіз активності в Інтернеті показує, що найпопулярнішими є листування електронною поштою, читання блогів і форумів, онлайн-банкінг, соціальні мережі, онлайн-магазини та миттєві повідомлення (рис. 2.104). Більше 50% чоловіків і жінок вибирають ці види. Більшість жінок віддають перевагу соціальним мережам як засобу спілкування. Оскільки вони в основному користуються смартфонами (рис. 2.104), ця інформація підтверджує, що їх більше цікавить швидкий і мобільний зв’язок, враховуючи їхню зайнятість на роботі, вдома та з дітьми. Чоловіки є активними користувачами в усіх інших онлайн-завданнях. Оскільки вони в основному використовують різні пристрої (рис. 2.104), разом із широким спектром діяльності, можна зробити висновок, що чоловіки проводять більше вільного часу в Інтернеті і не обмежуються лише комунікативними формами спілкування, як жінки.

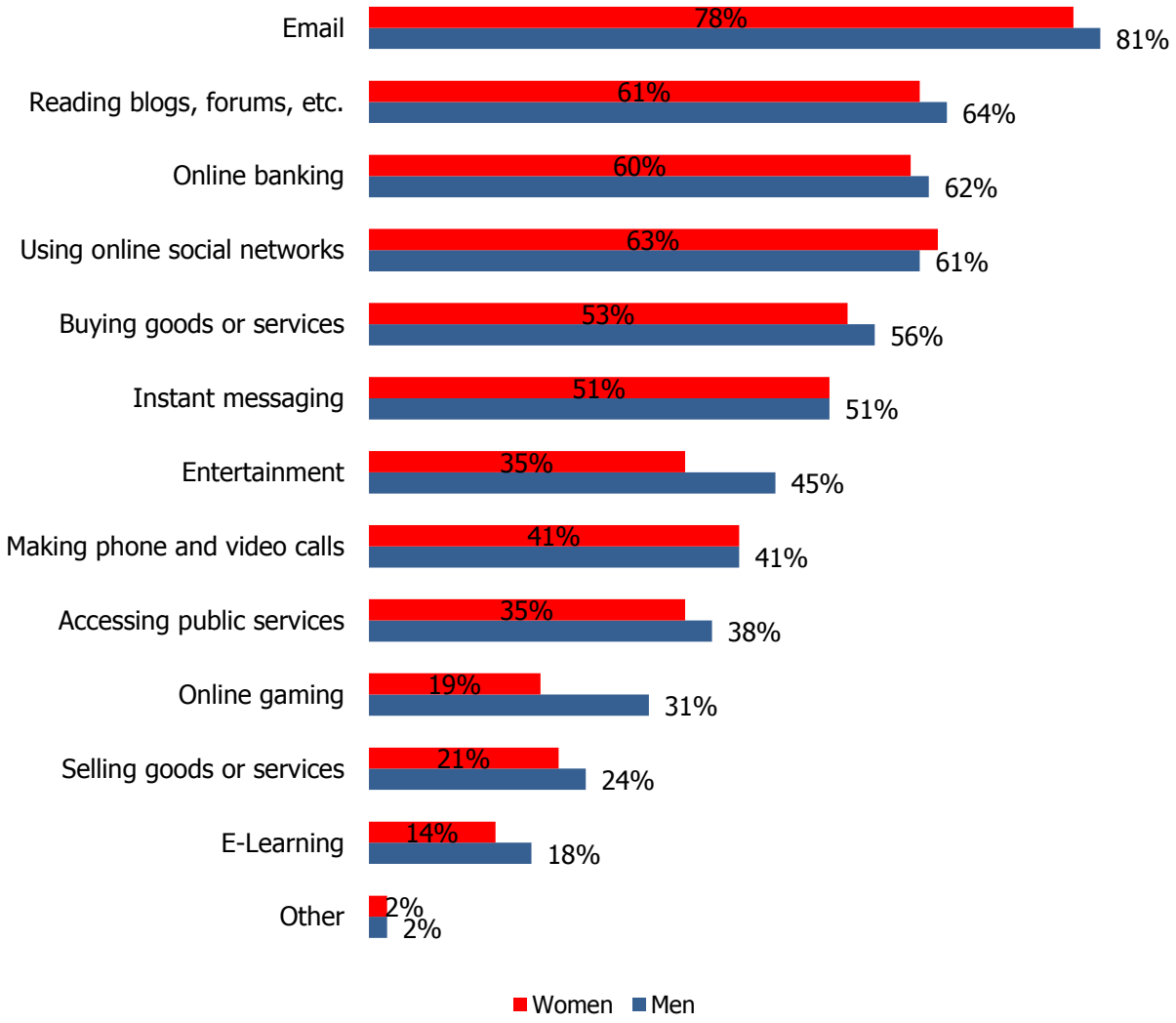


Рисунок 2.104 – Гендерний розподіл респондентів щодо їх активності в Інтернеті (створено за результатами опитування [**Ошибка! Источник ссылки не найден.**])

Високий відсоток активності електронної пошти можна віднести до найпопулярнішого кібершахрайства – фішингу, що відбувається шляхом надсилання електронних листів для отримання приватних даних користувача. Оскільки і чоловіки, і жінки переважно займаються цим видом діяльності, вони можуть швидко стати мішенню для кібершахраїв, хоча чоловіки будуть більш вразливими, ніж жінки.

Об’єктом кіберзагроз може бути кожен вид онлайн-діяльності. Тому кожен користувач повинен мати інформацію про ймовірність ризику стати жертвою

кіберзлочинців. 55% чоловіків вважають себе більш обізнаними про ризики кібершахрайства. З іншого боку, 55% жінок вважають себе менш обізнаними. Тобто чоловіки вважають себе більш впевненими у питаннях, пов’язаних із кіберзлочинністю, що може вплинути на зниження уваги до цієї проблеми. Оскільки жінки менш обізнані в цьому питанні, вони більше зосередяться на пошуку додаткових джерел інформації щодо особистого захисту.

Аналіз ситуацій, у яких респонденти постраждали від кіберзлочинності, показує, що чоловіки та жінки ставали жертвами у більшості випадків, коли отримували шахрайські повідомлення чи дзвінки (38% чоловіків та 33% жінок) або знаходили шкідливе програмне забезпечення на своїх комп’ютерах (31% чоловіків і 24% жінок) (рис. 2.105). Так, більше третини європейців хоча б раз стикалися з фішинговими, вішинговими та вірусними кібератаками.

Інші випадки стати жертвами відзначили 6%-13% користувачів, що значно менше, ніж для описаних вище ситуацій. Це пояснюється високими стандартами захисту інформації програмних додатків, за допомогою яких респонденти здійснюють онлайн-діяльність. Наприклад, користувач здійснює оплату транзакції через мобільний додаток банку або купує товари в електронних магазинах. У більшості випадків системи банків або інтернет-магазинів реалізують додаткові методи захисту, які вимагають іншої аутентифікації або багатоетапної перевірки, які шахраю важко обійти. Оскільки злочинці, окрім комп’ютерних навичок, також використовують психологічні чинники, такі як неуважність, довірливість та низька поінформованість користувача, такий підхід впливає на поширеність «Отримання шахрайських електронних листів або телефонних дзвінків».

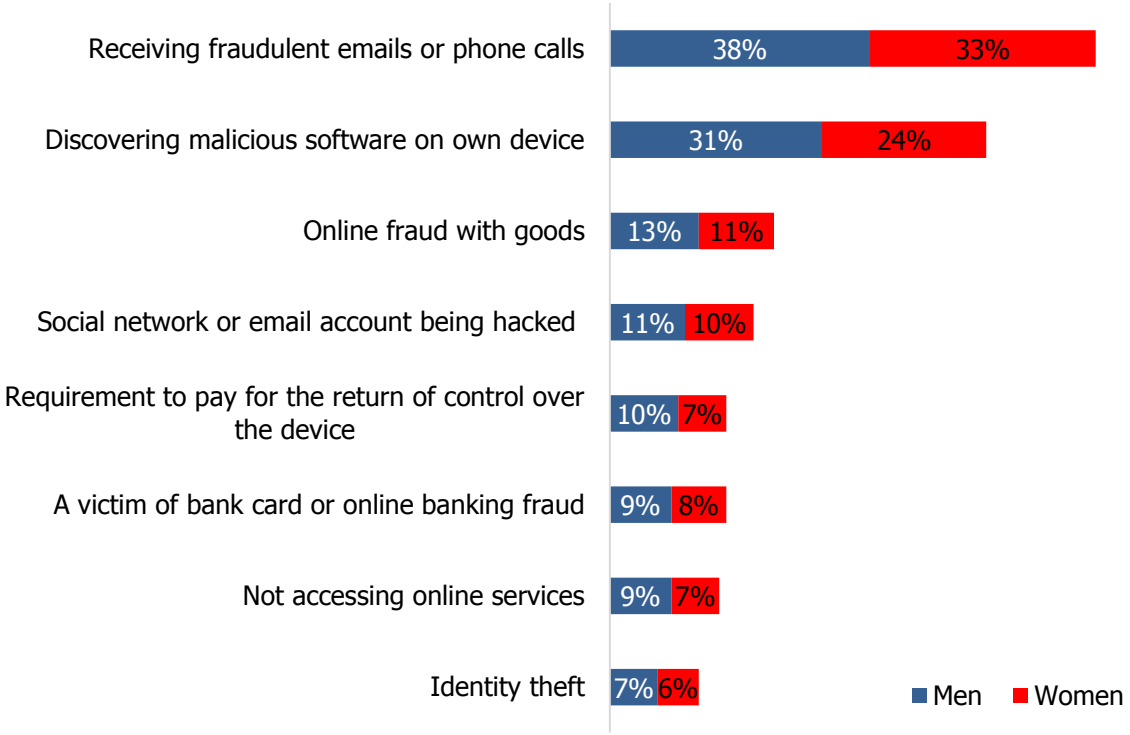


Рисунок 2.105 – Гендерний розподіл респондентів щодо ситуацій кібершахрайства (створено за результатами опитування [**Ошибка! Источник ссылки не найден.**])

Незалежно від типу кібершахрайства, чоловіки частіше, ніж жінки, ставали жертвами. Найбільш істотна різниця відчувається в ситуаціях з фішингом, вішингом і вірусними атаками. Це можна пояснити багатьма факторами: низька концентрація на особистому захисті, обізнаність у питаннях кібербезпеки, рівень довіри, освіта користувачів, вік, соціальний статус, рівень доходу, кількість часу, проведеного в Інтернеті, тощо. Однак ця статистика показує, що чоловіки більше вразливіші до кіберзлочинців, ніж жінки.

Рисунок 2.106 демонструє рівень занепокоєння респондентів щодо ситуацій кіберзлочинності. У всіх випадках жінки більше, ніж чоловіки, стурбовані можливим кібершахрайством. Їхні найбільші побоювання викликають шахрайство з онлайн-банкінгом, крадіжка особистих даних і зараження шкідливим програмним забезпеченням. Незалежно від ситуації жінки

можуть швидше відреагувати на потенційний злочин і вжити більш особистих заходів для протидії шахрайству.

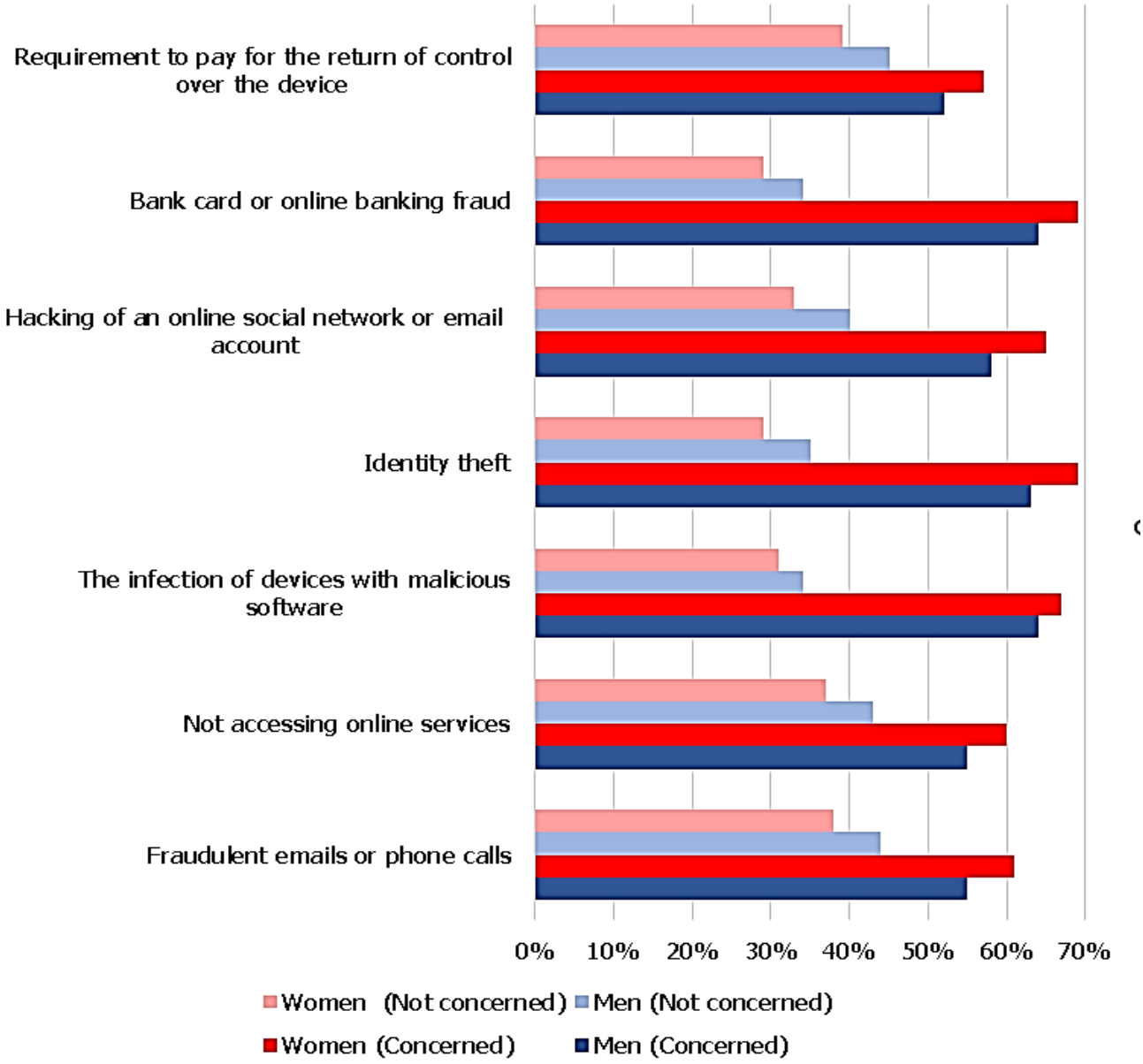


Рисунок 2.106 – Гендерний розподіл респондентів щодо впевненості в ситуаціях кібершахрайства (створено автором за результатами опитування [**Ошибка! Источник ссылки не найден.**])

На малюнку 2.107 показано, як респонденти захищаються від кібершахрайства.

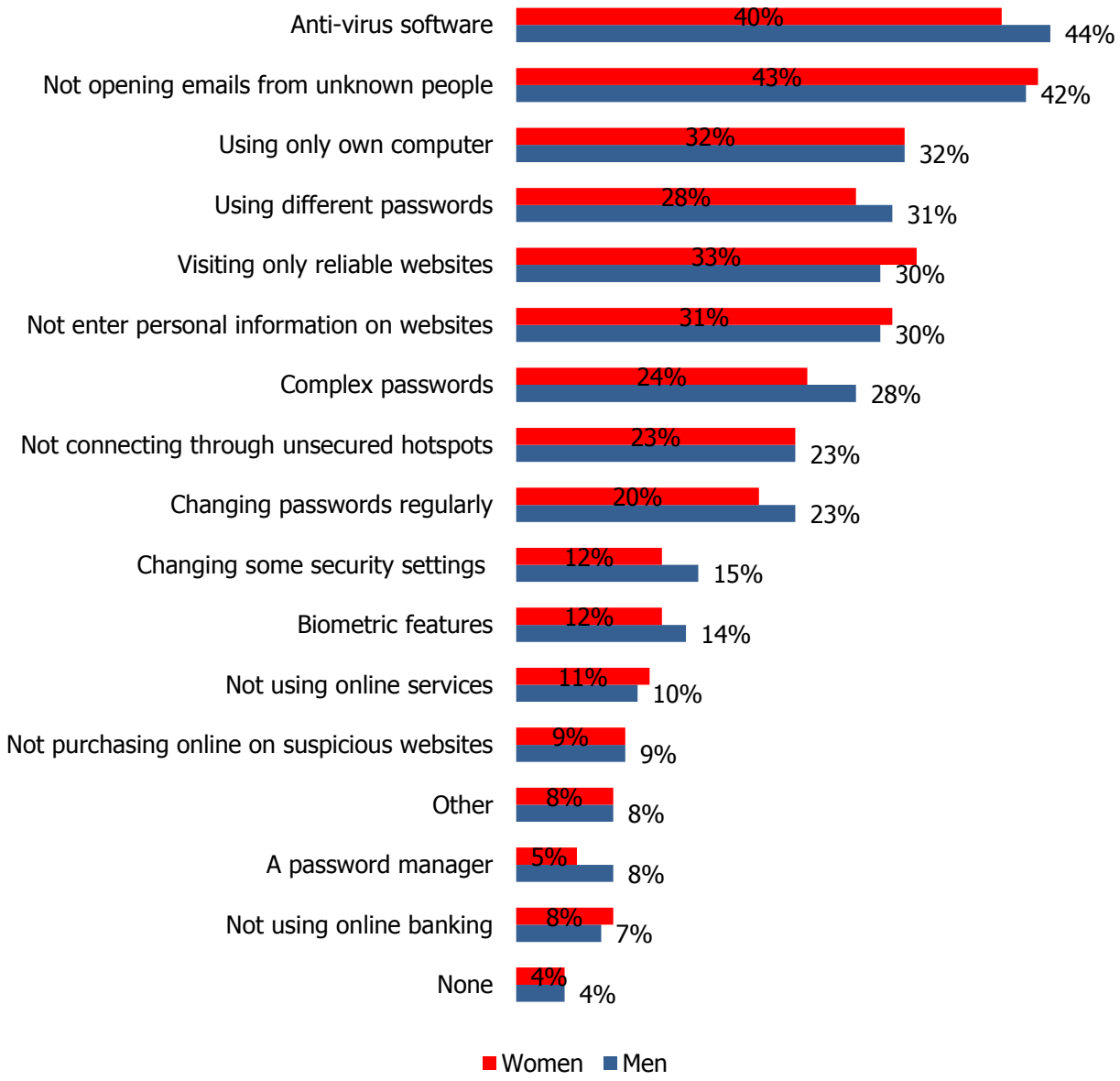


Рисунок 2.107 – Гендерний розподіл респондентів щодо способів захисту від кібершахрайства (створено автором за результатами опитування [**Ошибка! Источник ссылки не найден.**])

Жінки вважають за краще не відкривати пошту від незнайомців, не вводити особисту інформацію на сайтах, відвідувати лише надійні сайти, меншою мірою користуватися онлайн-сервісами та банкінгом. Чоловіки віддають перевагу антивірусним програмам, використовують різні і складні паролі, регулярно змінюють їх і налаштування безпеки, біометричні дані, менеджер паролів. Оскільки жінки рідше відвідують незнайомі сайти та не

відкривають незнайомі листи, це більшою мірою допомагає запобігти фішингу та отриманню особистих даних через фейкові сайти. Крім того, характер суб'єктивних оцінок, вжитих жінками, вказує на те, що вони не довіряють незнайомій або підозрілій інформації більше, ніж чоловіки, і віддають перевагу більш традиційним заняттям, ніж онлайн. Що стосується чоловіків, то, оскільки вони проводять багато часу в Інтернеті, вони більше, ніж жінки, намагаються використовувати більш складні та просунуті засоби особистого захисту, такі як біометрія, складні паролі та антивірусні програми.

За результатами гендерного аналізу буде сформовано портрет ймовірної жертви кіберзлочинців. Це може бути як чоловік, так і жінка, але жертвами в більшості ситуацій є чоловіки, ніж жінки. Жертви чоловічої статі є користувачами комп'ютерів і займаються різноманітною онлайн-діяльністю, переважно електронною поштою, що збільшує ймовірність зіткнутися з фішингом. Жінки-жертви — користувачі смартфонів, які проводять більше часу в соціальних мережах, але вони також можуть стати об'єктами фішингу та вішингу. Чоловіки-жертви вважають себе більш впевненими та усвідомлюють ризики кібершахрайства, ніж жінки-жертви, що може призвести до меншої уваги до підозрілої інформації. Менше жінок вважають себе добре поінформованими, що може вплинути на їхній підвищений інтерес до питань особистого захисту. Чоловіки-жертви менш стурбовані випадками шахрайства, які можуть змусити їх нехтувати різними заходами безпеки. Це впливає на їхній інтерес до впровадження методів захисту програмного забезпечення та ігнорування традиційних методів. Тобто ймовірною жертвою кібершахрая стане особа чоловічої статі, яка проводить значну кількість часу в Інтернеті, зі зниженою концентрацією уваги, підвищеною довірою до сторонньої інформації та впевненістю в технічних і програмних заходах безпеки.

Вирішення питань протидії кібершахрайству є актуальним як для наукової сфери, так і для практичної діяльності через швидку автоматизацію та цифровізацію багатьох процесів у державі, бізнесі, соціальному та особистому житті населення. Найефективнішим засобом боротьби з кібершахрайством є

сучасні математичні алгоритми, методи та інформаційні технології. Кіберзлочинці випереджають системи захисту та постійно вдосконалюють свої засоби боротьби з кіберзлочинністю. Однак вони також враховують різні аспекти потенційних жертв, щоб збільшити ефект своїх злочинів. Тому, коли компанії будують стратегії безпеки, вкрай важливо оцінювати різні характеристики жертв кібершахрайства, наприклад, стать.

У цьому дослідженні було проведено гендерний аналіз потенційних жертв кіберзлочинів, для якого використано результати опитування Європейської комісії. Дослідження показало, що чоловіки використовують більше різноманітних електронних пристроїв у повсякденному житті, ніж жінки. У той же час вони зацікавлені в електронному листуванні, читанні блогів і форумів, онлайн-банкінгу, покупках, іграх, розвагах і освіті. Жінки віддають перевагу смартфонам і спілкуванню в соціальних мережах. Тобто чоловіки проводять більше часу в Інтернеті, ніж жінки, тому вони частіше стають жертвами кібершахраїв, ніж жінки. Жінки вважають, що їм не вистачає знань про запобігання злочинам, і побоюються можливого шахрайства. Чоловіки вважають себе більш обізнаними про кіберризики, які можуть вплинути на їх довіру до інформації третіх сторін. Аналіз різних ситуацій кіберзлочинності показав, що чоловіки частіше стають жертвами, ніж жінки. Встановлено, що жінки дотримуються традиційних методів особистого захисту та не довіряють стороннім ресурсам, листам, сервісам і веб-сайтам. Сучасними методами особистої безпеки в основному користуються чоловіки. Тому, можливо, більше довіряють сторонній інформації в надії на системи захисту програмного забезпечення.

Підводячи підсумок, можна сказати, що потенційна жертва кібершахрайства – це людина, яка проводить значну кількість часу в Інтернеті, впевнена в заходах особистої безпеки та обізнаності про кіберризики, довіряє стороннім сервісам, сайтам і листам. Тобто може стати об’єктом вішингових, фішингових та вірусних кібератак. Результати гендерного аналізу можуть бути корисними при розробці систем моніторингу та кібербезпеки в компаніях, які

практикують онлайн-платіжні транзакції, наприклад, у банках, компаніях електронної комерції, індустрії розваг тощо. Ця інформація також може бути використана для створення портретів потенційних жертв кіберзлочинів для різних ділових та державних секторів.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**].

2.3.2 Розробка кіберпрофілів сучасних фінансових кіберзлочинів

У сучасному світі масштаби кіберзлочинності стрімко зростають. Це наслідок масової автоматизації різних сфер суспільства. Тому актуальною є проблема своєчасного виявлення та попередження кіберзлочинів, що дозволить суб'єктам господарювання та органам державної влади оперативно реагувати на масові загрози. На практиці використовується багато методів і засобів боротьби з кіберзлочинністю. Сьогодні є популярними програмні та кіберфізичні комплекси на основі сучасних математичних методів. Найбільш ефективними виявилися штучний інтелект, нейронні мережі, генетичні алгоритми, блокчейни, роботи тощо. Але, на нашу думку, поряд із такими потужними інструментами варто використовувати й профілювання кіберзлочинців.

Профілювання — це процес розробки профілів злочинців, які визначають потенційні загрози на основі заздалегідь визначених характеристик. Як правило, вони формуються за вже відомими випадками кіберзлочинів. Профілі використовують ретроспективний підхід до відображення подій, що дозволяє постійно оновлювати їх у разі появи нових даних про нові кіберзлочини.

Профілюванням повинні займатися спеціальні підрозділи кібербезпеки суб'єктів господарювання або державних установ. Оскільки саме вони збирають статистичну інформацію щодо кіберзлочинів, загроз, порушень, шахрайств, тощо. Але на практиці цьому виду інструменту кіберзахисту приділяють не досить багато уваги. Це пов'язано з трудомісткістю збору та обробки інформації, що потребує постійної реалізації даних процесів. Оскільки кіберзлочини удосконалюються з часом завдяки застосуванню новітніх технологій, які

впроваджують злочинці, то й процес профілювання повинен базуватися також на врахуванні великої кількості ознак кібершахраїв та злочинів. База даних ознак повинна поповнюватися новою інформацією, як результат здійснення різних видів кіберзагроз. При розробці профілів також треба враховувати, що вони корелюють як для злочину, так й для людини, яка його вчинила.

Основними характеристиками, які можна використовувати при побудові профілів, є поведінкові характеристики злочинців в Інтернеті. Ця необхідність викликана тим, що більшість фінансових та кіберзлочинів все ж таки реалізуються шляхом застосування онлайн-технологій. Це може відбуватися через соціальні мережі, Інтернет-магазини, блоги, онлайн-повідомлення, тощо. Відповідно, варто враховувати цю ознаку в процесі профілювання. Але тут є важливим в більшій мірі – частота відвідування конкретних інтернет-ресурсів, особливості використання тих чи інших інструментів і додатків, їх нестандартна та непередбачувана поведінка, проведення тіньових розрахунків та незаконних операцій з готівкою.

Можливе також використання такої характеристики, як відношення злочинця до державного законодавства країни. Усі злочинці нехтують дотриманням законів. Це пов’язано із тим, що більшість країн не мають чітко визначених заходів відповідальності за кіберзлочини. В основному за це передбачена адміністративна відповідальність, хоча в ряді країн, таких як США, Великобританія, за кіберзлочини впроваджена й кримінальна відповідальність. При визначенні профілю кіберзлочинця можна додавати й цю характеристику. Наприклад, на основі створеної бази даних кримінальних злочинців можна визначити тих, які мають спеціальну комп’ютерну освіту, або здійснювали кіберзлочини. Для таких зловмисників створюється «чорний лист», до якого можна постійно звертатися в процесі ідентифікації кіберзлочинців.

Географічне профілювання також є важливим, що дозволяє ідентифікувати злочинців за їхнім географічним розташуванням. Це можна зробити, відстежуючи джерела кібератак та IP-адреси, з яких надсилалися або надсилаються віруси, спам, кібератаки тощо. Існують спеціальні ресурси,

наприклад, як Лабораторія Касперського, які в режимі онлайн демонструють різні види кібератак, які направлені на конкретні країни та які відбуваються з інших країн світу. Використовуючи схожий принцип та можливості подібних ресурсів, можна визначити потенційні країни-атакери, тобто ті, з яких найбільше всього відбувається кібератак у світі. А також можна вивити країни-жертви, тобто ті, які піддаються кібератакам найбільше. Використання цієї інформації допоможе ідентифікувати клієнтів або транзакції, які відбуваються з країн-атакерів.

При розробці кіберпрофілю важливо враховувати такий аспект, як мотивація зловмисників. Тут можна виділити два її види. Перший пов'язаний з усвідомленим наміром кіберзлочинця вчинити злочин. Сюди можна віднести прагнення швидко отримати матеріальну вигоду, позбутися почуття залежності чи обмеженості, цікавість і допитливість, почуття задоволення від здобуття влади, потреба у визнанні чи самоствердженні, а також бажання висловити чи стверджувати політичні чи соціальні позиції, інтереси. Як правило, такі злочинці є хитрими та обізнаними в усіх аспектах кіберзлочинів і вони підуть до кінця, щоб отримати бажане. Це вид зловмисників дуже небезпечний для індивідів, бізнесу та держави. Їх виявлення потребуватиме більший обсяг зусиль та засобів, оскільки їх дії, як правило, є більш продуманими та організованими.

Другим мотиваційним напрямком виступають фінансові проблеми кіберзлочинця, пов’язані із необхідністю виплати кредитів, втратою роботи, низьким рівнем доходу, наявністю великої кількості дітей, хворих членів сім’ї, вирішенням проблем, пов’язаних із виплатою боргів в результаті азартних ігор, тощо. Як правило, такі зловмисники стають ними, виходячи з певних обставин в їхньому житті. Вони не можуть планувати такі злочини ідеально. Можливо, вони також не мають достатніх технічних знань для здійснення кіберзлочинів. Інструменти, якими вони користуються, є примітивними. Такі злочинці, як правило, вдаються до найбільш простих видів кіберзлочинів, наприклад, соціальна інженерія або розсилка програм-вимагачів. Знаходження таких індивідів є менш складним процесом, ніж у першому випадку, оскільки їх дії є

типовими та вони слідують заздалегідь придбаної інструкції або відомій їм техніці злочину, наприклад, підглянутої в Інтернеті.

Звичайно, що процес профілювання кіберзлочинців для різних видів кіберзлочинів буде відрізнятися в певних деталях, але підходи його формування будуть однаковими. Він використовує три види заходів, орієнтовані на особи, зловмисне програмне забезпечення або злочини.

Підхід, орієнтований на людину, передбачає здійснення аналізу дій кіберзлочинців та їх особистих характеристик. З цією метою використовуються різні матеріали, опубліковані в джерелах масової інформації та в Інтернеті, а також інформація, узята з професійних баз даних кіберзлочинців. В більшій мірі ці дані будуть відображати види та інструменти злочинів, характеристики жертв, особливості здійснення протиправних дій, тощо. Дана інформація дозволить сформувати реальне уявлення щодо особистості, яка скоїла злочин. Звичайно, що цей портрет може бути індивідуальним в кожному певному випадку, але більшість кіберзлочинців все одно матимуть багато спільного.

Підхід, який орієнтується на зловмисне програмне забезпечення, передбачає використання знань щодо схожих програм, які застосовують або застосовували кіберзлочинці. Збирається картотека відповідних програм на основі даних попереднього досвіду щодо кіберзагроз. Також сюди можуть включати і те забезпечення, яке потрібне для виявлення і протидії кіберзлочинів. Їх ядро направлене на аналіз патернів, які ідентифікують ситуації, що можуть бути ймовірною загрозою. Саме ці алгоритми можуть застосовуватися для виявлення інших ситуацій за умов коректування відповідних шаблонів перевірки.

Третій захід направлений конкретно на певний випадок кіберзлочину. Він застосовується з урахуванням тих самих методів, які реалізуються в попередніх двох підходах. Тобто є потреба у створенні відповідної бази даних кримінальних випадків із залученням комп’ютерних технологій, де відображаються їх ключові характеристики. Поєднання трьох підходів дає більший ефект, оскільки

одночасно спрямовується на суб’єкта, хто здійснює кіберзлочин, інструмент, за допомогою якого він реалізується, та об’єкт, на який спрямований злочин.

Для формування кіберпрофілів суттєве значення має використання криміналістичних методів ідентифікації традиційних злочинців, серед яких виділяють кримінально-розшукові, клінічні та статистичні методи. Підхід кримінального розслідування базується на проведенні різних видів експертизи для виявлення подібних випадків у минулому. Але такий підхід не буде ефективним на сто відсотків для формування профілю кіберзлочинців, оскільки цей вид злочинності є специфічним через швидкий розвиток технологій і методів, які використовують злочинці, що ускладнює їх ідентифікацію. Також він потребує постійного оновлення характеристик кіберзлочинців, кіберзлочинів та їх інструментів, але за часту це зробити складно за рахунок постійної модифікації технік та інструментів, які використовують злочинці. І завдяки цьому потрапити на місце злочину складно за рахунок віддаленості зловмисника або його маскування. Але в певних випадках проведення експертизи дозволятиме сформувати реальну картину подій.

Клінічний підхід забезпечує створення повної історії злочину, що дозволяє оцінити його основні характеристики. Його можна частково застосувати для виявлення кіберзлочинів, оскільки також є необхідність постійно оновлювати історію кіберзлочинів. Також може виникнути проблема у формуванні клінічної картини за рахунок відсутності відповідних фахівців, які б могли оцінити ознаки кіберзагрози. Деякі випадки навіть потребують сторонніх спеціалістів, які володіють знаннями та навиками використання відповідної мови програмування.

Найбільш ефективним є статистичний підхід, оскільки він передбачає використання спеціальних програмних засобів і статистичних методів, що дозволяє не тільки збирати відповідні характеристики, а й проводити відповідні розрахунки, які сприятимуть ідентифікації більш значної кількості злочинів. Коли інформації недостатньо, використовується непараметрична статистика. Коли доступна більш глибока інформація, регресії та байєсовські мережі можуть бути корисними в контексті профілювання. Методи класифікації та кластеризації

використовуються для формування профілів із типовими ознаками та стереотипним уявленням про злочинців, підозрюваних, свідків і жертв.

Статистичні методи кіберпрофілювання дають більший ефект, ніж інші, оскільки дозволяють вираховувати різні статистичні характеристики, моделювати ситуації потенційних кіберзагроз, прогнозувати ймовірність виникнення кіберзлочину, тощо. Звичайно, що використання тільки подібних методів не дозволить виявляти ситуації загроз на сто відсотків, але їх використання у сукупності з іншими підходами та методами профілювання сприятиме комплексній оцінці ситуацій в цілому або формуванню дій попередження, які дозволять підвищити увагу до конкретних транзакцій, або запитів, або користувачів системи.

В даній роботі було проведено формування кіберпрофілю потенційних злочинців, які здійснюють незаконні дії щодо кредитних операцій. У банківському секторі це досить серйозна проблема на сьогодні, оскільки пов’язана із спрощенням умов кредитування та отримання коштів онлайн різним категоріям населення за допомогою різних мобільних додатків. Існують також способи незаконно отримати кредити, використовуючи чужі дані, або цілеспрямовано отримати їх і не повернути, тому що у злочинця на меті здійснення кібершахрайства. Для реалізації даного завдання доцільно використати кластерний аналіз, який дозволить чітко сформувати профілі тих клієнтів, які будуть мати ознаки потенційної злочинної діяльності для банку.

Було застосовано такий різновид кластерного аналізу, як «Очікування-максимізація». Суть цього алгоритму полягає у виявленні кластерів (груп) клієнтів банку, які потенційно можуть бути пов’язані з кібершахрайством. Для його реалізації використовувалася клієнтська база даних одного з банків, яка містить понад 300 тис. спостережень. Кожен запис має 122 атрибути, які включають тип нерухомості клієнта, наявність автомобіля у клієнта, стать клієнта, кількість дітей, середній дохід і тип доходу клієнта, освіту, суму кредиту, суму щомісячний платіж тощо. В якості цільового атрибута використовується характеристика труднощів клієнта при виплаті кредиту. Якщо

його значення дорівнює «1», то, відповідно, у клієнта виникають ускладнення, які можуть сигналізувати про можливе кібершахрайство. Якщо значення «0» відповідає клієнту, то його профіль не викликає жодних підозр з боку кібербезпеки банку. Фрагмент вхідних даних представлено на рисунку 2.108.

Column1	Row	Description
1	SK_ID_CURR	ID of loan in our sample
2	TARGET	Target variable (1 - client with payment difficulties: he/she had late payment more than X days on at least one of the first Y installments of the loan in our sample, 0 - all other cases)
5	NAME_CONTRACT_TYPE	Identification if loan is cash or revolving
6	CODE_GENDER	Gender of the client
7	FLAG_OWN_CAR	Flag if the client owns a car
8	FLAG_OWN_REALTY	Flag if client owns a house or flat
9	CNT_CHILDREN	Number of children the client has
10	AMT_INCOME_TOTAL	Income of the client
11	AMT_CREDIT	Credit amount of the loan
12	AMT_ANNUITY	Loan annuity
13	AMT_GOODS_PRICE	For consumer loans it is the price of the goods for which the loan is given
14	NAME_TYPE_SUITE	Who was accompanying client when he was applying for the loan
15	NAME_INCOME_TYPE	Clients income type (businessman, working, maternity leave,...)
16	NAME_EDUCATION_TYPE	Level of highest education the client achieved
17	NAME_FAMILY_STATUS	Family status of the client
18	NAME_HOUSING_TYPE	What is the housing situation of the client (renting, living with parents, ...)
19	REGION_POPULATION_RELATIVE	Normalized population of region where client lives (higher number means the client lives in more populated region)
20	DAYS_BIRTH	Client's age in days at the time of application
21	DAYS_EMPLOYED	How many days before the application the person started current employment
22	DAYS_REGISTRATION	How many days before the application did client change his registration
23	DAYS_ID_PUBLISH	How many days before the application did client change the identity document with which he applied for the loan
24	OWN_CAR_AGE	Age of client's car
25	FLAG_MOBIL	Did client provide mobile phone (1=YES, 0=NO)
26	FLAG_EMP_PHONE	Did client provide work phone (1=YES, 0=NO)
27	FLAG_WORK_PHONE	Did client provide home phone (1=YES, 0=NO)
28	FLAG_CONT_MOBILE	Was mobile phone reachable (1=YES, 0=NO)
29	FLAG_PHONE	Did client provide home phone (1=YES, 0=NO)
30	FLAG_EMAIL	Did client provide email (1=YES, 0=NO)
31	OCCUPATION_TYPE	What kind of occupation does the client have
32	CNT_FAM_MEMBERS	How many family members does client have
33	REGION_RATING_CLIENT	Our rating of the region where client lives (1,2,3)
34	REGION_RATING_CLIENT_W_CITY	Our rating of the region where client lives with taking city into account (1,2,3)
35	WEEKDAY_APPR_PROCESS_START	On which day of the week did the client apply for the loan

Рисунок 2.108 – Фрагмент початкових даних для формування кіберпрофілю

Набір даних містить багато спостережень, які мають пропущену інформацію або мають викиди та екстремальні значення. В результаті, було відібрано 51 зміну з 122 та проведено їх якість. Для реалізації цього підходу використано програмно-аналітичний пакет «Deductor Studio Academic».

Результат оцінки на рисунку 2.109. З відібраних даних система виявила 4 атрибути є непридатними. Інші містять викиди та екстремальні значення. До них було застосовано процедури заповнення пропусків та усунення викидів, екстремальних значень. Результат даної процедури представлений на рисунку 2.110.

№	Столбец	Тип данных	Вид данных	Пропуски		Выбросы		Экстремальные		Кол-во уникальных	Качество данных	Резюме
				Кол-во	Действие	Кол-во	Действие	Кол-во	Действие			
6	FLAG_OWN_REALTY	ab Строковый	⋯ Дискретный							2	0,8999	Пригоден
7	CNT_CHILDREN	9.0 Вещественный	⋯ Дискретный			62	Заменять медианой	9	Заменять медианой	9	0,4115	Предобработка
8	AMT_INCOME_TOTAL	ab Строковый	⋯ Дискретный			441	Заменять наиболее ...	939	Заменять наиболее ...	416	0,5693	Предобработка
9	AMT_CREDIT	ab Строковый	⋯ Дискретный			1 906	Заменять наиболее ...	2 423	Заменять наиболее ...	2504	0,7854	Предобработка
10	AMT_ANNUITY	ab Строковый	⋯ Дискретный			2 673	Заменять наиболее ...			6119	0,8925	Предобработка
11	AMT_GOODS_PRICE	ab Строковый	⋯ Дискретный			1 832	Заменять наиболее ...	3 764	Заменять наиболее ...	426	0,6918	Предобработка
12	NAME_TYPE_SUITE	ab Строковый	⋯ Дискретный					1 479	Заменять наиболее ...	8	0,3171	Предобработка
13	NAME_INCOME_TYPE	ab Строковый	⋯ Дискретный			1 249	Заменять наиболее ...	10	Заменять наиболее ...	6	0,5800	Предобработка
14	NAME_EDUCATION_TYPE	ab Строковый	⋯ Дискретный					1 292	Заменять наиболее ...	5	0,4167	Предобработка
15	NAME_FAMILY_STATUS	ab Строковый	⋯ Дискретный			937	Заменять наиболее ...			5	0,7277	Предобработка
16	NAME_HOUSING_TYPE	ab Строковый	⋯ Дискретный			1 736	Заменять наиболее ...	1 817	Заменять наиболее ...	6	0,3284	Предобработка
17	REGION_POPULATION_RELATIVE	ab Строковый	⋯ Дискретный					3	Заменять наиболее ...	80	0,9255	Предобработка
18	DAYS_BIRTH	9.0 Вещественный	▬ Непрерывный								0,9596	Пригоден
19	DAYS_EMPLOYED	9.0 Вещественный	▬ Непрерывный								0,1327	Пригоден
20	DAYS_REGISTRATION	ab Строковый	⋯ Дискретный							9930	0,9792	Пригоден
21	DAYS_ID_PUBLISH	9.0 Вещественный	▬ Непрерывный								0,9422	Пригоден
22	FLAG_MOBIL	0/1 Логический	⋯ Дискретный							1	0,0000	Непригоден
✔ 23	FLAG_EMP_PHONE	0/1 Логический	⋯ Дискретный							2	0,5307	Пригоден
24	FLAG_WORK_PHONE	0/1 Логический	⋯ Дискретный							2	0,7914	Пригоден
25	FLAG_CONT_MOBILE	0/1 Логический	⋯ Дискретный					45	Заменять наиболее ...	2	0,0191	Предобработка
26	FLAG_PHONE	0/1 Логический	⋯ Дискретный							2	0,8032	Пригоден
27	FLAG_EMAIL	0/1 Логический	⋯ Дискретный					1 374	Заменять наиболее ...	2	0,3087	Предобработка
28	OCCUPATION_TYPE	ab Строковый	⋯ Дискретный			327	Заменять наиболее ...	221	Заменять наиболее ...	19	0,7510	Предобработка
29	CNT_FAM_MEMBERS	7 Дата/Время	⋯ Дискретный	1	Заменять мед...	55	Ограничивать	14	Ограничивать	10	0,5593	Предобработка
30	REGION_RATING_CLIENT	9.0 Вещественный	⋯ Дискретный							3	0,6736	Пригоден
31	REGION_RATING_CLIENT_W_CITY	9.0 Вещественный	⋯ Дискретный							3	0,6660	Пригоден
32	WEEKDAY_APPR_PROCESS_START	ab Строковый	⋯ Дискретный							7	0,9718	Пригоден
33	HOUR_APPR_PROCESS_START	9.0 Вещественный	▬ Непрерывный			33	Ограничивать				0,7937	Предобработка
34	REG_REGION_NOT_LIVE_REGION	0/1 Логический	⋯ Дискретный					433	Заменять наиболее ...	2	0,1268	Предобработка
35	REG_REGION_NOT_WORK_REGION	0/1 Логический	⋯ Дискретный					1 388	Заменять наиболее ...	2	0,3110	Предобработка
36	LIVE_REGION_NOT_WORK_REGION	0/1 Логический	⋯ Дискретный					1 056	Заменять наиболее ...	2	0,2538	Предобработка
37	REG_CITY_NOT_LIVE_CITY	0/1 Логический	⋯ Дискретный							2	0,5247	Пригоден
38	REG_CITY_NOT_WORK_CITY	0/1 Логический	⋯ Дискретный							2	0,8848	Пригоден
39	LIVE_CITY_NOT_WORK_CITY	0/1 Логический	⋯ Дискретный							2	0,7632	Пригоден
40	ORGANIZATION_TYPE	ab Строковый	⋯ Дискретный			978	Заменять наиболее ...	1 706	Заменять наиболее ...	58	0,6917	Предобработка
41	EXT_SOURCE_2	ab Строковый	⋯ Дискретный							21988	0,9900	Пригоден
42	FONDKAPREMONT_MODE	ab Строковый	⋯ Дискретный					1 548	Заменять наиболее ...	5	0,4928	Предобработка
43	HOUSETYPE_MODE	ab Строковый	⋯ Дискретный					255	Заменять наиболее ...	4	0,5332	Предобработка
44	TOTALAREA_MODE	ab Строковый	⋯ Дискретный					11 119	Заменять наиболее ...	2644	0,5123	Предобработка
45	WALLSMATERIAL_MODE	ab Строковый	⋯ Дискретный					1 562	Заменять наиболее ...	8	0,5746	Предобработка
46	EMERGENCYSTATE_MODE	ab Строковый	⋯ Дискретный					223	Заменять наиболее ...	3	0,6677	Предобработка
47	OBS_30_CNT_SOCIAL_CIRCLE	7 Дата/Время	⋯ Дискретный	13 097	Оставить без и...	263	Оставить без измен...			12	0,0000	Непригоден
48	DEF_30_CNT_SOCIAL_CIRCLE	7 Дата/Время	⋯ Дискретный	21 134	Оставить без и...	35	Оставить без измен...	9	Оставить без измен...	6	0,0000	Непригоден
49	OBS_60_CNT_SOCIAL_CIRCLE	7 Дата/Время	⋯ Дискретный	13 164	Оставить без и...	254	Оставить без измен...			12	0,0000	Непригоден
50	DEF_60_CNT_SOCIAL_CIRCLE	7 Дата/Время	⋯ Дискретный	22 030	Оставить без и...	95	Оставить без измен...	18	Оставить без измен...	5	0,0000	Непригоден

Рисунок 2.109 – Результат оцінки якості початкових даних

№	Столбец	Тип данных	Вид данных	Пропуски		Выбросы		Экстремальные		Кол-во уникальных	Качество данных	Резюме
				Кол-во	Действие	Кол-во	Действие	Кол-во	Действие			
1	SK_ID_CURR	9.0 Вещественный	Непрерывный								0,9999	Пригоден
2	TARGET	9.0 Вещественный	Дискретный							1	0,0000	Непригоден
3	NAME_CONTRACT_...	ab Строковый	Дискретный							1	0,0000	Непригоден
4	CODE_GENDER	ab Строковый	Дискретный							2	0,9855	Пригоден
5	FLAG_OWN_CAR	ab Строковый	Дискретный							2	0,8875	Пригоден
6	FLAG_OWN_REALTY	ab Строковый	Дискретный							2	0,8999	Пригоден
7	CNT_CHILDREN	9.0 Вещественный	Дискретный							3	0,7405	Пригоден
8	AMT_INCOME_TOTAL	ab Строковый	Дискретный			873	Заменять наиболее ...	60	Заменять наиболее ...	46	0,7961	Предобработка
9	AMT_CREDIT	ab Строковый	Дискретный			1 445	Заменять наиболее ...	10 076	Заменять наиболее ...	571	0,7492	Предобработка
10	AMT_ANNUITY	ab Строковый	Дискретный			2 019	Заменять наиболее ...	16 129	Заменять наиболее ...	3446	0,8442	Предобработка
11	AMT_GOODS_PRICE	ab Строковый	Дискретный			643	Заменять наиболее ...	5 239	Заменять наиболее ...	49	0,7051	Предобработка
12	NAME_TYPE_SUITE	ab Строковый	Дискретный							2	0,5328	Пригоден
13	NAME_INCOME_TYPE	ab Строковый	Дискретный							3	0,7805	Пригоден
14	NAME_EDUCATION_...	ab Строковый	Дискретный							2	0,6379	Пригоден
15	NAME_FAMILY_STA...	ab Строковый	Дискретный							4	0,7415	Пригоден
16	NAME_HOUSING_TY...	ab Строковый	Дискретный							1	0,0000	Непригоден
17	REGION_POPULATI...	ab Строковый	Дискретный							79	0,9280	Пригоден
18	DAYS_BIRTH	9.0 Вещественный	Непрерывный								0,9596	Пригоден
19	DAYS_EMPLOYED	9.0 Вещественный	Непрерывный								0,1327	Пригоден
20	DAYS_REGISTRATION	ab Строковый	Дискретный							9930	0,9792	Пригоден
21	DAYS_ID_PUBLISH	9.0 Вещественный	Непрерывный								0,9422	Пригоден
22	FLAG_MOBIL	0/1 Логический	Дискретный							1	0,0000	Непригоден
23	FLAG_EMP_PHONE	0/1 Логический	Дискретный							2	0,5307	Пригоден
24	FLAG_WORK_PHONE	0/1 Логический	Дискретный							2	0,7914	Пригоден
25	FLAG_CONT_MOBILE	0/1 Логический	Дискретный					45	Заменять наиболее ...	2	0,0191	Предобработка
26	FLAG_PHONE	0/1 Логический	Дискретный							2	0,8032	Пригоден
27	FLAG_EMAIL	0/1 Логический	Дискретный					1 374	Заменять наиболее ...	2	0,3087	Предобработка
28	OCCUPATION_TYPE	ab Строковый	Дискретный							13	0,8183	Пригоден
29	CNT_FAM_MEMBERS	Дата/Время	Непрерывный								0,4611	Пригоден
30	REGION_RATING_C...	9.0 Вещественный	Дискретный							3	0,6736	Пригоден
31	REGION_RATING_C...	9.0 Вещественный	Дискретный							3	0,6660	Пригоден
32	WEEKDAY_APPR_P...	ab Строковый	Дискретный							7	0,9718	Пригоден
33	HOUR_APPR_PROC...	9.0 Вещественный	Непрерывный								0,8521	Пригоден
34	REG_REGION_NOT_...	0/1 Логический	Дискретный					433	Заменять наиболее ...	2	0,1268	Предобработка
35	REG_REGION_NOT_...	0/1 Логический	Дискретный					1 388	Заменять наиболее ...	2	0,3110	Предобработка
36	LIVE_REGION_NOT_...	0/1 Логический	Дискретный					1 056	Заменять наиболее ...	2	0,2538	Предобработка
37	REG_CITY_NOT_LIV...	0/1 Логический	Дискретный							2	0,5247	Пригоден
38	REG_CITY_NOT_WO...	0/1 Логический	Дискретный							2	0,8848	Пригоден
39	LIVE_CITY_NOT_WO...	0/1 Логический	Дискретный							2	0,7632	Пригоден
✔ 40	ORGANIZATION_TYPE	ab Строковый	Дискретный			1 033	Заменять наиболее ...	726	Заменять наиболее ...	19	0,7594	Предобработка
41	EXT_SOURCE_2	ab Строковый	Дискретный							21988	0,9900	Пригоден
42	FONDKAPREMONT_...	ab Строковый	Дискретный							2	0,7367	Пригоден
43	HOUSETYPE_MODE	ab Строковый	Дискретный							2	0,9819	Пригоден
44	TOTALAREA_MODE	ab Строковый	Дискретный							1	0,0000	Непригоден
45	WALLSMATERIAL_M...	ab Строковый	Дискретный							3	0,8237	Пригоден
46	EMERGENCYSTATE...	ab Строковый	Дискретный							2	0,9920	Пригоден

Рисунок 2.110 – Кінцевий результат оцінки якості початкових даних

Хоча проведені процедури не дозволили повністю усунути всі недоліки, але якість масиву даних значно покращилася. Також було виявлено ще непридатні для аналізу дані, які будуть не враховані в подальшому моделюванні.

Далі було проведено кластерний аналіз «Очікування-максимізація», який використовувався як алгоритм для формування профілю кібершахраїв із кредитними операціями. Це ітераційний метод, який формує групи за максимальною правдоподібністю або максимальними апостеріорними оцінками параметрів. Перевагою цього алгоритму є пошук прихованих змінних, які можуть сильно впливати на формування цільової змінної, що може бути корисним, оскільки аналітик або фахівець з кібербезпеки можуть не виявити всіх можливих ознак. У результаті було сформовано 10 кластерів користувачів, які можна проаналізувати на предмет можливого кібершахрайства з кредитними операціями. Результати зв’язків між кластерами представлені на рисунку 2.111. На рисунку 2.111 можна побачити міцність кластерів, максимальну похибку розпізнання та середню похибку розпізнання.

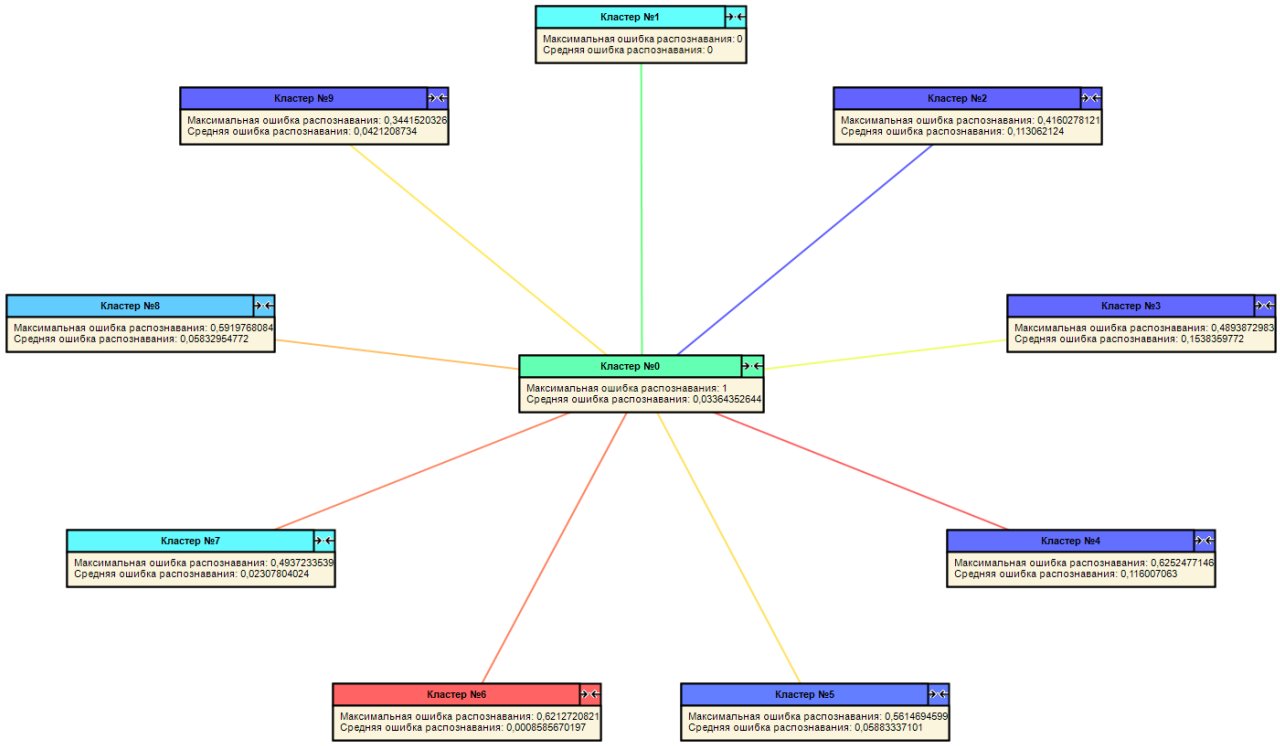


Рисунок 2.111 – Результати зв’язків між кластерами

Найбільш потужними є кластери № 4, 3, 1, 7, 9. Можна було б зменшити їх кількість до 5, але залишимо дану конфігурацію, оскільки це дозволить нам сформувати більш детальні профілі злочинців, виходячи із наповнюваності кластерів. За умови отримання нової інформації, такий обсяг профілів сприятиме їх уточненню та більш детальній класифікації кіберзлочинців.

Фрагмент отриманих профілів, виходячи з ознаки кластерів, представлено на рисунку 2.112. Всі інші частини профілів наведено у додатку Д.

Найбільш наповненими є перші п’ять кластерів. Тобто при ідентифікації потенційних кіберзлочинців, ймовірність того, що вони належатимуть даним кластерам є більшою, ніж у інших випадках. Вони формуватимуть основний набір характеристик, які відповідатимуть потенційним загрозам. Рисунок 2.112 показує, що значення типу будинку є невизначеним у більшості випадків (атрибут “HOUSETYPE_MODE”). Візуалізація наповненості кластерів клієнтів за даною характеристикою представлена на рисунку 2.113. Це пов’язано із відсутністю вхідних даних і зазначенням їх як невизначений тип.

Рисунок 2.112 – Фрагмент профілів кіберзлочинців

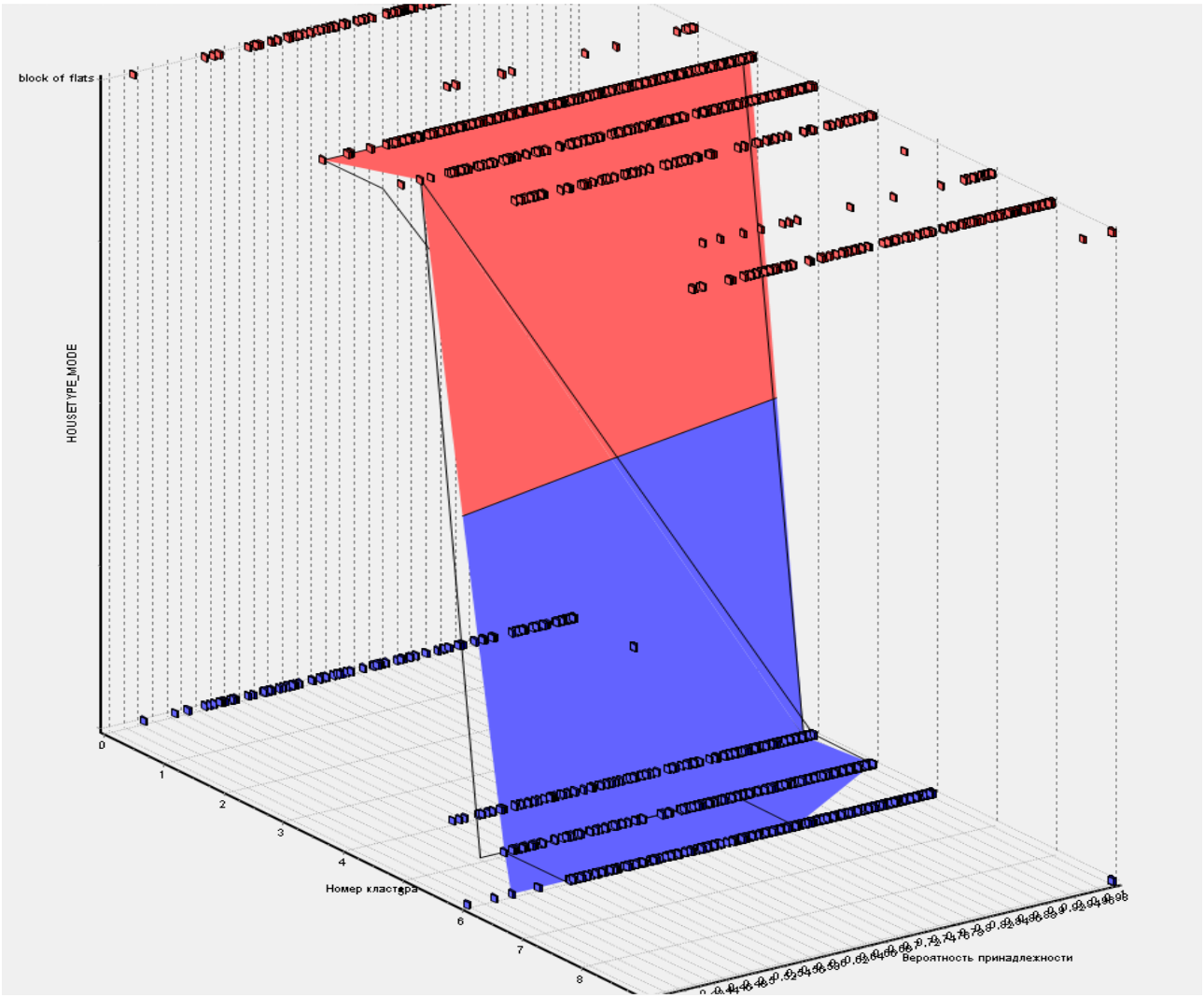


Рисунок 2.113 – Візуалізація характеристики “HOUSETYPE_MODE”

Але клієнти, які відносяться до 3, 1 та 7 кластеру і які були ідентифіковані, як шахраї, в переважній більшості мали «Block of Flats». Це можна пояснити тим, що більшість клієнтів банку є власниками цього типу нерухомості. Саме такі клієнти звертаються до банків за кредитами на купівлю будинку. Відповідно, вони можуть належати до групи кібершахраїв.

Якщо аналізувати змінну “FONDKAPREMONT_MODE”, то також видно, що умови якості житла, яким володіє індивід, є значними для 3, 1 та 7 кластерів. Візуалізація наповненості кластерів клієнтів за даною характеристикою представлена на рисунку 2.114. Для інших вони є невизначеними, томи при аналізі характеристик можуть бути не враховані. Але потенційні зловмисники, які відповідають кластерам 3, 1 та 7, можуть мати мотивацією до здійснення

злочину квартирне питання – або взяття коштів на поліпшення житлових умов, або зміни типу житла. Звісно, що якщо людина намагається узяти кредит за цими цілями, то це не означає, що їй банк повинен відмовити. Отримання інформації, що клієнт може бути потенційним шахраєм, оскільки попадає в даний кластер за цими ознаками, означатиме проведення додаткових перевірок та аналізу за іншими критеріями.

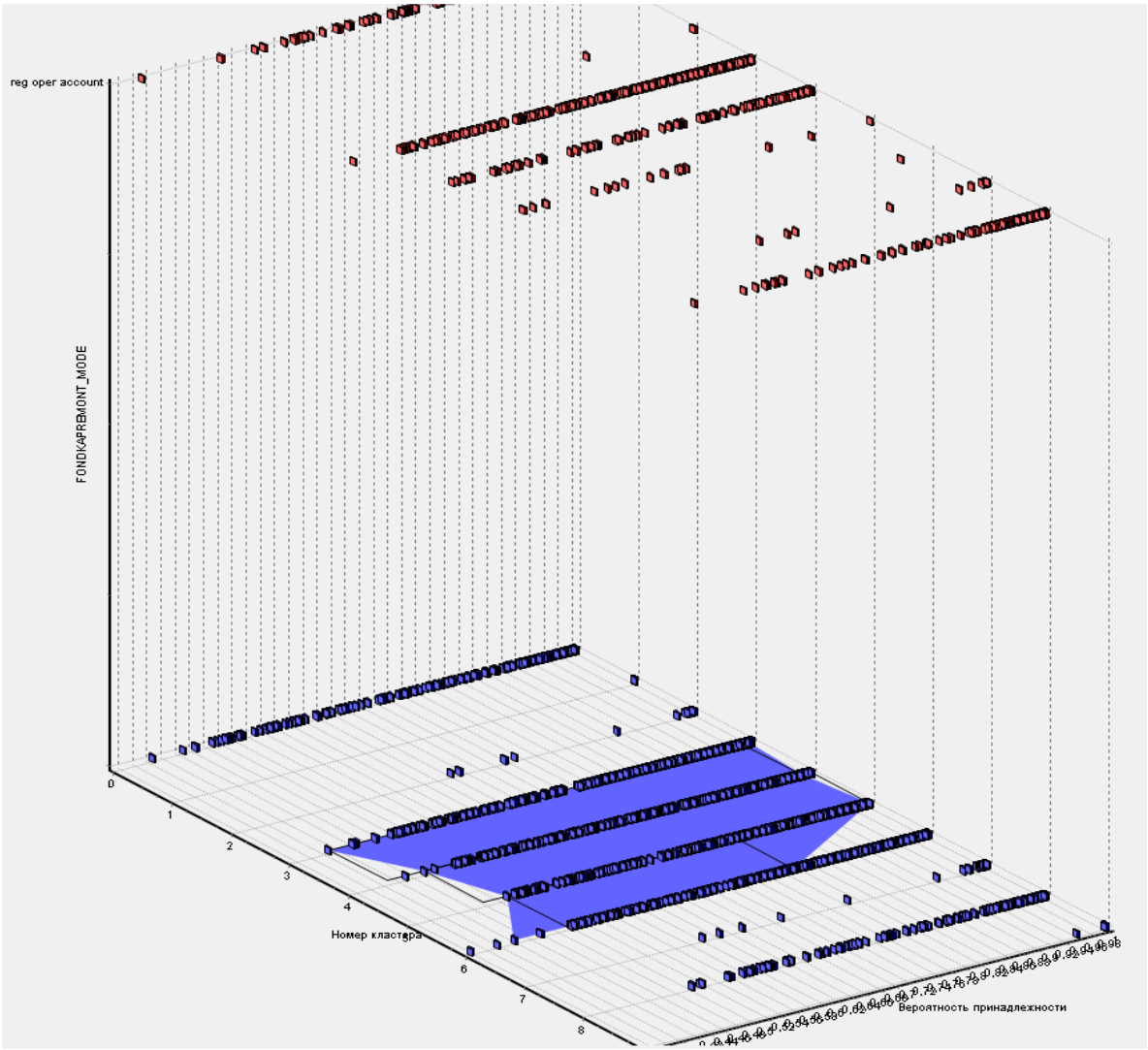


Рисунок 2.114 – Візуалізація характеристики “FONDKAPREMONT_MODE”

Що стосується наступної ознаки, такої як “ORGANIZATION_TYPE”, то вона свідчитиме про вид діяльності клієнта. Візуалізація наповненості кластерів клієнтів за даною характеристикою представлена на рисунку 2.115. Можна побачити, що 1-й кластер сформували переважно ті, чий вид діяльності ідентифікується як “XNA”. Оскільки використана для дослідження база даних не

містить розшифровок, то важко сказати, яка діяльність має дану абревіатуру. До 4, 3 та 7 кластерів увійшли ті, хто класифікується як “Business Entity Type 3”. Також 20,2% 4-го кластеру сформовані клієнтами, які є самозайнятими. Ймовірно, що ці зловмисники мають проблеми з їх видом діяльності, які потребують додаткових інвестицій, або фінансових засобів на розширення, або погашення боргів, тому вони й вдаються до злочинних схем із незаконним отриманням або неповерненням кредитних коштів.

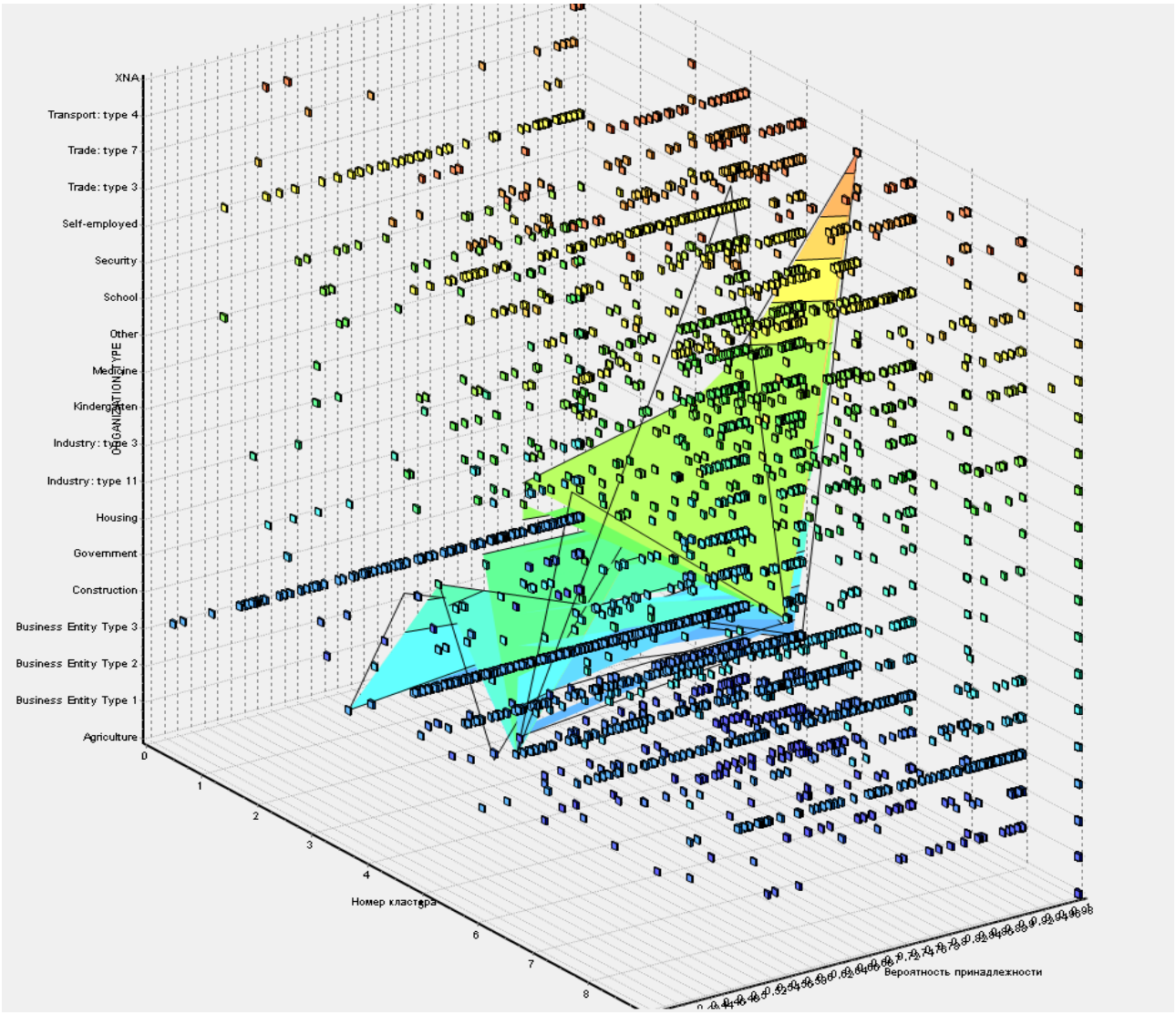


Рисунок 2.115 – Візуалізація характеристики “ORGANIZATION_TYPE”

При формуванні профіля важливою ознакою є сімейний статус “NAME_FAMILY_STATUS”. Візуалізація наповненості кластерів клієнтів за даною характеристикою представлена на рисунку 2.116. Значення кластерів

показують, що переважна більшість клієнтів банку є одруженими / заміжніми. На нашу думку, до аналізованого виду кібершахрайств можуть вдаватися саме сімейні люди, які звертаються до злочину завдяки неможливості утримувати сім'ю або наявності хворого члена сім'ї, або задля реалізації бажань, що потребують значних фінансових ресурсів.

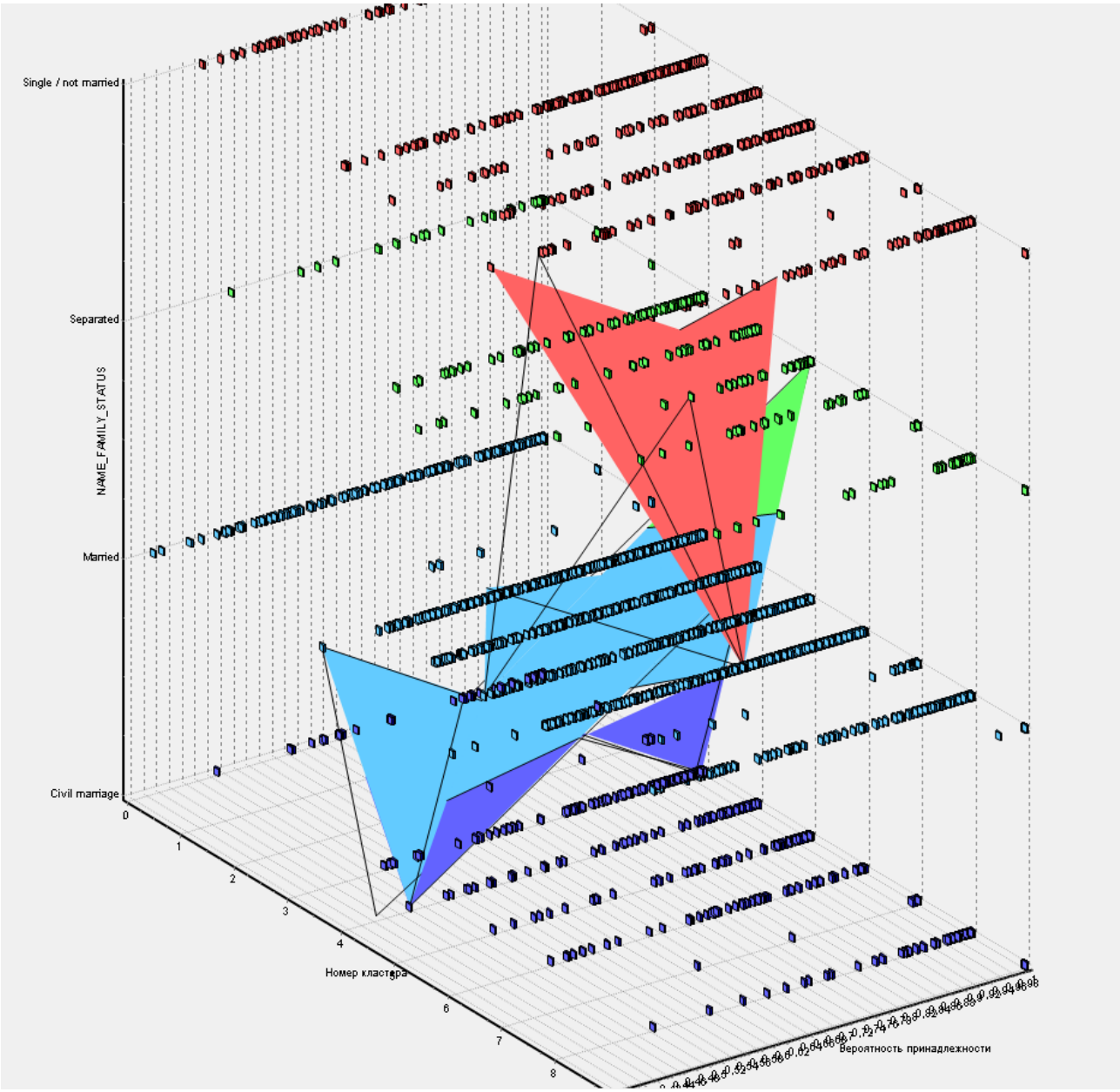


Рисунок 2.116 – Візуалізація характеристики "NAME_FAMILY_STATUS"

Цікавим виявилася така характеристика, як тип освіти. Візуалізація наповненості кластерів клієнтів за даною характеристикою представлена на рисунку 2.117. Виявилося, що шахрайством можуть займатися ті, хто завершили

середню школу або мають вищу освіту. Переважно зловмисники відносяться до першої категорії (завершили середню школу). Ці особистості можливо не мають певних досягнень у житті, тому для швидкої самореалізації вони можуть вдатися до найпростішого виду кібершахрайств, такого як шахрайства з кредитними операціями. Відповідно, такі злочинці слідують типовим схемам і можуть вираховуватися швидше. Але приблизно 10-20% клієнтів мають вищу освіту, при цьому кожен кластер містить таких потенційних злочинців. Ця категорія є достатньо вмотивованою, оскільки, можливо, намагаються здійснити злочин заради самоствердження або підтримки гострих почуттів. Виявляти таких кандидатів досить складно, оскільки вони мають не тільки освіту, але й роботу, що гарантує формуванню гарної кредитної історії.

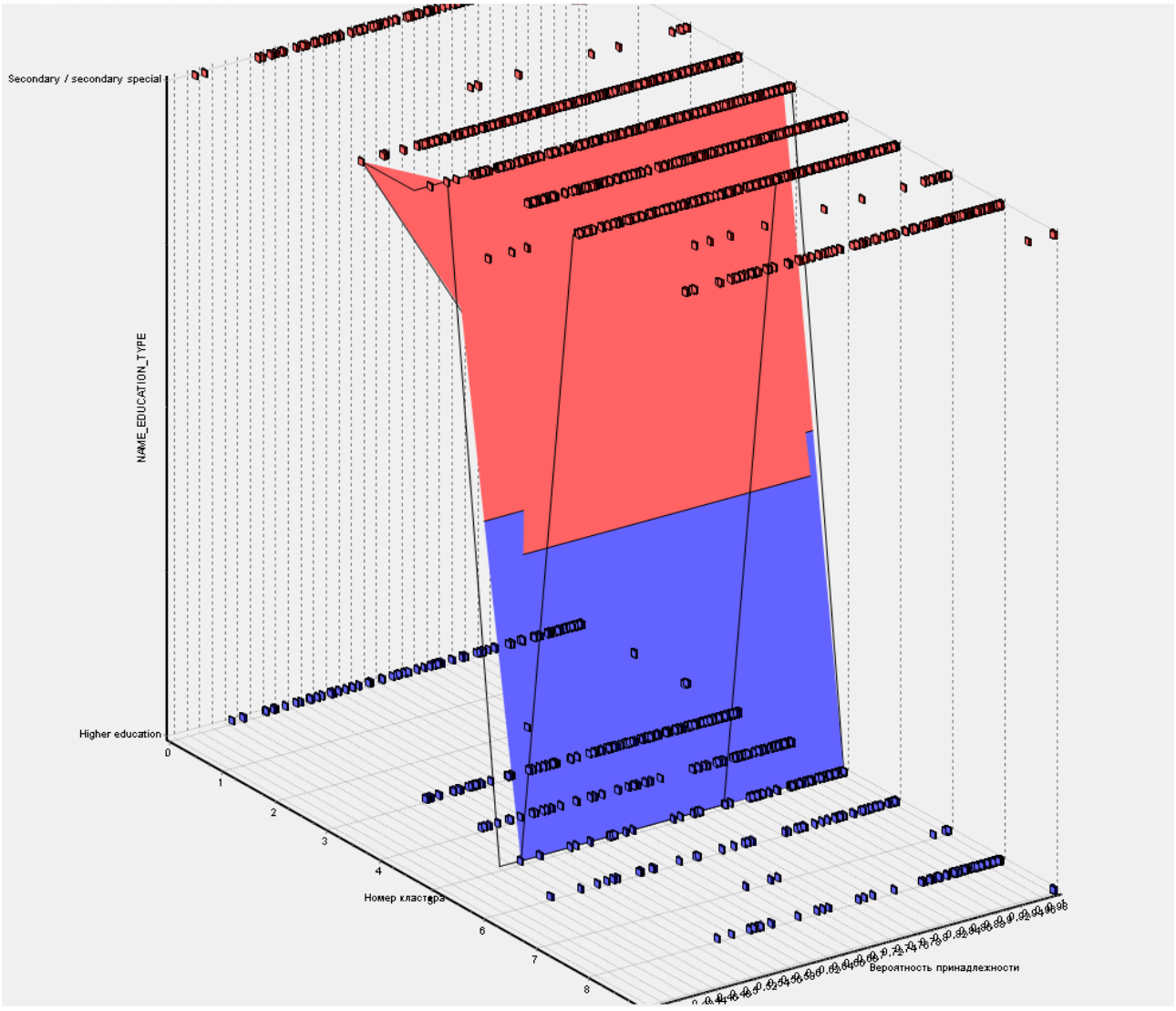


Рисунок 2.117 – Візуалізація характеристики “NAME_EDUCATION_TYPE”

Що стосується отриманих доходів “NAME_INCOME_TYPE”, то візуалізація наповненості кластерів клієнтів за даною характеристикою представлена на рисунку 2.118. Виявляється, що переважна кількість зловмисників отримують дохід від роботи або в результаті участі в комерційних об’єднаннях. Але цікавість викликає кластер під номером 1, куди попали ті шахраї, які отримують пенсії. З урахуванням визначення попередніх ознак виявляється, що даний кластер містить зловмисників, які знаходяться на пенсії, завершили середню школу, мають сім’ю, та можливо проблеми житлового характеру. Тобто, даний кластер є високо ризикованим щодо надання кредиту та його неповернення. Навіть, якщо клієнт і не має мотивації щодо цілеспрямованого шахрайства, то, можливо, різні фактори сприятимуть тому, що вони не зможуть повернути кошти банку.

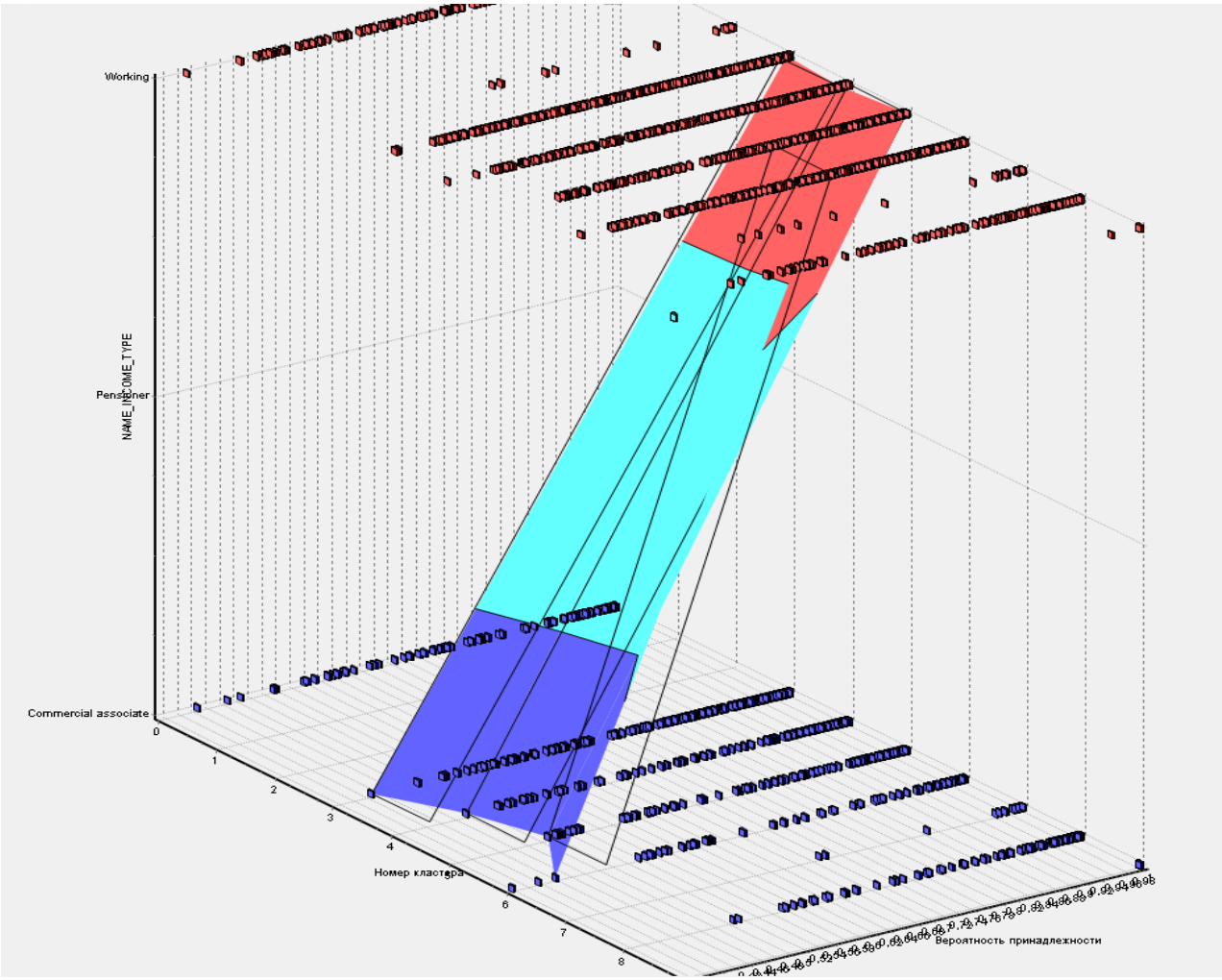


Рисунок 2.118 – Візуалізація характеристики “NAME_INCOME_TYPE”

Наступною характеристикою є “WEEKDAY_APPR_PROCESS_START”, для якої візуалізація наповненості кластерів клієнтів представлена на рисунку 2.119. Встановлено, що переважна кількість клієнтів звертається за кредитом у вівторок, а найменша кількість відповідає вихідним дням. На нашу думку, дана ознака можливо буде мало інформативною для формування кіберпрофілю у випадку таких злочинів як кібератака або соціальна інженерія. Щодо злочинів, пов’язаних з кредитним шахрайством, дана характеристика може дозволити тільки отримати додаткову оцінку завантаженості працівників банку на обробку заявок. Тобто дане знання може сприятиме формуванню додаткових організаційних заходів для підвищення уваги щодо неправильного прийняття рішення стосовно видачі кредиту.

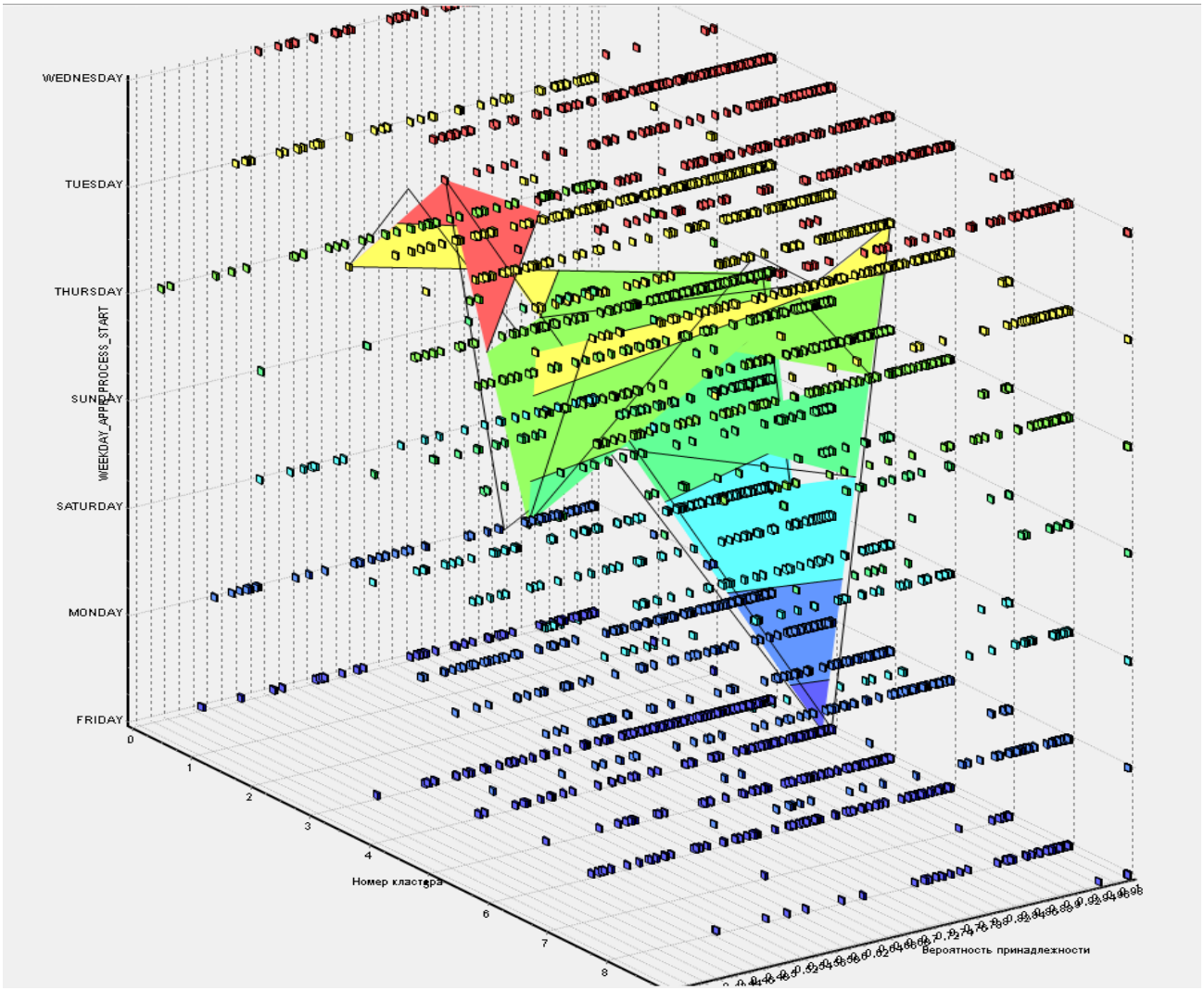


Рисунок 2.119 – Візуалізація характеристики “WEEKDAY_APPR_PROCESS_START”

В додатках представлені інші характеристики, які можна використовувати для формування та аналізу профілів кіберзлочинців. Для різних ситуацій та видів злочинів дані характеристики можуть бути замінені на ті, які є актуальними для даного випадку. Але запропонований в даній роботі підхід до профілювання кіберзлочинів є також ефективний у боротьбі з кіберзлочинністю для громадян, підприємств та державних установ поряд із програмним забезпеченням, математичними та технічними інструментами. У їх розвитку вкрай важливо враховувати різні характеристики, такі як поведінкові, географічні, психологічні та соціальні. Для їх формування профілів важливо застосовувати криміналістичні, слідчі, клінічні та статистичні методи, які сприятимуть виявленню кіберзагроз за різними аспектами їх виникнення.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**, **Ошибка! Источник ссылки не найден.**, **Ошибка! Источник ссылки не найден.**].

2.3.3 Алгоритми розпізнавання поведінки кібершахраїв

Стрімке зростання кібершахрайств за останнє десятиліття пов’язане з різними чинниками, такими як масова цифровізація різних процесів у суспільстві, автоматизація бізнес-процесів суб’єктів господарювання, створення та впровадження розумних міст, комп’ютеризація державних структур тощо. Усе це створило сприятливі умови для розвитку та використання нових доступних комп’ютерних технологій, які використовуються злочинцями для викрадення приватної інформації, отримання доступу до акаунтів інших людей, зміни даних, пошкодження комп’ютерних мереж та виведення з ладу інформаційних систем. Хоча проблема боротьби з кіберзлочинністю не є головною для світу, її вирішення вкрай необхідно для суспільства.

Одним із інструментів протидії кібершахрайству є математичні методи та моделі, які використовуються для розробки алгоритмів розпізнавання поведінки кіберзлочинців. Існує безліч підходів, які можна застосувати для вирішення тієї чи іншої задачі протидії та виявлення шахрайств. Найбільш узагальненими є

статистичні методи. Їх принциповими особливостями для дослідження даних є усереднення характеристик вибірки. При дослідженні реальних процесів, наприклад, в банківській справі, зазначені характеристики є фіктивними величинами. Також статистика вивчає масові явища, тобто не поодинокі випадки, а сукупність фактів. При цьому кількісну розмірність показників неможливо досліджувати без їх якісної визначеності. Головним завданням статистичного дослідження є виявлення закономірностей, залежностей (взаємозалежностей) між явищами, перевірка гіпотези, розбиття матриці даних на підмножини (класи), аналіз залежності однієї величини від іншої, тощо.

Вивчаючи більш детально метод статистичного аналізу, наприклад авторами Nuha M., Mahmud S., Sattar A., визначаються ключові фактори зростання шахрайства в секторах електронного та мобільного банкінгу. Використовуючи кореляційний аналіз, автори виявили значну залежність між недостатньою обізнаністю та ймовірністю негативних результатів від шахрайства. Результати дослідження демонструють, що 86,3 % жертв електронного або мобільного банківського шахрайства раніше не знали про цей тип злочинності, в результаті чого зловмисники використовували тактику та емоційні маніпуляції для отримання конфіденційної інформації клієнтської бази замість процесу злому на основі кодування [**Ошибка! Источник ссылки не найден.**].

Статистичні методи дозволяють зробити тільки певні припущення в ситуаціях із кіберзлочинами. Більш сучасні системи протидії кіберзагрозам використовують інтелектуальний аналіз даних (Data Mining). Він є концепцією, мета якої полягає в аналізі даних різної природи, просіюванні величезних обсягів збережених даних, що можуть бути неточними, неповними, суперечливими, різнорідними [**Ошибка! Источник ссылки не найден.**]. Також інтелектуальний аналіз даних – це процес обробки значного масиву інформації, з метою виявлення в них латентних правил і закономірностей; процес виявлення в «сирих» даних: раніше невідомих, нетривіальних, практично корисних,

доступних інтерпретації знань, необхідних для прийняття рішень [**Ошибка! Источник ссылки не найден.**].

Систематизовано підходи до методів інтелектуального аналізу даних, які використовувалися вченими у наукових працях для виявлення кібершахрайств у банках (таблиця 2.6).

Таблиця 2.6 – Підходи до методів інтелектуального аналізу даних для виявлення кібершахрайств у банках

№ з/п	П.І.Б. науковців	Методи інтелектуального аналізу
Вітчизняні дослідники та науковці		
1.	Яровенко Г.М., Сковронська А.І., Бояджян М.М.	метод дерева рішень (decision trees), нейронні мережі, логіт-регресія
2.	Syniavska O., Dekhtyar N., Deyneka O., Zhukova T., Syniavska O.	класична модель Лотки-Вольтерра з логістичним зростанням та динамічна модель Холлінга-Таннера
Зарубіжні дослідники та науковці		
3.	Lekha K. Chitra, Prakasam S.	k-means, Influenced Association Classifier, J48 Prediction tree
4.	Vinayakumar R., Alazab M., Soman K. P., Poornachandran P., Al-Nemrat A., Venkatraman S.	нейронна мережа (DNN), тип моделі глибинного навчання
5.	Nuha M., Mahmud S., Sattar A.	статистичний метод Data Mining (кореляційний аналіз)
6.	Kanimozhi V., Prem Jacob T.	метод штучного інтелекту
7.	Fedotova G. V., Gontar A. A., Titov V. A., Kurbanov A. K., Kuzmina E. V.	методи вертикального, горизонтального, фінансового, трендового, систематизації, аналогії, порівняння
8.	Akinbowale O. E., Klingelhöfer H. E., Zerihun M. F.	метод BSC (збалансування системи показників)
9.	Alshamasi S., Menai M.	cluster analysis

Із швидким розвитком інформаційних технологій з’явилися різновиди інтелектуального аналізу, такі як дерева рішень, нейронні мережі, логіт-регресія, система міркувань на основі аналогічних прикладів, метод штучного інтелекту, метод BSC (збалансування системи показників), еволюційне програмування, генетичні алгоритми, візуалізація багатовимірних даних, алгоритми обмеженого перебору, предметно-орієнтовні аналітичні системи, тощо [**Ошибка! Источник ссылки не найден.**].

Методи інтелектуального аналізу в банківській справі використовуються для виявлення фактів шахрайства з дебетовими та кредитними картками клієнтів, електронного та мобільного банкінгу; сегментація клієнтів для результативної маркетингової політики; прогнозування змін клієнтури, побудова моделей прогнозування обсягів споживання відповідних послуг, тощо. Це великий клас традиційних підходів (перевірка гіпотез, факторний аналіз, кореляційний, канонічний, регресійний аналіз, кластеризація) та сучасні методи (дерева класифікації, багатомірне шкалювання, структурне моделювання, дискримінантний, логлінійний, дисперсійний аналіз, компоненти дисперсії, побудова класифікаційних, асоціативних правил). Для реалізації методів інтелектуального аналізу використовують різні версії статистичних пакетів, такі як Statistica, SPSS, SAS, STATGRAPICS, STADIA, Python, R, Deductor, тощо.

Alshamasi S., Menai M. для виявлення порушень стилю написання в документі, виявлення позиції зміни авторів вимагає декомпозиції тексту на його авторські компоненти. Найкращий метод для поставленої задачі є групування текстового документу в стилістично однорідні кластери, елементи об’єднуються включаючи схожі за стилем фрагменти тексту. Застосований метод кластеризації авторства всередині документу відіграє важливу роль в розслідуванні кіберзлочинів у банках, судовій лінгвістиці, тощо [**Ошибка! Источник ссылки не найден.**].

Результати кластеризації часто не піддаються змістової інтерпретації, не відповідають інтуїтивним очікуванням, набутому досвіду за даною проблематикою. Кластеризуючи показники бази даних слід аналізувати позитивні та негативні аспекти кластеризації, обрати максимально прийнятні алгоритми. Швидкий розвиток комп’ютерної технології породжує нові методи обробки інформації, наука про бази даних зростає, обсяги інформації невпинно збільшуються. З цих причин сучасним статистичним методам складно адекватно опрацювати значні масиви даних.

Використання нейромереж дає можливість виявити приховані взаємозв’язки, у процесі роботи системи виділити класи за подібністю об’єктів.

Так, використання нейронних мереж дозволяє обробити вагомі масиви даних, використовувати алгоритми об’єднаних методів ієрархічної кластеризації з іншими методами. Нейронні мережі здатні знаходити розв’язки навіть за відсутності закономірностей та залежностей між перемінними, за відсутності апріорних відомостей про вибірку даних. Зазначимо, статистичний аналіз та математичні методи поступають у адекватному вирішенні відповідних задач. Нейронні мережі відзначаються потенціальною швидкодією, яка формується на основі масового паралелізму обробки масиву даних.

Особлива привабливість нейронних мереж зумовлена здатністю давати точні прогнози, точність прогнозування значно вища відносно статистичного аналізу. Також, до переваг віднесемо можливість працювати з неповними даними, здатність до навчання (експерт вільний у виборі математичної моделі, адже її побудова відбувається адаптивно під час навчання), висока точність, реалізація нелінійних відображень, здатність системи до адаптації (реакція та пристосування до зміни навколишнього середовища),

Lekha K. Chitra, Prakasam S. досліджують дані про кіберзлочинність на основі методів інтелектуального аналізу даних, а саме, такі як K-Means, Influenced Association Classifier, J48 Prediction tree. Метою аналізу є розпізнавання закономірностей кіберзлочинів, щоб передбачити злочинність, передбачити злочинну діяльність і запобігти їй. Науково-методичний підхід алгоритму K-Means вибирає початкові центроїди, щоб класифікатор міг отримувати дані та формулювати прогнози кіберзлочинів за допомогою алгоритму J48. Об’єднання методів безсумнівно дасть покращений, об’єднаний і точний результат показників кіберзлочинності в банківському секторі, щоб подолати та запобігти кібератаці та прогнозування неплатежів, виявлення підроблених транзакцій тощо [**Ошибка! Источник ссылки не найден.**].

Vinayakumar R., Alazab M., Soman K. P., Poornachandran P., Al-Nemrat A., Venkatraman S. досліджують глибоку нейронну мережу (DNN), тип моделі глибинного навчання, для розробки гнучкої та ефективної IDS, що виявляє та класифікує непередбачувані кібератаки. Даний тип дослідження визначає

найкращий алгоритм, який адекватно виявляє майбутні кібератаки. Комплексна оцінка експериментів DNN та різних класичних класифікаторів машинного навчання сповіщає на різних загальнодоступних наборах даних про зловмисне програмне забезпечення [**Ошибка! Источник ссылки не найден.**].

Kanimozhi V., Prem Jacob T. запропонований методичний підхід має виявити класифікацію ботнет-атаки, яка становить серйозну загрозу для фінансового сектора та банківських послуг. Використаний метод штучного інтелекту відіграє важливу роль у виявленні кібератак, системи виявлення вторгнень (IDS). Дослідження проводиться до реалістичного набору даних виявлення вторгнень кіберзахисту (CSE-CIC-IDS2018), створеного в 2018 році Канадським інститутом кібербезпеки (CIC) на AWS (веб-сервіси Amazon). Оцінка точності 99,97%, коефіцієнт помилкових позитивних результатів 0,001, тобто використаний метод штучного інтелекту виявлення атак ботнету є значущим, може бути запропонований для аналізу мережевого трафіку в реальному часі [**Ошибка! Источник ссылки не найден.**].

Syniavska O., Dekhtyar N., Deyneka O., Zhukova T., Syniavska O. запропонована математична модель присвячена питанню протидії кібератакам у сфері електронного банкінгу. В основі моделювання закладена класична модель Лотки-Вольтерра з логістичним зростанням та динамічна модель Холлінга-Таннера. На основі теорії біфуркації виділено типи фіксованих точок: сідло, стабільний вузол, стабільний вироджений вузол та лінії стабільних фіксованих точок, що малоймовірно зустрічаються в реальному житті. Розглянуту модель можна використовувати для теоретичних та емпіричних досліджень протидії кібератакам банківського сектору [**Ошибка! Источник ссылки не найден.**].

Fedotova G. V., Gontar A. A., Titov V. A., Kurbanov A. K., Kuzmina E. V. використовують методи вертикального, горизонтального, фінансового та трендового Data Mining масиву даних для оцінки динаміки та тренду розвитку кіберзлочинності в банківській сфері [**Ошибка! Источник ссылки не найден.**]. Проведено аналіз ситуації з кіберзлочинністю банківської системи, розглянуті механізми забезпечення безпеки особистих рахунків, розробка таргетів

впровадження інформаційної безпеки платіжних систем банківської сфери. Методи інтелектуального аналізу даних, а саме, систематизації, аналогії, порівняння використані авторами при формуванні висновків та рекомендацій щодо розглянутих напрямів підвищення економічної безпеки інформаційних банківських систем.

Akinbowale O. E., Klingelhöfer H. E., Zerihun M. F. використовують метод BSC (збалансування системи показників) для аналізу впливу кіберзлочинності на банківський сектор. На основі аналізу, щоб запобігти значним збиткам від кіберзлочинності, запропонована система оповіщення. Споживачами можуть бути банки, клієнти шляхом ефективного впровадження та інтеграції технології великих даних у їхню систему для пом'якшення негативного впливу. кіберзлочинності [**Ошибка! Источник ссылки не найден.**].

У банківському секторі існує серйозна проблема, пов'язана із кібершахрайствами щодо здійснення кредитних операцій. Оскільки методи інтелектуального аналізу є найбільш ефективними у боротьбі із кіберзлочинністю, то було розроблено алгоритми виявлення кіберзлочинної операції на основі побудови різних регресій, дерева рішень та нейронної мережі.

Перший алгоритм базується на побудові регресій. У випадку із шахрайствами доцільно застосувати саме логістичну регресію, але для наших даних виявилося неможливим її побудувати через незбіжність матриці. Тому було прийнято рішення щодо побудови узагальненої регресії, яка дозволить зробити прогноз не у бінарному вигляді, а в інтервалі від 0 до 1. За наявності значень, близьких до 1, вважається ознака відповідною зловживанню, в протилежному випадку, вона не вважається відповідною зловживанню. Регресію було побудовано з урахуванням також й фіктивних змінних, оскільки набір має велику кількість категоріальних змінних, які було перетворено на фіктивні. Результат узагальненої регресійної моделі представлений на рисунку 2.120. Рівень p-значущості для всіх змінних є меншим 0,05, тобто рівняння складатиметься із статистично значущих величин. Коефіцієнт детермінації дорівнює 0,775, що в принципі свідчить про непогану якість моделі.

OLS Regression Results

Dep. Variable:	y	R-squared:	0.775
Model:	OLS	Adj. R-squared:	0.775
Method:	Least Squares	F-statistic:	5335.
Date:	Thu, 08 Dec 2022	Prob (F-statistic):	0.00
Time:	20:13:28	Log-Likelihood:	4017.0
No. Observations:	193384	AIC:	-7782.
Df Residuals:	193258	BIC:	-6500.
Df Model:	125		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.3710	0.007	184.444	0.000	1.356	1.386
CNT_CHILDREN	-0.0739	0.002	-36.548	0.000	-0.078	-0.070
AMT_INCOME_TOTAL	2.234e-09	8.23e-10	2.714	0.007	6.2e-10	3.85e-09
AMT_CREDIT	1.43e-07	8.49e-09	16.835	0.000	1.26e-07	1.6e-07
AMT_ANNUITY	2.78e-07	6.29e-08	4.418	0.000	1.55e-07	4.01e-07
AMT_GOODS_PRICE	-1.731e-07	9.42e-09	-18.384	0.000	-1.92e-07	-1.55e-07
DAYS_EMPLOYED	6.704e-06	2.79e-07	24.047	0.000	6.16e-06	7.25e-06
CNT_FAM_MEMBERS	0.0617	0.002	34.713	0.000	0.058	0.065
REGION_RATING_CLIENT	-0.0231	0.003	-7.941	0.000	-0.029	-0.017
REGION_RATING_CLIENT_W_CITY	0.0249	0.003	8.452	0.000	0.019	0.031
HOUR_APPR_PROCESS_START	-0.0019	0.000	-10.478	0.000	-0.002	-0.002
NAME_CONTRACT_TYPE_Cash loans	-0.0487	0.004	-12.731	0.000	-0.056	-0.041
NAME_CONTRACT_TYPE_Revolving loans	-0.0777	0.004	-18.509	0.000	-0.086	-0.069
CODE_GENDER_F	-0.0629	0.002	-36.519	0.000	-0.066	-0.059
CODE_GENDER_M	-0.0434	0.002	-23.715	0.000	-0.047	-0.040
FLAG_OWN_CAR_N	-0.0601	0.002	-36.586	0.000	-0.063	-0.057
FLAG_OWN_CAR_Y	-0.0997	0.002	-51.133	0.000	-0.104	-0.096
FLAG_OWN_REALTY_N	-0.0872	0.002	-44.361	0.000	-0.091	-0.083
FLAG_OWN_REALTY_Y	-0.0681	0.002	-41.674	0.000	-0.071	-0.065
NAME_TYPE_SUITE_Children	-0.1091	0.008	-13.685	0.000	-0.125	-0.093
NAME_TYPE_SUITE_Family	-0.1125	0.003	-41.072	0.000	-0.118	-0.107
NAME_TYPE_SUITE_Group of people	-0.0740	0.025	-2.985	0.003	-0.123	-0.025
NAME_TYPE_SUITE_Other_A	-0.1194	0.013	-8.964	0.000	-0.145	-0.093
NAME_TYPE_SUITE_Other_B	-0.1112	0.010	-11.544	0.000	-0.130	-0.092
NAME_TYPE_SUITE_Spouse, partner	-0.1177	0.004	-27.509	0.000	-0.126	-0.109
NAME_TYPE_SUITE_Unaccompanied	-0.0684	0.002	-39.181	0.000	-0.072	-0.065
NAME_INCOME_TYPE_Businessman	-0.1427	0.084	-1.700	0.089	-0.307	0.022
NAME_INCOME_TYPE_Commercial associate	-0.1076	0.002	-53.949	0.000	-0.112	-0.104
NAME_INCOME_TYPE_Maternity leave	-0.1334	0.237	-0.563	0.574	-0.598	0.331
NAME_INCOME_TYPE_State servant	-0.0878	0.003	-25.912	0.000	-0.094	-0.081
NAME_INCOME_TYPE_Student	-0.1630	0.090	-1.817	0.069	-0.339	0.013
NAME_INCOME_TYPE_Working	-0.0690	0.002	-43.108	0.000	-0.072	-0.066
NAME_EDUCATION_TYPE_Academic degree	-0.1618	0.030	-5.481	0.000	-0.220	-0.104
NAME_EDUCATION_TYPE_Higher education	-0.1036	0.002	-47.031	0.000	-0.108	-0.099
NAME_EDUCATION_TYPE_Incomplete higher	-0.1149	0.004	-29.209	0.000	-0.123	-0.107
NAME_EDUCATION_TYPE_Lower secondary	-0.0900	0.009	-9.873	0.000	-0.108	-0.072
NAME_EDUCATION_TYPE_Secondary / secondary special	-0.0632	0.002	-38.518	0.000	-0.066	-0.060
NAME_FAMILY_STATUS_Civil marriage	-0.1603	0.003	-58.961	0.000	-0.166	-0.155
NAME_FAMILY_STATUS_Married	-0.1353	0.002	-77.191	0.000	-0.139	-0.132
NAME_FAMILY_STATUS_Separated	-0.1087	0.003	-33.060	0.000	-0.115	-0.102
NAME_FAMILY_STATUS_Single / not married	-0.0907	0.002	-36.351	0.000	-0.096	-0.086
NAME_FAMILY_STATUS_Widow	-0.1479	0.005	-30.694	0.000	-0.157	-0.138
NAME_HOUSING_TYPE_Co-op apartment	-0.0665	0.011	-5.883	0.000	-0.089	-0.044
NAME_HOUSING_TYPE_House / apartment	-0.0423	0.002	-23.078	0.000	-0.046	-0.039
NAME_HOUSING_TYPE_Municipal apartment	-0.0678	0.004	-17.497	0.000	-0.075	-0.060
NAME_HOUSING_TYPE_Office apartment	-0.0897	0.008	-10.726	0.000	-0.106	-0.073
NAME_HOUSING_TYPE_Rented apartment	-0.0568	0.007	-8.477	0.000	-0.070	-0.044
NAME_HOUSING_TYPE_With parents	-0.0604	0.004	-16.082	0.000	-0.068	-0.053
OCCUPATION_TYPE_Accountants	-0.2913	0.004	-75.091	0.000	-0.299	-0.284
OCCUPATION_TYPE_Cleaning staff	-0.2803	0.005	-53.581	0.000	-0.291	-0.270
OCCUPATION_TYPE_Cooking staff	-0.2834	0.005	-54.843	0.000	-0.294	-0.273
OCCUPATION_TYPE_Core staff	-0.2447	0.003	-79.875	0.000	-0.251	-0.239
OCCUPATION_TYPE_Drivers	-0.2386	0.003	-73.936	0.000	-0.245	-0.232
OCCUPATION_TYPE_HR staff	-0.2650	0.013	-20.519	0.000	-0.290	-0.240
OCCUPATION_TYPE_High skill tech staff	-0.2792	0.004	-78.292	0.000	-0.286	-0.272
OCCUPATION_TYPE_IT staff	-0.2872	0.013	-21.401	0.000	-0.313	-0.261
OCCUPATION_TYPE_Laborers	-0.2062	0.002	-97.241	0.000	-0.210	-0.202
OCCUPATION_TYPE_Low-skill Laborers	-0.2197	0.009	-24.926	0.000	-0.237	-0.202
OCCUPATION_TYPE_Managers	-0.2613	0.003	-89.404	0.000	-0.267	-0.256
OCCUPATION_TYPE_Medicine staff	-0.2749	0.005	-53.376	0.000	-0.285	-0.265
OCCUPATION_TYPE_Private service staff	-0.3049	0.007	-44.512	0.000	-0.318	-0.291
OCCUPATION_TYPE_Realty agents	-0.3057	0.012	-25.755	0.000	-0.329	-0.282
OCCUPATION_TYPE_Sales staff	-0.2516	0.003	-94.429	0.000	-0.257	-0.246
OCCUPATION_TYPE_Secretaries	-0.2671	0.009	-30.272	0.000	-0.284	-0.250
OCCUPATION_TYPE_Security staff	-0.2583	0.006	-46.931	0.000	-0.269	-0.247
OCCUPATION_TYPE_Waiters/barmen staff	-0.2767	0.010	-27.917	0.000	-0.296	-0.257

Рисунок 2.120 – Результати побудованої узагальненої регресії (початок)

ORGANIZATION_TYPE_Advertising	-0.3818	0.016	-23.670	0.000	-0.413	-0.350
ORGANIZATION_TYPE_Agriculture	-0.3875	0.015	-26.396	0.000	-0.416	-0.359
ORGANIZATION_TYPE_Bank	-0.3843	0.007	-56.184	0.000	-0.398	-0.371
ORGANIZATION_TYPE_Business Entity Type 1	-0.3897	0.005	-76.269	0.000	-0.400	-0.380
ORGANIZATION_TYPE_Business Entity Type 2	-0.3858	0.004	-97.820	0.000	-0.393	-0.378
ORGANIZATION_TYPE_Business Entity Type 3	-0.2964	0.002	-142.228	0.000	-0.301	-0.292
ORGANIZATION_TYPE_Cleaning	-0.3882	0.023	-16.797	0.000	-0.433	-0.343
ORGANIZATION_TYPE_Construction	-0.3707	0.005	-78.912	0.000	-0.380	-0.361
ORGANIZATION_TYPE_Culture	-0.3956	0.019	-21.133	0.000	-0.432	-0.359
ORGANIZATION_TYPE_Emergency	-0.4101	0.016	-24.973	0.000	-0.442	-0.378
ORGANIZATION_TYPE_Government	-0.3884	0.004	-87.467	0.000	-0.397	-0.380
ORGANIZATION_TYPE_Hotel	-0.3938	0.014	-28.001	0.000	-0.421	-0.366
ORGANIZATION_TYPE_Housing	-0.4130	0.007	-59.994	0.000	-0.426	-0.399
ORGANIZATION_TYPE_Industry: type 1	-0.4087	0.012	-35.176	0.000	-0.431	-0.386
ORGANIZATION_TYPE_Industry: type 10	-0.3994	0.031	-13.017	0.000	-0.460	-0.339
ORGANIZATION_TYPE_Industry: type 11	-0.4124	0.007	-57.910	0.000	-0.426	-0.398
ORGANIZATION_TYPE_Industry: type 12	-0.4215	0.018	-22.996	0.000	-0.457	-0.386
ORGANIZATION_TYPE_Industry: type 13	-0.3995	0.061	-6.522	0.000	-0.520	-0.279
ORGANIZATION_TYPE_Industry: type 2	-0.4377	0.015	-29.751	0.000	-0.467	-0.409
ORGANIZATION_TYPE_Industry: type 3	-0.4039	0.007	-55.707	0.000	-0.418	-0.390
ORGANIZATION_TYPE_Industry: type 4	-0.3951	0.013	-31.243	0.000	-0.420	-0.370
ORGANIZATION_TYPE_Industry: type 5	-0.4192	0.013	-31.296	0.000	-0.445	-0.393
ORGANIZATION_TYPE_Industry: type 6	-0.4425	0.038	-11.786	0.000	-0.516	-0.369
ORGANIZATION_TYPE_Industry: type 7	-0.4159	0.009	-43.851	0.000	-0.435	-0.397
ORGANIZATION_TYPE_Industry: type 8	-0.2662	0.075	-3.548	0.000	-0.413	-0.119
ORGANIZATION_TYPE_Industry: type 9	-0.4144	0.006	-65.382	0.000	-0.427	-0.402
ORGANIZATION_TYPE_Insurance	-0.4044	0.014	-29.124	0.000	-0.432	-0.377
ORGANIZATION_TYPE_Kindergarten	-0.3839	0.005	-78.271	0.000	-0.394	-0.374
ORGANIZATION_TYPE_Legal Services	-0.3731	0.018	-20.687	0.000	-0.408	-0.338
ORGANIZATION_TYPE_Medicine	-0.3740	0.005	-79.411	0.000	-0.383	-0.365
ORGANIZATION_TYPE_Military	-0.4046	0.008	-50.557	0.000	-0.420	-0.389
ORGANIZATION_TYPE_Mobile	-0.3518	0.018	-19.126	0.000	-0.388	-0.316
ORGANIZATION_TYPE_Other	-0.3811	0.004	-101.841	0.000	-0.388	-0.374
ORGANIZATION_TYPE_Police	-0.3974	0.008	-50.638	0.000	-0.413	-0.382
ORGANIZATION_TYPE_Postal	-0.4205	0.009	-47.057	0.000	-0.438	-0.403
ORGANIZATION_TYPE_Realtor	-0.3115	0.016	-19.238	0.000	-0.343	-0.280
ORGANIZATION_TYPE_Religion	-0.3998	0.053	-7.535	0.000	-0.504	-0.296
ORGANIZATION_TYPE_Restaurant	-0.3589	0.009	-39.170	0.000	-0.377	-0.341
ORGANIZATION_TYPE_School	-0.4010	0.005	-80.566	0.000	-0.411	-0.391
ORGANIZATION_TYPE_Security	-0.3829	0.008	-50.124	0.000	-0.398	-0.368
ORGANIZATION_TYPE_Security Ministries	-0.3957	0.009	-44.663	0.000	-0.413	-0.378
ORGANIZATION_TYPE_Self-employed	-0.3208	0.003	-125.543	0.000	-0.326	-0.316
ORGANIZATION_TYPE_Services	-0.3772	0.009	-42.413	0.000	-0.395	-0.360
ORGANIZATION_TYPE_Telecom	-0.4029	0.014	-28.669	0.000	-0.430	-0.375
ORGANIZATION_TYPE_Trade: type 1	-0.3748	0.018	-21.282	0.000	-0.409	-0.340
ORGANIZATION_TYPE_Trade: type 2	-0.3975	0.008	-51.227	0.000	-0.413	-0.382
ORGANIZATION_TYPE_Trade: type 3	-0.3609	0.006	-58.289	0.000	-0.373	-0.349
ORGANIZATION_TYPE_Trade: type 4	-0.4726	0.042	-11.258	0.000	-0.555	-0.390
ORGANIZATION_TYPE_Trade: type 5	-0.4704	0.045	-10.483	0.000	-0.558	-0.382
ORGANIZATION_TYPE_Trade: type 6	-0.3970	0.014	-28.382	0.000	-0.424	-0.370
ORGANIZATION_TYPE_Trade: type 7	-0.3725	0.004	-84.000	0.000	-0.381	-0.364
ORGANIZATION_TYPE_Transport: type 1	-0.3946	0.029	-13.679	0.000	-0.451	-0.338
ORGANIZATION_TYPE_Transport: type 2	-0.4077	0.008	-51.645	0.000	-0.423	-0.392
ORGANIZATION_TYPE_Transport: type 3	-0.3382	0.010	-32.387	0.000	-0.359	-0.318
ORGANIZATION_TYPE_Transport: type 4	-0.3816	0.005	-73.222	0.000	-0.392	-0.371
ORGANIZATION_TYPE_University	-0.3939	0.010	-40.649	0.000	-0.413	-0.375
HOUSETYPE_MODE_block of flats	-0.0534	0.004	-14.316	0.000	-0.061	-0.046
HOUSETYPE_MODE_specific housing	-0.0719	0.008	-9.044	0.000	-0.087	-0.056
HOUSETYPE_MODE_terraced house	-0.0794	0.009	-8.885	0.000	-0.097	-0.062

Omnibus:	85722.900	Durbin-Watson:	1.933
Prob(Omnibus):	0.000	Jarque-Bera (JB):	383793.968
Skew:	2.207	Prob(JB):	0.00
Kurtosis:	8.306	Cond. No.	4.42e+08

Рисунок 2.120 – Результати побудованої узагальненої регресії (продовження)

Оскільки модель регресії не є ефективною при побудові складних алгоритмів, у випадку, коли змінних дуже багато, де також не враховується їх нормальний розподіл, то було побудовано наступні моделі регресій, такі як LASSO, RIDGE та Elastic Net. Їх оцінки є менш зміщеними, що може дати кращі

результати в процесі прогнозування цільової змінної. Результати оцінок побудованих видів регресії представлені на рисунках 2.118-2.120. Значення коефіцієнту детермінації для LASSO регресії дорівнює 0,485, що говорить про не досить хорошу якість моделі. З урахуванням того, що оцінки є не такими зміщеними, як у випадку із узагальненою регресією, то не рекомендується застосовувати даний вид для алгоритму виявлення кіберзлочину.

Результат Elastic Net регресії (рис. 2.121) показує значення критерію детермінації, яке дорівнює 0,645. Тобто якість моделі є середньою і її результати можна брати до уваги в процесі протидії кіберзагроз. Найбільш ефективною виявилася RIDGE регресія, для якої коефіцієнт детермінації дорівнює 0,775 (рис. 2.123). І хоча його значення відповідає аналогічному для узагальненої регресії, у випадку для даних кібершахрайств цей вид регресії буде більш ефективним. Тому із запропонованих видів регресії обираємо RIDGE регресію.

```
[ 0.          -0.           0.          -0.          -0.          -0.
  0.00709414   0.          -0.          -0.          -0.           0.
 -0.          -0.17002758 -0.07068182 -0.07034849 -0.14391311 -0.12767708
 -0.11409848 -0.          -0.02288539 -0.          -0.          -0.
 -0.          -0.          -0.          -0.12148423 -0.          -0.04697268
 -0.          -0.06914145 -0.          -0.10276462 -0.          -0.
 -0.0052286  -0.05142318 -0.12352886 -0.03566322 -0.06989978 -0.00110099
 -0.          -0.          -0.          -0.          -0.          -0.
 -0.          -0.          -0.          -0.02688197 -0.          -0.
 -0.00278777 -0.          -0.04276375 -0.          -0.00402532 -0.
 -0.          -0.          -0.01437364 -0.          -0.          -0.
 -0.          -0.          -0.          -0.          -0.          -0.02625839
 -0.          -0.          -0.          -0.          -0.          -0.
 -0.          -0.          -0.          -0.          -0.          -0.
 -0.          -0.          -0.          -0.          -0.          -0.
 -0.          -0.          -0.          -0.          -0.          -0.
 -0.          -0.          -0.          -0.          -0.          -0.
 -0.          -0.          -0.          -0.          -0.          -0.
 -0.013564    -0.          -0.          -0.          -0.          -0.
 -0.          -0.          -0.          -0.          -0.          -0.
 -0.          -0.          -0.          -0.          -0.          -0.        ]
MSE train: 0.5145631, test: 0.5141511
R^2 train: 0.4854337, test: 0.4858315
```

Рисунок 2.121 – Результати побудованої LASSO регресії

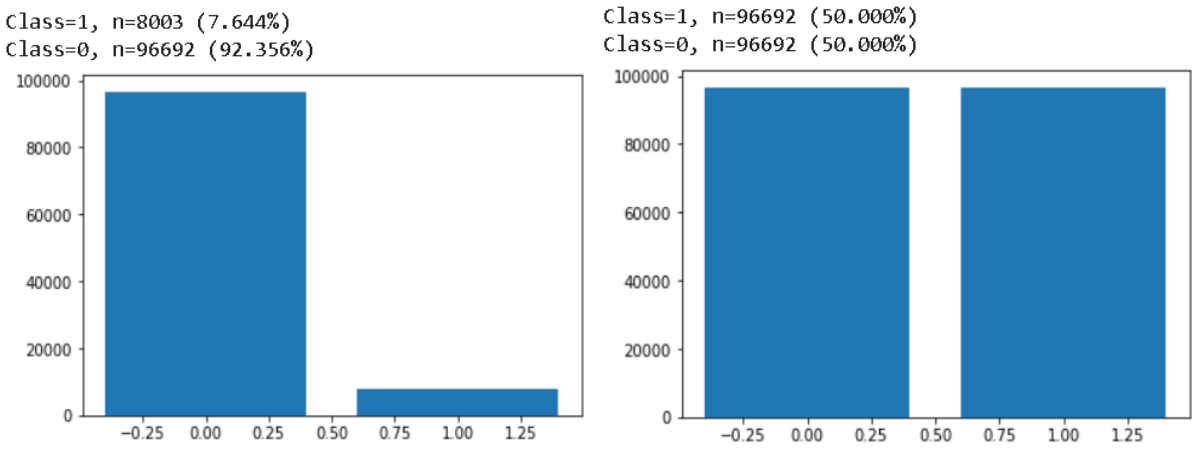
```
[ 0.         -0.01195471  0.         -0.         -0.         -0.
  0.02756828  0.         -0.         -0.         -0.          0.
 -0.         -0.14838179 -0.07969526 -0.0889707  -0.14068873 -0.12803405
 -0.1151974  -0.         -0.0554265  -0.         -0.         -0.
 -0.01131085 -0.03925029 -0.         -0.13516669 -0.         -0.06805381
 -0.         -0.09783352 -0.         -0.12103528 -0.0247224  -0.
 -0.0524418  -0.08499625 -0.15366117 -0.07023511 -0.10681343 -0.03914577
 -0.         -0.01679104 -0.00547173 -0.         -0.         -0.
 -0.03612119 -0.02410465 -0.02056971 -0.09074713 -0.0542794  -0.
 -0.05793286 -0.         -0.10552381 -0.         -0.06719768 -0.03489841
 -0.00047365 -0.         -0.07605948 -0.         -0.02778745 -0.
 -0.         -0.         -0.         -0.         -0.01205588 -0.04883392
 -0.         -0.         -0.         -0.         -0.         -0.00404568
 -0.         -0.         -0.         -0.         -0.         -0.
 -0.         -0.         -0.         -0.         -0.         -0.
 -0.         -0.         -0.         -0.         -0.         -0.
 -0.0172396  -0.         -0.         -0.02055019 -0.         -0.
 -0.         -0.         -0.         -0.         -0.         -0.
 -0.03792866 -0.         -0.         -0.         -0.         -0.
 -0.         -0.         -0.         -0.00306903 -0.         -0.
 -0.         -0.         -0.         -0.         -0.         -0.        ]
MSE train: 0.355, test: 0.355
R^2 train: 0.645, test: 0.645
```

Рисунок 2.122 – Результати побудованої Elastic Net регресії

```
[ 0.         -0.0935649   0.00296911  0.11892204  0.00761869 -0.12913517
  0.02907017  0.10521608 -0.02273391  0.02370381 -0.01219976 -0.03006902
 -0.04096166 -0.06171671 -0.03961367 -0.05895927 -0.08629575 -0.06966483
 -0.06604109 -0.01495996 -0.05761171 -0.0056295  -0.00891418 -0.01353051
 -0.03252201 -0.05715769 -0.00186482 -0.08727753 -0.00061832 -0.03567371
 -0.00140046 -0.06880264 -0.00650444 -0.08127442 -0.03521828 -0.00921614
 -0.06114914 -0.08044876 -0.1354045  -0.04329814 -0.0566609  -0.03632243
 -0.00629059 -0.03384819 -0.02165712 -0.01174937 -0.0118305  -0.0197203
 -0.0961641  -0.06422306 -0.06753818 -0.13185843 -0.10355067 -0.02263505
 -0.1039079  -0.02503478 -0.15910371 -0.0281121  -0.13233491 -0.08087302
 -0.05384816 -0.02883609 -0.14721941 -0.03208359 -0.06368    -0.03264186
 -0.02671673 -0.02959126 -0.06487991 -0.08927476 -0.12119003 -0.23233962
 -0.01987878 -0.09207939 -0.02309125 -0.04150343 -0.02694944 -0.10920965
 -0.02940132 -0.06811247 -0.03740234 -0.01335678 -0.06399749 -0.02604187
 -0.00782066 -0.0322204  -0.0620455  -0.03387124 -0.03377689 -0.01207303
 -0.04952477 -0.00639443 -0.07446286 -0.03104605 -0.09803155 -0.02329821
 -0.11949942 -0.0584315  -0.02155258 -0.12947565 -0.05920128 -0.05335531
 -0.02226548 -0.00756368 -0.04304942 -0.10174997 -0.06590341 -0.05091359
 -0.18651761 -0.05025029 -0.03154031 -0.02348703 -0.05775551 -0.06606384
 -0.01226029 -0.01163781 -0.03195566 -0.10147148 -0.01543445 -0.05851367
 -0.03635182 -0.0854006  -0.04534433 -0.01753266 -0.00932793 -0.00966927]
MSE train: 0.225, test: 0.225
R^2 train: 0.775, test: 0.775
```

Рисунок 2.123 – Результати побудованої RIDGE регресії

В якості наступного алгоритму пропонуємо застосування дерева рішень. Оскільки маємо справу із бінарною цільовою змінною, то доцільно буде побудувати класифікаційне дерево рішень. Але спочатку треба провести аналіз збалансованості змінних, що є важливим для побудови даного виду моделей. Для збалансування масиву даних була застосована техніка передискретизації синтетичної меншості. Результати незбалансованого початкового набору даних та даних після збалансування представлені на рисунку 2.124.



а) б)

Рисунок 2.124 – Дані до (a) і після (b) методу передискретизації синтетичної меншості

На рисунку 2.124а чітко видно, що дані, які відповідають випадкам, ідентифікованим як кіберзлочин, дорівнюють всього 7,64%. Тобто набір даних не є збалансованим. Застосування техніки передискретизації синтетичної меншості дозволило отримати збалансований набір даних (рис. 2.124б). Також, побудова дерев рішень на різних видах збалансованої та незбалансованої вибірок згодом підтвердило, що ефективніше даний алгоритм будувати на основі саме збалансованого масиву даних.

Для побудови дерева рішень необхідно визначити його глибину, що сприятиме формуванню такої моделі, значення якої можна інтерпретувати та використати для моделювання. Тому проведено визначення точності поділу гілок дерева рішень для незбалансованих даних за допомогою тесту Джині та ентропії. Результати представлені на рисунку 2.125.

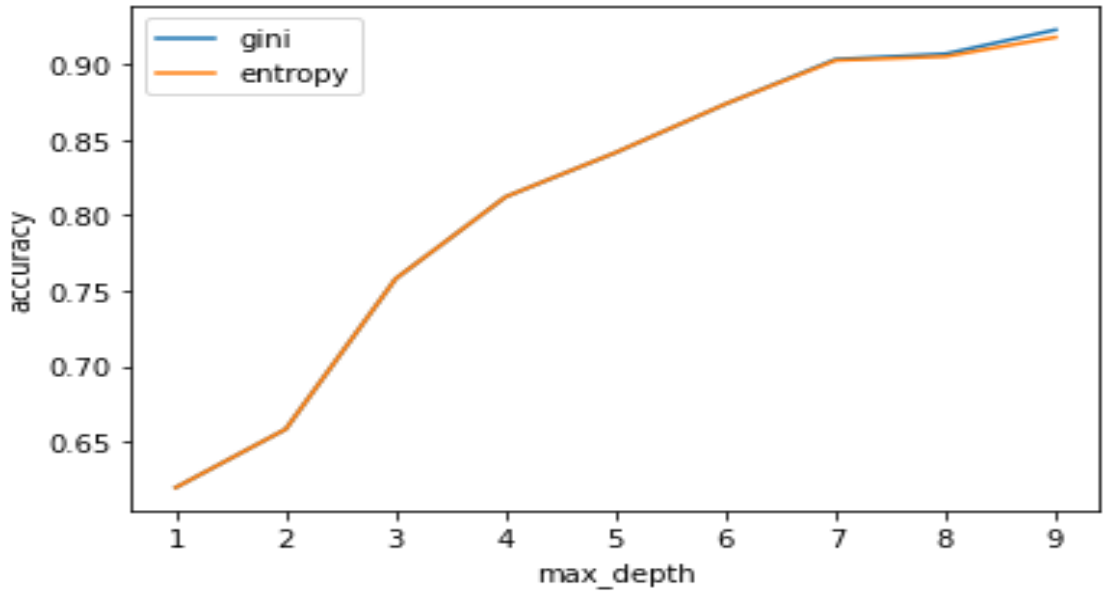


Рисунок 2.125 – Визначення точності поділу гілок дерева рішень за допомогою тесту Джині та ентропії

Дані рисунку 2.125 показують, що максимальної точності модель досягатиметься при глибині 9, але практика показує, що велике дерево рішень важко застосовувати для інтерпретації. Пожертвуємо часткою точності моделі з рахунок її спрощення. Тому обираємо глибину рівну 7 за умови точності 0,9, що є досить гарним показником. При цьому такий рівень досягається як із використанням ентропійного коефіцієнту, так і показника Джині. Для побудови дерева рішень обираємо показник ентропії. Результат його точності представлено на рисунку 2.126, а самої конфігурації моделі на рисунку 2.127.

```
Confusion Matrix:
[[19408   161]
 [ 3574 15534]]
Classification Report:
              precision    recall  f1-score   support

           0       0.84      0.99      0.91     19569
           1       0.99      0.81      0.89     19108

    accuracy                           0.90     38677
   macro avg       0.92      0.90      0.90     38677
weighted avg       0.92      0.90      0.90     38677

Accuracy: 0.9034309796519896
```

Рисунок 2.126 – Оцінки якості дерева рішень

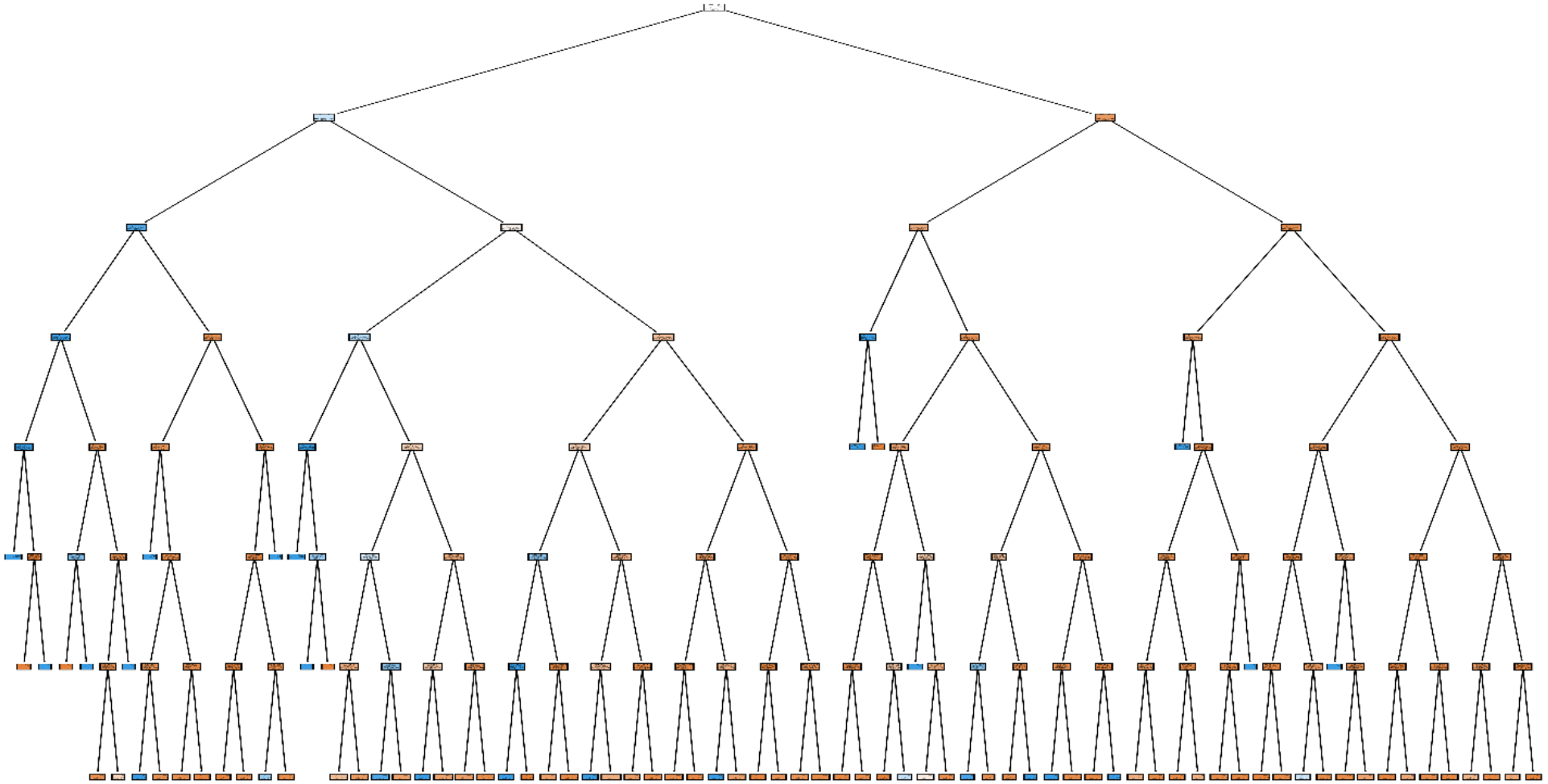


Рисунок 2.127 – Алгоритм розпізнавання поведінки кібершахраїв на основі моделі дерева прогнозних рішень

316

Загальна якість моделі для класів «0» та «1» є високою і дорівнює 0,9034. Дерево рішень дасть правильний прогноз з імовірністю 90,34%. Точність для позитивних рангів коливається від 0,84 до 0,99, що вказує на високу ймовірність того, що модель робить багато точних оптимістичних прогнозів і меншу кількість неправильних позитивних класифікацій. Параметр чутливості для всіх класів становить від 0,81 до 0,99, що підтверджує високу здатність моделі правильно визначати позитивні ранги. Оскільки ми не отримали суттєво відмінних значень точності та повторного виклику, показник F1 має високі значення, які наближаються до 1, що вказує на гарне поєднання точності та запам’ятовування моделі. Таким чином, запропонована модель є дуже якісною.

Оскільки дана модель за показником точності перевищує регресійні моделі, то можна зробити висновок, що класифікаційне дерево рішень буде більш ефективним.

Визначимо третій алгоритм, який передбачає побудову нейронної мережі. Для даного дослідження було використано Deductor Academic, оскільки воно дозволяє здійснити візуалізацію нейронної мережі, що важко виконати, застосовуючи інші аналітичні пакети.

Математичну модель нейронної мережі з урахуванням вхідних та вихідних змінних щодо кібершахрайств з кредитними операціями можна представити наступним чином (формули 2.17-2.19):

$$h_1^{(2)} = f\left(w_{1_1}^{(1)} x_1 + w_{1_2}^{(1)} x_2 + \cdots + w_{1_{126}}^{(1)} x_{126} + b_1^{(1)}\right), \quad (2.17)$$

$$h_2^{(2)} = f\left(w_{2_1}^{(1)} x_1 + w_{2_2}^{(1)} x_2 + \cdots + w_{2_{126}}^{(1)} x_{126} + b_2^{(1)}\right), \quad (2.18)$$

$$y\left(\frac{p}{1-p}\right) = f\left(w_1^{(2)} h_1^{(2)} + w_2^{(2)} h_2^{(2)}\right) \quad (2.19)$$

де $f(\cdot)$ – активаційна функція вузла, в нашому випадку сигмоїдна (логістична) функція;

$h_1^{(2)}$ – вихід першого вузла у другому шарі нейронної мережі, входами у якій є вихід першого вузла, тобто $\left(w_{1_1}^{(1)}x_1 + w_{1_2}^{(1)}x_2 + \cdots + w_{1_{126}}^{(1)}x_{126}\right)$ та вільний член для даних першого шару $b_1^{(1)}$. Ці входи складаються та передаються в активаційну функцію для розрахунку виходу першого вузла. Інший вузол $h_2^{(2)}$ формується аналогічно;

y – вихід другого вузла у третьому шарі, в якому беруться зважені виходи вузлів другого шару $h_1^{(2)}, h_2^{(2)}$. Для кінцевого виходу p відповідає цільовій змінній, що дорівнює 0, $1 - p$ – цільовій змінній, що дорівнює 1.

В якості активаційної функції для прихованих шарів та виходів застосовано сигмоїдальну (логістичну) функцію. Логістична функція для активації вихідних вузлів має вигляд (формула 2.20):

$$OUT = \frac{1}{1+\exp\left(-a \times net\right)}, \tag{2.20}$$

де OUT – виходи вузлів нейронної мережі у другому та третьому шарах, тобто $h_1^{(2)}, h_2^{(2)}$ та y;

net – сума вхідних сигналів, помножена на відповідні ваги для другого та третього шару, наприклад, $\left(w_{1_1}^{(1)}x_1 + w_{1_2}^{(1)}x_2 + \cdots + w_{1_{126}}^{(1)}x_{126} + b_1^{(1)}\right)$ для $h_1^{(1)}$ (див. формули 2.17-2.19);

a – ступінь крутизни логістичної функції.

Візуалізація отриманої нейронної мережі представлена на рисунку 2.128, де можна побачити, що на вході маємо 126 змінних та 3 шари. Третій шар відповідає прогнозованому значенню змінної, що сигналізує або випадок кіберзлочину, або його відсутності.

Представлена конфігурація нейронної мережі є на перший погляд дуже спрощеною. Перевіримо якість отриманої моделі, результати якої представлені на рисунку 2.129.

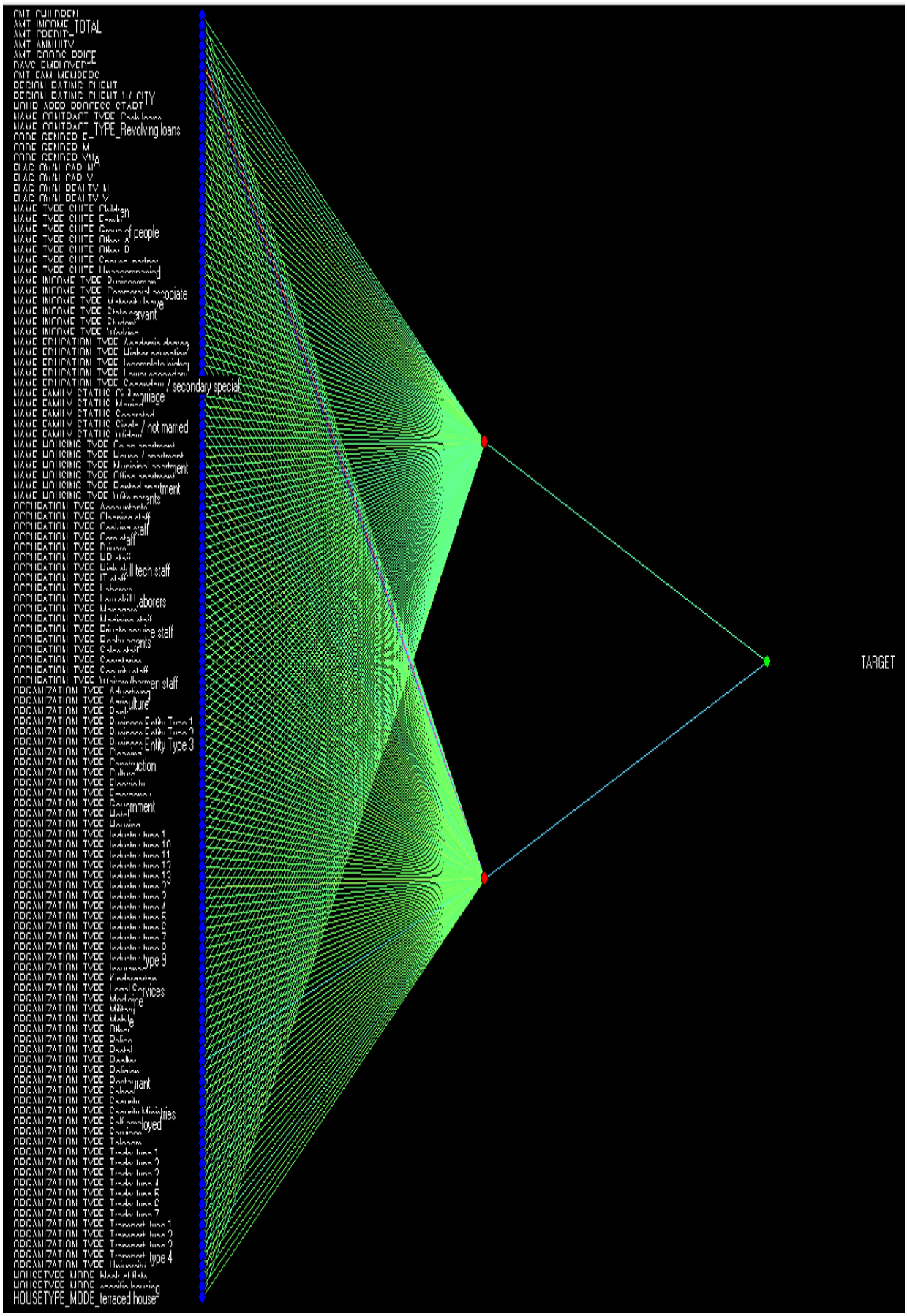


Рисунок 2.128 – Алгоритм розпізнавання поведінки кібершахраїв на основі нейронної моделі

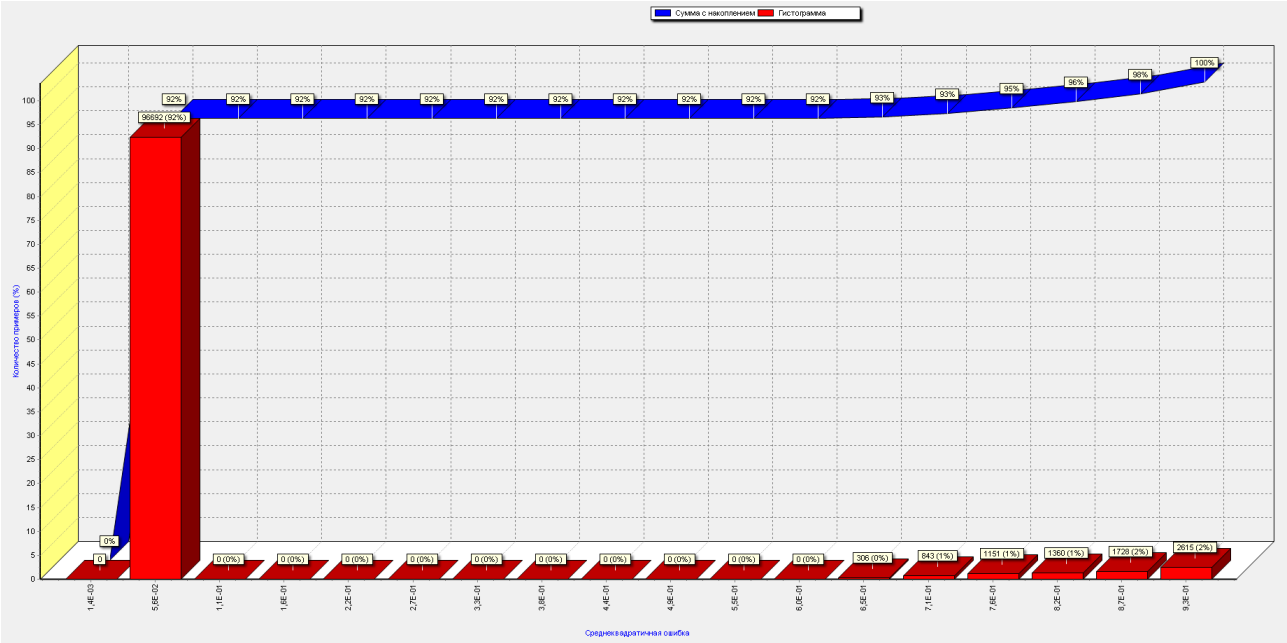


Рисунок 2.129 – Оцінка якості нейронної моделі

Графік на рисунку 2.129 показує, що середньоквадратична похибка практично дорівнює 0 для 92% прикладів. Навіть із накопиченням експериментів до 100% похибка не буде перевищувати 0,05. Отримані значення свідчать про достатньо високу якість алгоритму, побудованого на основі нейронної мережі.

Таким чином, методи виявлення та протидії кіберзагрозам є актуальними на сьогодні, особливо у контексті побудови відповідних алгоритмів розпізнавання поведінки кібершахраїв. У цьому контексті найбільш ефективними є алгоритми, засновані на математичних методах. У цьому дослідженні представлено алгоритми на основі регресійних моделей, класифікаційного дерева рішення та нейронної моделі. Отримані моделі демонструють досить непогані показники якості. Виявилося, що дерево рішень та нейронна мережа є більш точними в оцінці поведінки кібершахраїв. Саме тому пропонується їх використовувати на практиці в банківських установах для виявлення кіберзагроз на основі сформованих профілів кіберзлочинців.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден., Ошибка! Источник ссылки не найден., Ошибка! Источник ссылки не найден.**].

2.4 Методика прогнозування кібершахрайських атак на комп’ютерні системи, мережеву та хмарну інфраструктуру фінансової установи

2.4.1 Розробка моделей прогнозування кібершахрайських атак

За останні два десятиліття Четверта промислова революція призвела до стрімкого зростання інформаційних та комунікаційних технологій у світі та їх активного впровадження у різні сфері життєдіяльності суспільства. З одного боку, це сприяло і сприяє виникненню позитивних тенденцій, таких як цифрова трансформація бізнесу, розвиток сфери Інтернету речей, економіки спільного користування, віртуалізації ІТ-інфраструктури, 3Д-маркетинг, поява та використання криптовалют, блокчейнів, штучного інтелекту, ай-трекінгу, тощо. З іншого боку, комп’ютеризація та цифровізація процесів призвела до появи такого негативного явища, як кіберзлочиннність, що супроводжується паралельним зростанням цифрової грамотності населення та зниженням вартості технологій для вчинення кіберзлочинів. Так, інсталяція шкідливого програмного забезпечення на темному веб-ринку вартує 1 долар, а персональні дані будь-якої людини можна отримати всього за 3 долара [**Ошибка! Источник ссылки не найден.**]. Тобто, будь-хто може стати кіберзлочинцем або отримати доступ до будь-яких конфіденційних даних за невелику ціну.

Про актуальність проблеми кіберзлочинів та кібершахрайств свідчить й інша статистика, яка показує динамічне зростання її негативних наслідків за останні роки. Так, середня вартість витоку даних у світі від кіберінцидентів у 2022 році склала 4,35 мільйонів доларів США, що збільшилась приблизно на 24.29% у порівнянні із 2014 роком (3,5 мільйонів доларів США) [**Ошибка! Источник ссылки не найден.**]. При цьому найбільш постраждалими є такі сектори, як охорона здоров’я (10,1 мільйонів доларів США), фінанси (5,97 мільйонів доларів США), фармацевтика (5,01 мільйонів доларів США), технології (4,97 мільйонів доларів США), енергетика (4,72 мільйони доларів США), послуги (4,7 мільйонів

доларів США) та промисловість (4,47 мільйона доларів США) [**Ошибка! Источник ссылки не найден.**].

Те, що сьогодні проблема кіберзлочинності є значущою для світу, свідчать й намагання багатьох IT-компаній світу протидіяти їй шляхом створення відповідних рішень для кіберзахисту інформації та комп'ютерної інфраструктури та формування відповідно ринку. У 2022 році очікується дохід від кіберрішень та кіберпослуг у розмірі 159,84 мільярдів доларів США, що на 14,88% перевищує даний показник у 2021 році та на 91,68% у 2014 році [**Ошибка! Источник ссылки не найден.**]. При цьому прогнозується збільшення ринку кібербезпеки у 2027 році на 86,87% до 298,7 мільярдів доларів США [**Ошибка! Источник ссылки не найден.**]. За оцінками експертів зростатиме й ринок страхування від кіберінцидентів. У 2018 році його обсяг сягнув 4 мільярдів доларів США, у 2020 – 9 мільярдів доларів США, а у 2025 році його обсяг прогнозується рівним 20 мільярдів доларів США [**Ошибка! Источник ссылки не найден.**].

Боротьба із кіберзлочинністю є світовою проблемою, тому для її вирішення створено спеціальні організації, діяльність яких спрямована на формування механізму для забезпечення кібербезпеки. Серед них можна виділити Управління ООН з наркотиків і злочинності, Міжурядова група експертів відкритого складу з кіберзлочинності, Комісія із запобігання злочинності та кримінального правосуддя, Міжнародний союз електрозв'язку, Міжнародна організація кримінальної поліції, тощо. Також створюються регіональні та приватні організації, які займаються питаннями кіберзахисту.

Окрім інституційних механізмів світовими організаціями запроваджено ряд програм і ініціатив. У 2016 році країни-члени НАТО визнали кібербезпеку галуззю, якою повинен опікуватися Альянс на рівні із захистом на суші, повітрі та у морі, та прийняли оборонний мандат [**Ошибка! Источник ссылки не найден.**]. У 2021 році на саміті НАТО була запропонована та схвалена нова Комплексна політика кіберзахисту [**Ошибка! Источник ссылки не найден.**]. United Nations розробила програму «Кібербезпека та нові технології», спрямовану на розробку та

посилення заходів боротьби із кібертероризмом для країн-членів та приватних компаній [**Ошибка! Источник ссылки не найден.**].

У зв’язку із війною, яку розпочала Російська Федерація проти України, багато країн стали впроваджувати посилені заходи щодо кібербезпеки. Наприклад, The White House в Інформаційному бюлетені виклав відповідні кроки для приватних організацій для забезпечення протидії кібератакам, які можуть бути наслідками кібервійни [**Ошибка! Источник ссылки не найден.**]. National Cyber Security Centre підготував та опублікував нові вказівки, спрямовані на підтримку стійкості персоналу, якого повинні дотримуватися компанії в умовах кіберзагроз, ініційованих військовою агресією [**Ошибка! Источник ссылки не найден.**].

Таким чином, проблема боротьби із кіберзлочинністю є актуальною і інтерес до неї з часом тільки зростає, особливо в умовах світових пандемій та військових агресій. В даних умовах важливо приділяти увагу різним напрямкам її вирішення – інституційному, правовому, організаційному, методичному, тощо, що потребує системного підходу до їх дослідження і реалізації, як на практичному, так й на науковому рівнях.

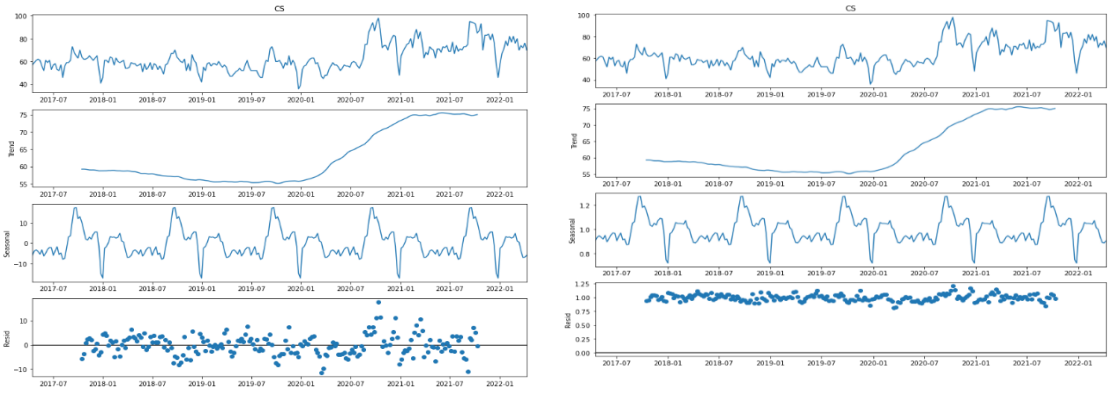
Можливості протидіяти кіберзлочинам передбачають необхідність прогнозувати потенційні кібератаки або кібершахрайства. Для вирішення даної проблеми найбільш ефективним є застосування математичних методів і моделей, які пропонуються науковцями різних світових дослідницьких шкіл. Серед них можна виділити традиційні економетричні методи дослідження, такі як регресійний аналіз [**Ошибка! Источник ссылки не найден.**], методи структурних рівнянь [**Ошибка! Источник ссылки не найден.**], VAR та VEC моделювання [**Ошибка! Источник ссылки не найден.**]. У дослідженнях набули популярності також методи нечітких множин [**Ошибка! Источник ссылки не найден.**], гравітаційне моделювання [**Ошибка! Источник ссылки не найден.**], Data mining [**Ошибка! Источник ссылки не найден.**], машинне навчання [**Ошибка! Источник ссылки не найден.**], штучний інтелект [**Ошибка! Источник ссылки не найден.**]. У даному досліджені в якості вхідних даних будуть використані інформаційні тренди найбільш популярних видів

кіберзлочинів. Оскільки їх значення представлятимуть собою часові ряди, то для їх прогнозування доцільно використовувати саме економетричні методи, які є простими у реалізації та дають точні результати на коротко- та середньострокову перспективи.

Для дослідження і прогнозування тенденцій кіберзлочинів сформовано набір вхідних даних на основі запитів інструментарію Google Trends. Сюди увійшли найбільш популярні звернення Інтернет-користувачів до термінів “Кібератаки на комп’ютерні системи фінансової установи” (CS), “Кібератаки на мережеву інфраструктуру фінансової установи” (NI), “Кібератаки на хмарну інфраструктуру фінансової установи” (CI) за період з 16.04.2017 по 10.04.2022 в розрізі потижневих рівнів.

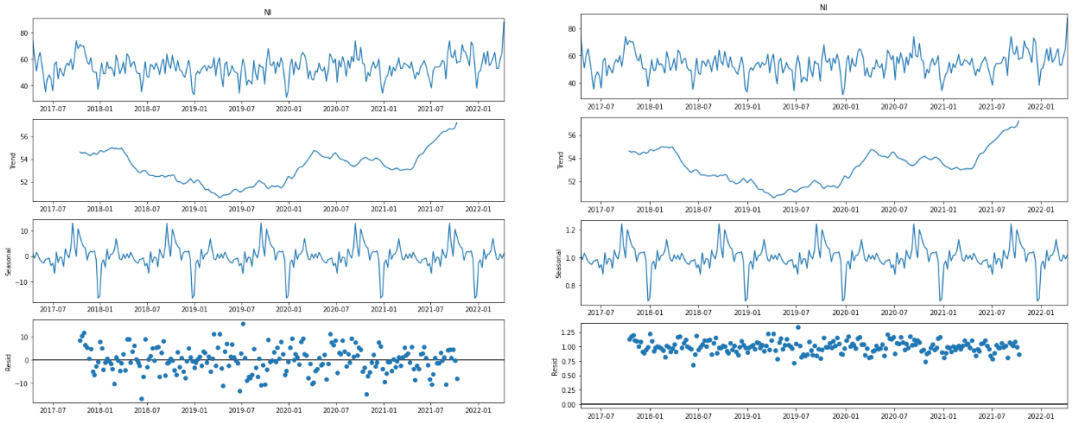
Цю інформацію було обрано, виходячи з наступних міркувань. Масові кібератаки, як правило, здійснюються по відношенню до суб’єктів економіки певної країни або країн. Віддзеркаленням цих подій є зростання зацікавленості Інтернет-користувачів у мережі щодо даних подій. Часовий розрив між реальним кіберзлочином та активністю в Інтернеті не може бути досить великим, оскільки реакція користувачів на значущі події у країні та світі є миттєвою. Офіційні джерела, які займаються збором, обробкою та публікацією статистичних даних, як правило, публікують її із значною затримкою у часі та в агрегованому вигляді. Тому у даному випадку інформаційні тренди, що відображають звернення Інтернет-користувачів, є швидким відкликом реальних подій. Відповідно, їх дослідження дозволить досить точно прогнозувати можливі кіберзлочини у світі.

Декомпозиція досліджуваних часових трендів запитів користувачів глобальної мережі з урахуванням сезонної, трендової та випадкової компонент для адитивної та мультиплікативної моделей представлена на рисунках 2.130-2.131.



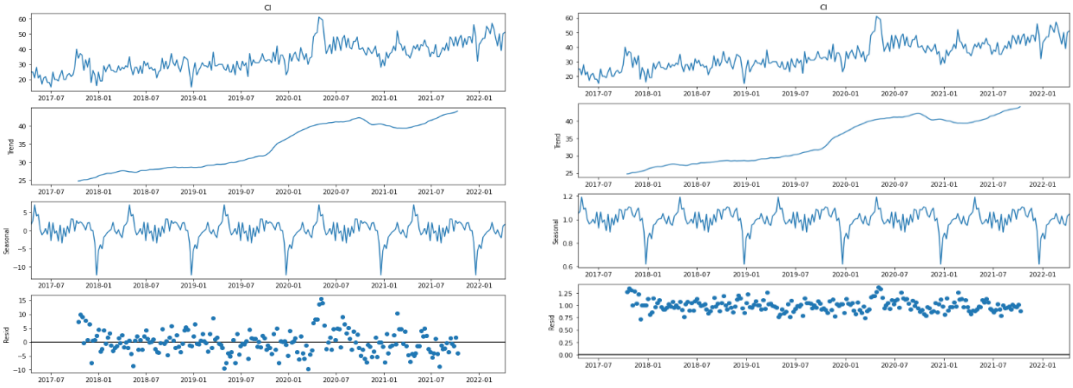
(**a**) (**b**)

Рисунок 2.130 – Декомпозиція (фактичні дані, трендова, сезонна та випадкова компоненти) часового ряду “Кібератаки на комп’ютерні системи фінансової установи”: (a) Адитивна модель; (b) Мультиплікативна модель



(**a**) (**b**)

Рисунок 2.131 – Декомпозиція (фактичні дані, трендова, сезонна та випадкова компоненти) часового ряду “Кібератаки на мережеву інфраструктуру фінансової установи”: (a) Адитивна модель; (b) Мультиплікативна модель



(**a**) (**b**)

Рисунок 2.132 – Декомпозиція (фактичні дані, трендова, сезонна та випадкова компоненти) часового ряду “Кібератаки на хмарну інфраструктуру фінансової установи”: (a) Адитивна модель; (b) Мультиплікативна модель

Аналіз декомпозиції ряду “Кібератаки на комп’ютерні системи фінансової установи” (рис. 2.130a-b) свідчить, що наявність трендової складової оцінити візуально досить складно, тому є необхідність застосування тестів для перевірки ряду на стаціонарність. Також потребує перевірці наявність викидів. Ряд містить сезонну компоненту та щільність розподілу залишків показує, що модель є адитивною. Аналіз декомпозиції “Кібератаки на мережеву інфраструктуру фінансової установи” (рис. 2.131a-b) показує наявність сезонної компоненти, а щільність розподілу залишків свідчить на користь адитивного процесу. Щодо трендової компоненти, то візуальний аналіз не дозволяє зробити висновок про її відсутність чи наявність. На рисунках 2.132a-b представлена декомпозиція ряду “Кібератаки на хмарну інфраструктуру фінансової установи”, яка демонструє чітку наявність тренду, сезонної складової та відповідність адитивному процесу.

Тобто, досліджувані дані представляють собою часові ряди, моделювання яких можливе за допомогою моделей експоненційного згладжування або авторегресійних моделей в залежності від доведення стаціонарності чи нестаціонарності процесу.

Дане прогнозування інформаційних трендів показників кіберзлочинів передбачається здійснення наступних етапів.

Етап 1. Здійснення перевірки часових рядів на наявність та відсутність аномальних значень та проведення їх відповідного коректування. Для реалізації даного етапу буде застосовано статистичний метод Z-score. Z-score вимірює відстань між значенням спостереження та середнім значенням ряду за допомогою стандартних відхилень і розраховується за формулою (2.21):

$$z = (x - \mu)/\sigma\,, \qquad (2.21)$$

де x – фактичне значення спостереження;

μ – середнє значення ряду;

σ – середньоквадратичне відхилення.

Розраховані значення Z-score порівнюють з пороговими (–3 та +3). Якщо одно із значень є більшим за +3 або меншим за –3, то дане спостереження є викидом.

Етап 2. Перевірка компонент сезонності для часових рядів CS, NI, CI, виконавши тест QS.

Сезонна стійкість виникає, коли процес є майже періодичним протягом сезону. У цьому випадку ми можемо думати, що середній рівень часового ряду x_t моделюється як:

$$x_t = S_t + w_t, \tag{2.22}$$

де S_t це сезонна складова, яка дещо змінюється від року до року відповідно до випадкового блукання:

$$S_t = S_{t-12} + v_t, \tag{2.23}$$

де w_t та v_t є некорельованими процесами білого шуму.

Для перевірки наявності компоненти сезонності в часових рядах CS, NI, CI пропонується використовувати тест QS та його застосування на базі пакету R. Ідея оцінки тесту QS базується на співвідношенні:

$$QS = n \cdot (n+2) \cdot \left(\frac{R_s^2}{n-s} + \frac{R_{2s}^2}{n-2s}\right), \tag{2.24}$$

де n – кількість спостережень у часовому ряду та s – періодичність даних (12 у цьому випадку з місячними даними);

R_s^2та R_{2s}^2 позначають автокореляції, отримані для відповідного часового ряду. Ця статистика приблизно відповідає розподілу χ2 з 2 ступенями свободи.

Для виконання тесту QS на сезонність у часовому ряді використовується функція:

$$qs(x, freq = NA, diff = T, residuals = F, autoarima = T), \quad (2.25)$$

де x - часові ряди;

freq - періодичність часового ряду;

diff – різницевий ряд;

residuals - залишки моделі;

autoarima - автоматичний.

Етап 3. Здійснення перевірки стаціонарності ряду шляхом застосування методу різниць середніх рівнів. Даний тест перевіряє гіпотезу про однорідність дисперсій частин часового ряду та гіпотезу про відсутність тренду. Застосування даного тесту є обґрунтованим для вхідних даних, оскільки на графіках трендів можна побачити, що дані не є однорідними протягом всього періоду часу та мають перегин. Для його реалізації ряд необхідно розділити на дві частини з приблизно однаковою кількістю точок та обчислити їх дисперсію (2.26):

$$\sigma_1^2 = \frac{\sum_{t=1}^{n_1}(Y_{t_1} - \overline{Y_1})^2}{n_1 - 1}; \sigma_2^2 = \frac{\sum_{t=1}^{n_2}(Y_{t_2} - \overline{Y_2})^2}{n_2 - 1}, \quad (2.26)$$

де σ_1^2, σ_2^2 – дисперсії двох частин часового ряду;

Y_{t_1}, Y_{t_2} – фактичні значення двох частин часового ряду;

$\overline{Y_1}$, $\overline{Y_2}$ – середнє значення двох частин часового ряду;

n_1, n_2 – кількість спостережень в 1-й та 2-й частинах часового ряду.

Перевірка гіпотези на однорідність ряду здійснюється за допомогою критерія Фішера (2.27):

$$F = \begin{cases} \sigma_1^2/\sigma_2^2\,,\ \sigma_1^2 > \sigma_2^2 \\ \sigma_2^2/\sigma_1^2\,,\ \sigma_2^2 > \sigma_1^2 \end{cases}, \tag{2.27}$$

де F – розраховане значення критерію Фішера. Якщо його значення менше табличного, визначеного для рівня значущості 0.05 та $(n_1 - 1)$, $(n_2 - 1)$ – ступенів вільності, то приймається гіпотеза про однорідність дисперсій, в протилежному випадку метод не дає відповіді на запитання про наявність чи відсутність тренду.

Перевірка гіпотези щодо відсутності тренду проводиться за допомогою критерію Стьюдента (2.28):

$$t = \frac{|\overline{Y_1} - \overline{Y_2}|}{\sqrt{\frac{(n_1-1)\cdot\sigma_1^2 + (n_2-1)\cdot\sigma_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \tag{2.28}$$

де t – розраховане значення критерію Ст’юдента. Якщо його значення менше табличного, визначеного для рівня значущості 0.05 та $(n_1 + n_2 - 2)$ – ступенів вільності, то приймається гіпотеза щодо відсутності тренду, в протилежному випадку тренд присутній.

Етап 4. Оскільки аналізовані ряди є нестаціонарні, тому для прогнозування інформаційних тенденцій кіберзлочинів буде обрано моделі експоненційного згладжування.

Проста модель експоненційного згладжування має вигляд (2.29):

$$S_t = \alpha \cdot X_t + (1 - \alpha) \cdot S_{t-1}, \tag{2.29}$$

де S_t, S_{t-1} – експоненційно згладжене значення в момент часу t та $(t-1)$ відповідно $(t=\overline{1,n})$;

α – параметр згладжування, який приймає значення від нуля (коли ігноруються усі поточні спостереження) до одиниці (коли повністю ігноруються усі попередні спостереження);

X_t – рівень часового ряду в момент часу t.

В даній роботі буде побудовано наступні різновиди моделей експоненційного згладжування:

- адитивна модель циклічності (2.30):

$$S_t = \alpha \cdot X_t + (1-\alpha) \cdot S_{t-1} + I_{t-p}, I_t = I_{t-p} + (1-\alpha) \cdot e_t, \quad (2.30)$$

де I_t, I_{t-p} – згладжений сезонний фактор у момент часу t та $t-p$ (довжина сезону);

e_t – залишки у момент часу t;

- тренд-циклічна адитивна модель з лінійним трендом (2.31):

$$S_t = LT_t + \alpha \cdot X_t + (1-\alpha) \cdot S_{t-1} + I_{t-p}, I_t = I_{t-p} + (1-\alpha) \cdot e_t, \quad (2.31)$$

де LT_t – лінійний тренд (значення в момент часу t);

- тренд-циклічна адитивна модель з експоненційним трендом (2.32):

$$S_t = ET_t + \alpha \cdot X_t + (1-\alpha) \cdot S_{t-1} + I_{t-p}, I_t = I_{t-p} + (1-\alpha) \cdot e_t, \quad (2.32)$$

де ET_t – експоненціальний тренд (значення в момент часу t);

- тренд-циклічна адитивна модель із затухаючим трендом (2.33):

$$S_t = DT_t + \alpha \cdot X_t + (1-\alpha) \cdot S_{t-1} + I_{t-p}, I_t = I_{t-p} + (1-\alpha) \cdot e_t, \quad (2.33)$$

де DT_t – затухаючий тренд (значення в момент часу t);

- мультиплікативна модель циклічності (2.34):

$$S_t = (\alpha \cdot X_t + (1-\alpha) \cdot S_{t-1}) \cdot I_{t-p}, I_t = I_{t-p} + \delta \cdot (1-\alpha) \cdot e_t / S_t, \quad (2.34)$$

де δ – сезонний параметр параметром згладжування, який зазначається лише для сезонних моделей;

- мультиплікативна тренд-циклічна модель з лінійним трендом (2.35):

$$S_t = LT_t \cdot (\alpha \cdot X_t + (1-\alpha) \cdot S_{t-1}) \cdot I_{t-p}, I_t = I_{t-p} + \delta \cdot (1-\alpha), \quad (2.35)$$

- мультиплікативна тренд-циклічна модель з експоненційним трендом (2.36):

$$S_t = ET_t \cdot (\alpha \cdot X_t + (1-\alpha) \cdot S_{t-1}) \cdot I_{t-p}, I_t = I_{t-p} + \delta \cdot (1-\alpha), \quad (2.36)$$

- мультиплікативна тренд-циклічна модель із затухаючим трендом (2.37):

$$S_t = DT_t \cdot (\alpha \cdot X_t + (1-\alpha) \cdot S_{t-1}) \cdot I_{t-p}, I_t = I_{t-p} + \delta \cdot (1-\alpha). \quad (2.37)$$

Хоча в результаті візуального аналізу початкових даних було доведено, що вони слідують адитивному процесу, буде побудовано також й мультиплікативні моделі експоненційного згладжування для математичного обґрунтування отриманих висновків.

Етап 5. На останньому етапі даного дослідження буде здійснено оцінку точності прогнозів показників «Кібератаки на комп'ютерні системи фінансової установи», «Кібератаки на мережеву інфраструктуру фінансової установи», «Кібератаки на хмарну інфраструктуру фінансової установи», розрахованих за побудованими моделями експоненційного згладжування. Для цього будуть розраховані помилки: «Mean Error», «Mean Absolute Error», «Sums of Squares», «Mean Square», «Mean Percentage Error» та «Mean Absolute Percentage Error».

На першому етапі запропонованої методології прогнозування інформаційних трендів кіберзлочинів було проаналізовано часові ряди на наявність аномальних значень. Для реалізації статистичного методу Z-score було використано мову програмування Python. В результаті для ряду показника «Кібератаки на комп'ютерні системи фінансової установи» виявлено одне аномальне значення, для «Кібератаки на мережеву інфраструктуру фінансової установи» – 5, для «Кібератаки на хмарну інфраструктуру фінансової установи» – 3. Виявлені значення були замінені на середньоарифметичні, узяті для спостережень, що є попереднім та наступним до аномального.

На другому етапі було застосовано QS тест, який було здійснено із використанням мови програмування R. В результаті встановлено, що значення циклічної компоненти для трьох рядів динаміки дорівнює 48, що також підтверджується візуалізацією сезонної складової на рисунках 2.130-2.132.

На третьому етапі побудовано автокореляційні функції часових рядів для проведення їх візуального аналізу на стаціонарність. Результати представлені на рисунках 2.133-2.135:

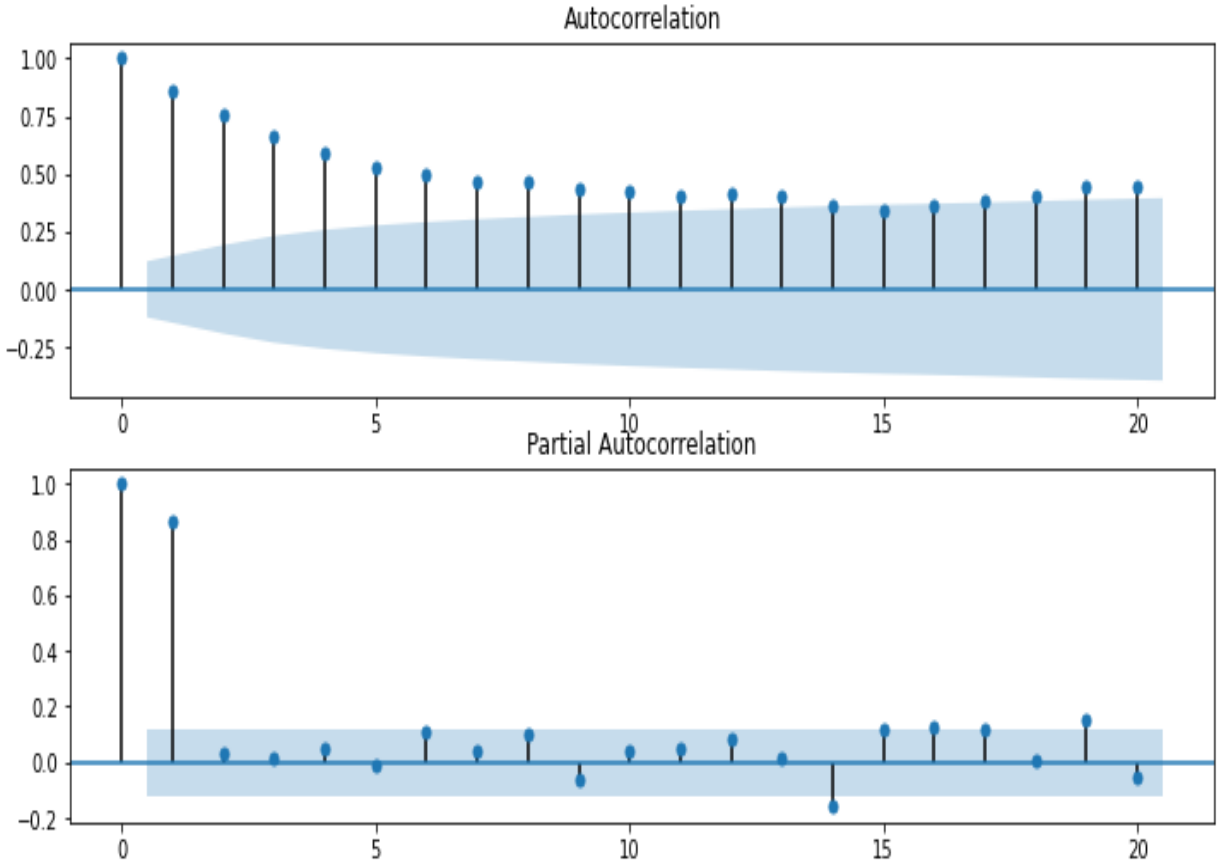


Рисунок 2.133 – Графіки автокореляційної функції та функції часткової автокореляції для показника «Кібератаки на комп’ютерні системи фінансової установи»

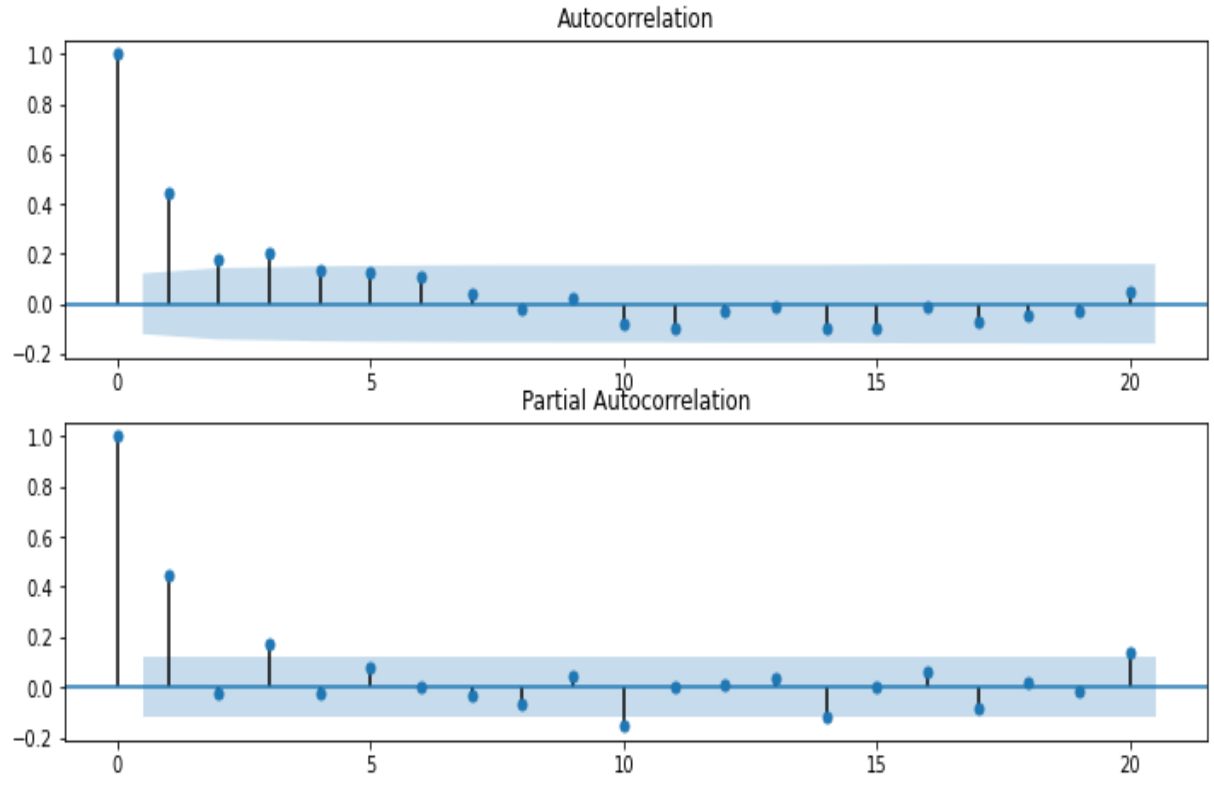


Рисунок 2.134 – Графіки автокореляційної функції та функції часткової автокореляції для показника «Кібератаки на мережеву інфраструктуру фінансової установи»

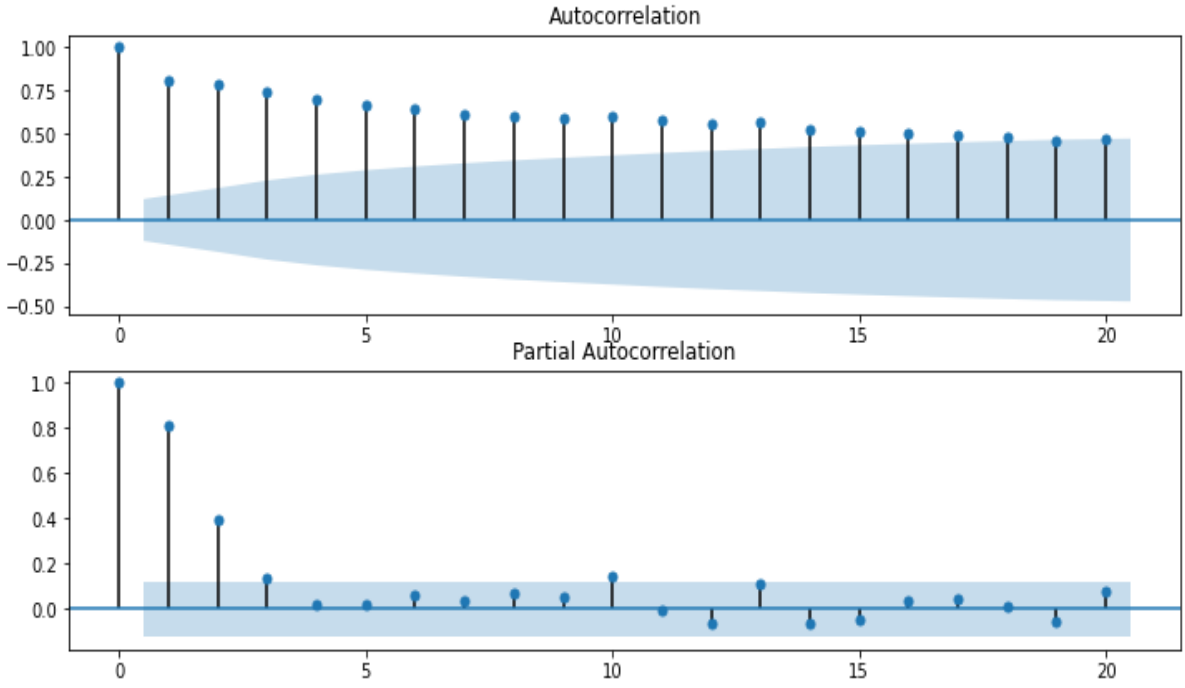


Рисунок 2.135 – Графіки автокореляційної функції та функції часткової автокореляції для показника «Кібератаки на хмарну інфраструктуру фінансової установи»

Аналізуючи отримані графіки, було зроблено попередній висновок, що ряди «Кібератаки на комп'ютерні системи фінансової установи» (рис. 2.133) та «Кібератаки на хмарну інфраструктуру фінансової установи» (рис. 2.135) є нестаціонарними, оскільки автокореляційні коефіцієнти для перших рівнів є статистично значущими. Що стосується ряду «Кібератаки на мережеву інфраструктуру фінансової установи», то однозначно не можна стверджувати, що ряд є стаціонарний чи нестаціонарний, оскільки значення автокореляційної функції для першого рівня дорівнює 0.5, що свідчить тільки про помітний рівень зв'язку і не дозволяє з впевненістю зробити висновок про стаціонарність. Тому було проведено тест різниць середніх рівнів із використанням програмного забезпечення MS Excel, результати якого представлені в таблиці 2.7.

Таблиця 2.7 – Результати проведеного тесту різниць середніх рівнів

Критерії та висновки	CS	NI	CI
F розрахований	4.6901	1.1489	1.8905
F критичний	1.3374	1.3374	1.3374
Результат перевірки гіпотези на однорідність ряду	Гіпотезу однорідності відхилено	Гіпотезу однорідності прийнято	Гіпотезу однорідності відхилено
t розрахований	9.1187	2.3558	18.5668
t критичний	1.9692	1.9692	1.9692
Результат перевірки гіпотези щодо відсутності тенденції	Тенденція є	Тенденція є	Тенденція є

Результати проведеного тесту показують, що ряди «Кібератаки на комп'ютерні системи фінансової установи» та «Кібератаки на хмарну інфраструктуру фінансової установи» є неоднорідними та містять тренд. Для ряду «Кібератаки на мережеву інфраструктуру фінансової установи» було підтверджено наявність тренду, хоча він й виявився однорідним. Таким чином, можна застосувати клас моделей експоненційного згладжування для даних дослідження.

На четвертому етапі було побудовано моделі експоненційного згладжування для прогнозування інформаційних трендів запитів щодо кібератак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи. З цією метою було використано інструментарій аналітичного пакету STATISTICA. Результати отриманих прогнозних моделей представлені у таблиці 2.8.

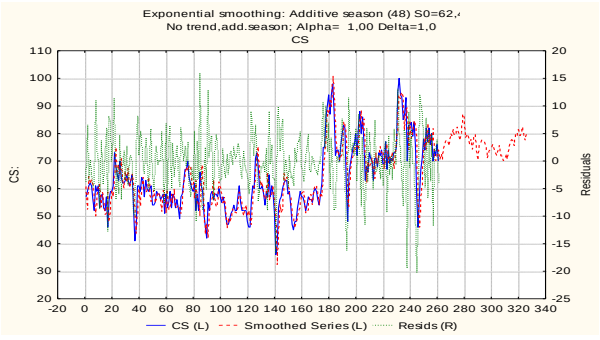
Таблиця 2.8 – Прогнозні моделі експоненціального згладжування інформаційних трендів запитів щодо кібератак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи

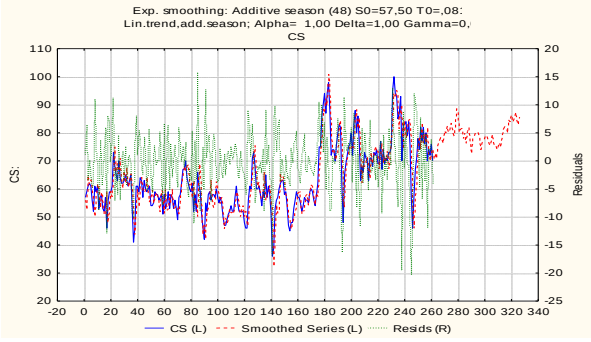
Індикатор	Модель	Властивості моделі
Кібератаки на комп'ютерні системи фінансової установи	Модель 1	Additive season (48); S0=62,42; No trend; Alpha= 1,00; Delta=1,00
	Модель 2	Additive season (48); S0=57,50; T0=0,0815; Linear trend; Alpha= 1,00; Delta=1,00; Gamma=0,00
	Модель 3	Additive season (48); S0=60,70; T0=0,9991; Exponential trend; Alpha= 1,00; Delta=1,00; Gamma=0,00
	Модель 4	Multiplicative season (48); S0=62,42; No trend; Alpha=0,883; Delta=0,125

Індикатор	Модель	Властивості моделі
	Модель 5	Multiplicative season (48); S0=57,50; T0=0,0815; Linear trend; Alpha=0,887; Delta=0,109; Gamma=0,00
	Модель 6	Multiplicative season (48) S0=60,70; T0=0,9991; Exponential trend; Alpha=0,887; Delta=0,114; Gamma=0,00
Кібератаки на мережеву інфраструктуру фінансової установи	Модель 1	Additive season (48); S0=53,90; No trend; Alpha=0,569; Delta=0,00
	Модель 2	Additive season (48); S0=56,88; T0=-0,012; Linear trend; Alpha=0,564; Delta=0,00; Gamma=0,00
	Модель 3	Additive season (48); S0=58,86; T0=0,9984; Exponential trend; Alpha=0,573; Delta=0,00; Gamma=0,00
	Модель 4	Additive season (48); S0=74,05; T0=-0,728; Damped trend; Alpha=0,361; Delta=0,00; Phi=0,017
	Модель 5	Multiplicative season (48); S0=53,90; No trend; Alpha=0,518; Delta=0,00
	Модель 6	Multiplicative season (48); S0=56,88; T0=-0,012; Linear trend; Alpha=0,518; Delta=0,00; Gamma=0,00
	Модель 7	Multiplicative season (48); S0=58,86; T0=0,9984; Exponential trend; Alpha=0,527; Delta=0,00; Gamma=0,00
	Модель 8	Multiplicative season (48); S0=71,43; T0=-0,618; Damped trend; Alpha=0,328; Delta=0,00; Phi=0,020
Кібератаки на хмарну інфраструктуру фінансової установи	Модель 1	Additive season (48); S0=33,73; No trend; Alpha=0,763; Delta=0,00
	Модель 2	Additive season (48); S0=25,94; T0=0,0667; Linear trend; Alpha=0,756; Delta=0,00; Gamma=0,00
	Модель 3	Additive season (48); S0=27,21; T0=1,000; Exponential trend; Alpha=0,761; Delta=0,00; Gamma=0,00
	Модель 4	Multiplicative season (48); S0=33,73; No trend; Alpha=1,00; Delta=1,00
	Модель 5	Multiplicative season (48); S0=25,94; T0=0,0667; Linear trend; Alpha=1,00; Delta=1,00; Gamma=0,00
	Модель 6	Multiplicative season (48); S0=27,21; T0=1,000; Exponential trend; Alpha=0,815; Delta=0,00; Gamma=0,00

Результати виявлених циклічних компонент розглянутих 3 часових рядів представлені в таблиці Е.1 Додатку Е.

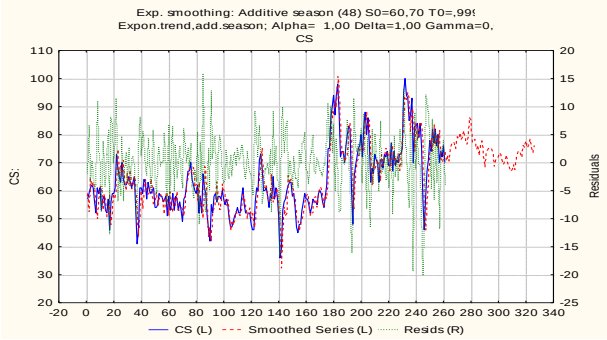
Зобразимо результати моделювання на рисунках 2.136–2.138, як співвідношення фактичних, теоретичних та прогнозних рівнів показника «Кібератаки на комп'ютерні системи фінансової установи». Прогнозні значення відображають період з 16.04.2017 по 09.07.2023 рр.

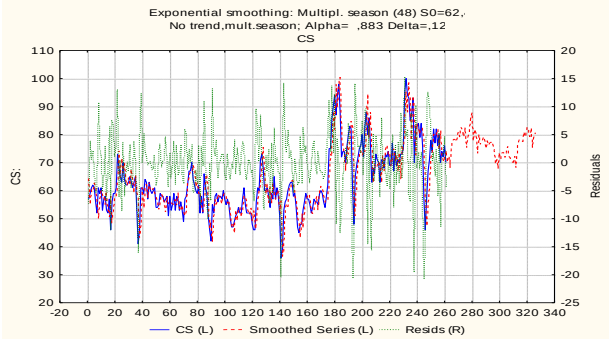




(a) (b)

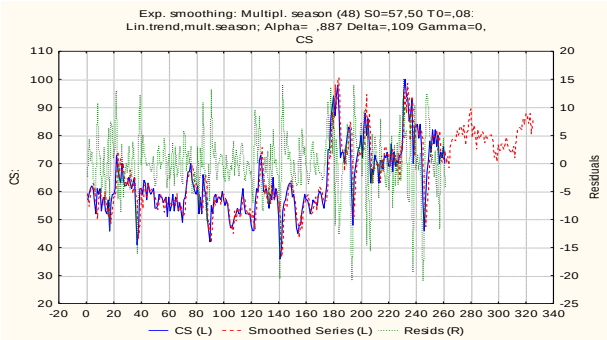
Рисунок 2.136 – Співвідношення фактичних, теоретичних та прогнозних рівнів показника «Кібератаки на комп'ютерні системи фінансової установи»: (a) Адитивна модель циклічності; (b) Тренд-циклічна адитивна модель з лінійним трендом

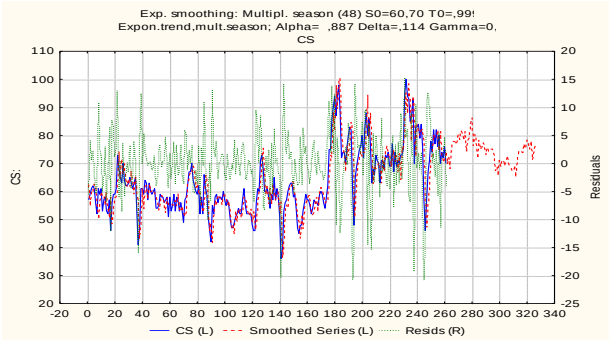




(a) (b)

Рисунок 2.137 – Співвідношення фактичних, теоретичних та прогнозних рівнів показника «Кібератаки на комп'ютерні системи фінансової установи»: (a) Тренд-циклічна адитивна модель з експоненційним трендом; (b) Мультиплікативна модель циклічності



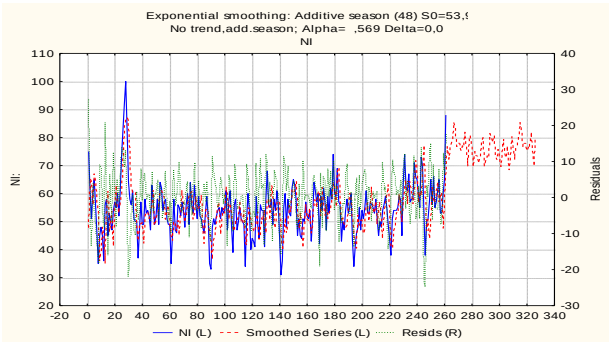


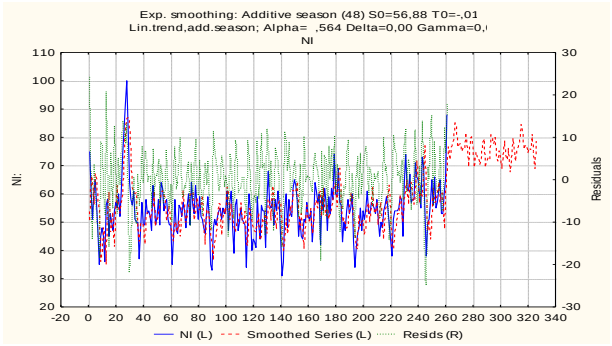
(a) (b)

Рисунок 2.138 – Співвідношення фактичних, теоретичних та прогнозних рівнів показника «Кібератаки на комп'ютерні системи фінансової установи»: (a) Мультиплікативна тренд-циклічна модель з лінійним трендом; (b) Мультиплікативна тренд-циклічна модель з експоненційним трендом

Представимо результати експоненційного моделювання на рисунках 2.139–2.142, як співвідношення фактичних, теоретичних та прогнозних рівнів показника

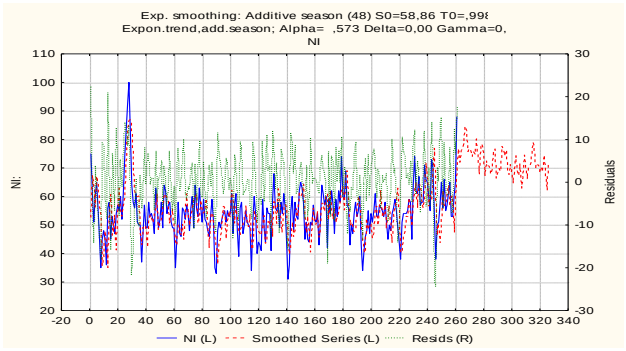
«Кібератаки на мережеву інфраструктуру фінансової установи». Прогнозні значення відображають період з 16.04.2017 по 09.07.2023 рр.

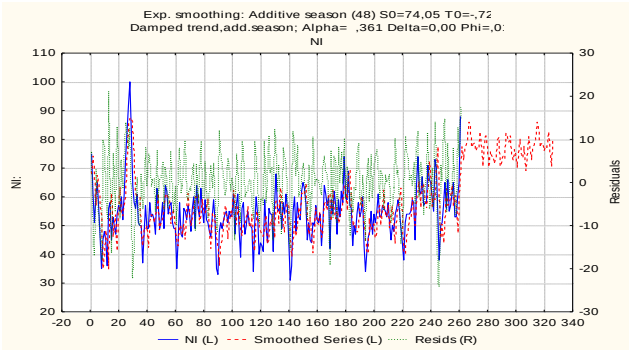




(a) (b)

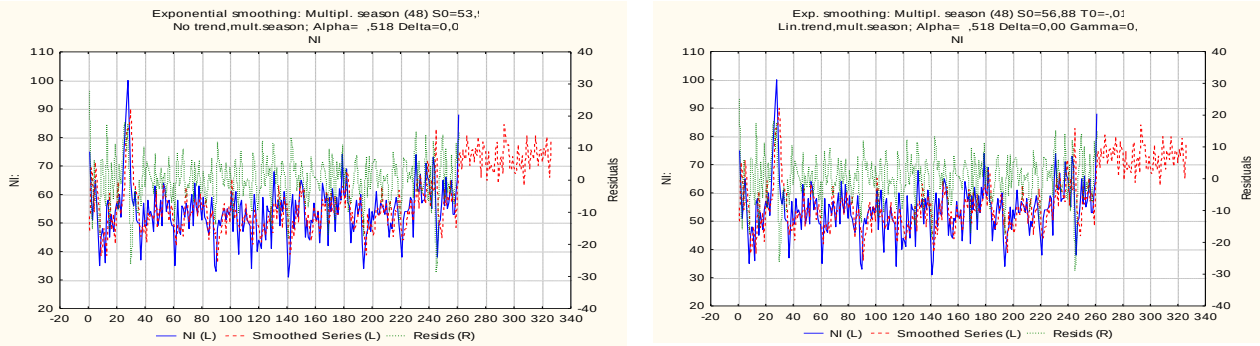
Рисунок 2.139 – Співвідношення фактичних, теоретичних та прогнозних рівнів показника «Кібератаки на мережеву інфраструктуру фінансової установи»: (a) Адитивна модель циклічності; (b) Тренд-циклічна адитивна модель з лінійним трендом





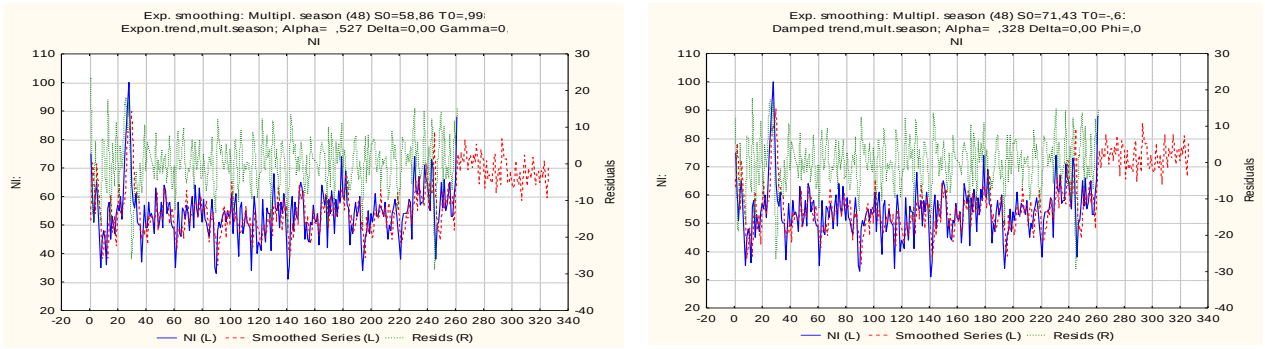
(a) (b)

Рисунок 2.140 – Співвідношення фактичних, теоретичних та прогнозних рівнів показника «Кібератаки на мережеву інфраструктуру фінансової установи»: (a) Тренд-циклічна адитивна модель з експоненційним трендом; (b) Тренд-циклічна адитивна модель з затухаючим трендом



(a) (b)

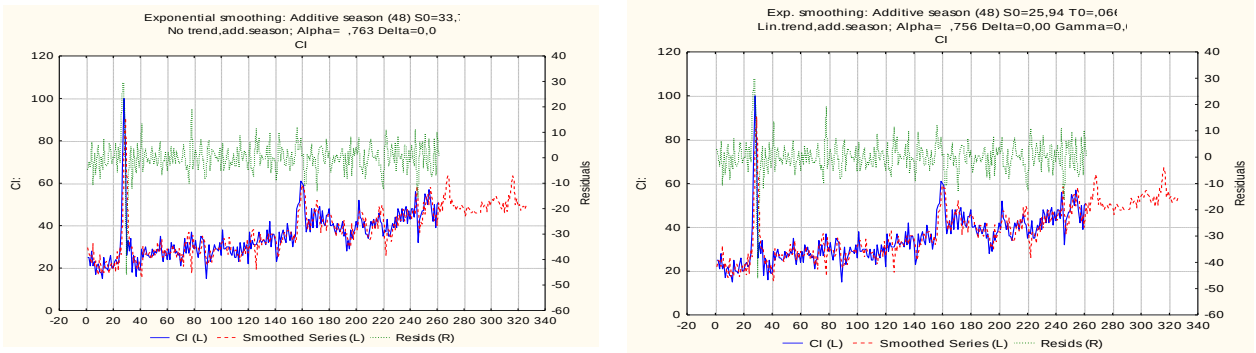
Рисунок 2.141 – Співвідношення фактичних, теоретичних та прогнозних рівнів показника «Кібератаки на мережеву інфраструктуру фінансової установи»: (a) Мультиплікативна модель циклічності; (b) Мультиплікативна тренд-циклічна модель з лінійним трендом



(a) (b)

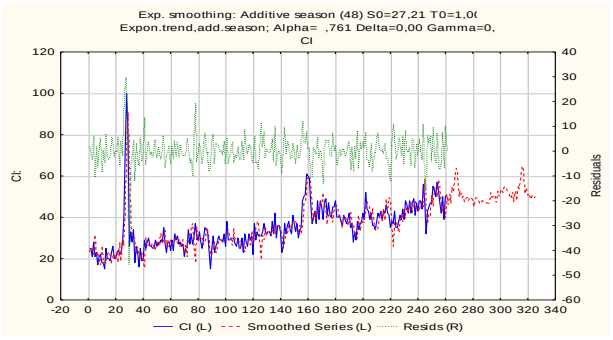
Рисунок 2.142 – Співвідношення фактичних, теоретичних та прогнозних рівнів показника «Кібератаки на мережеву інфраструктуру фінансової установи»: (a) Мультиплікативна тренд-циклічна модель з експоненційним трендом; (b) Мультиплікативна тренд-циклічна модель з затухаючим трендом

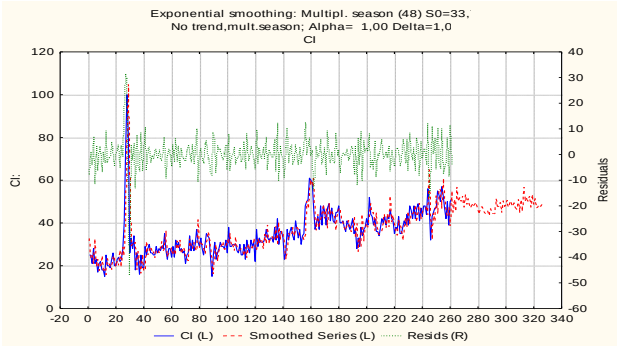
На рисунках 2.143-2.145 представлені результати моделювання співвідношення фактичних, теоретичних та прогнозних рівнів показника «Кібератаки на хмарну інфраструктуру фінансової установи». Прогнозні значення відображають період з 16.04.2017 по 09.07.2023 рр.



(a) (b)

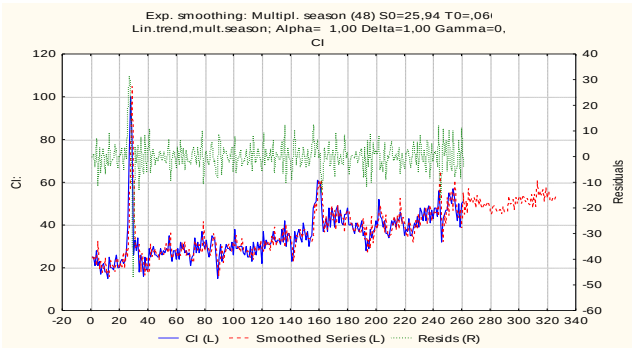
Рисунок 2.143 – Співвідношення фактичних, теоретичних та прогнозних рівнів показника «Кібератаки на хмарну інфраструктуру фінансової установи»: (a) Адитивна модель циклічності; (b) Тренд-циклічна адитивна модель з лінійним трендом

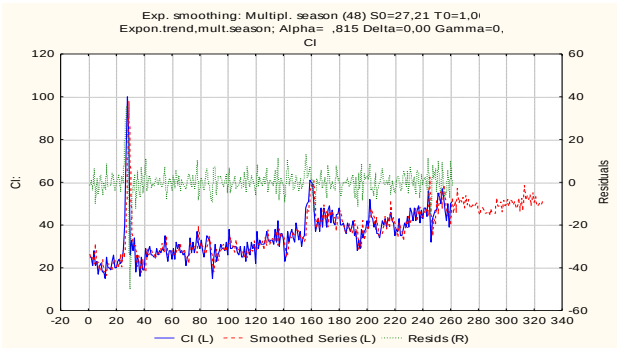




(a) (b)

Рисунок 2.144 – Співвідношення фактичних, теоретичних та прогнозних рівнів показника «Кібератаки на хмарну інфраструктуру фінансової установи»: (a) Тренд-циклічна адитивна модель з експоненційним трендом; (b) Мультиплікативна модель циклічності





(a) (b)

Рисунок 2.145 – Співвідношення фактичних, теоретичних та прогнозних рівнів показника «Кібератаки на хмарну інфраструктуру фінансової установи»: (a) Мультиплікативна тренд-циклічна модель з лінійним трендом; (b) Мультиплікативна тренд-циклічна модель з експоненційним трендом

Розраховані прогнозні значення показників кібератак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи за період з 17.04.2022 по 09.07.2023 систематизуємо у вигляді таблиці Е.2 та представимо у Додатку Е. Розроблення моделі прогнозування кібератак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи вимагало проведення перевірки точності обчислених прогнозних рівнів. Тому на п'ятому етапі проведено аналіз наступного переліку показників: «Mean Error», «Mean Absolute

Error», «Sums of Squares», «Mean Square», «Mean Percentage Error», «Mean Absolute Percentage Error» (таблиці 2.9 – 2.11 відповідно для трьох розглянутих напрямків кібершахрайських атак).

Таблиця 2.9 – Показники точності прогнозів показника «Кібератаки на комп'ютерні системи фінансової установи»

Назва помилки	Модель 1	Модель 2	Модель 3	Модель 4	Модель 5	Модель 6
Mean Error	0,0539	-0,0088*	0,1152	0,0229	-0,0472	0,0917
Mean Absolute Error	4,3026	4,2923*	4,2948	4,3488	4,3289	4,3471
Sums of Squares	8174,6881	8160,2652	8160,2379*	9340,3966	9300,4093	9312,8601
Mean Square	31,3206	31,2654	31,2653*	35,7870	35,6338	35,6815
Mean Percentage Error	-0,3187	-0,4186	-0,2200*	-0,4462	-0,5556	-0,3336
Mean Absolute Percentage Error	6,9939	6,9803	6,9776*	7,0286	6,9983	7,0197

* Найменші значення похибок виділені сірим кольором

Таблиця 2.10 – Показники точності прогнозів показника «Кібератаки на мережеву інфраструктуру фінансової установи»

Назва помилки	Модель 1	Модель 2	Модель 3	Модель 4	Модель 5	Модель 6	Модель 7	Модель 8
Mean Error	0,1481	0,1503	0,2695	0,0162*	0,0476	0,0507	0,1810	-0,0699
Mean Absolute Error	5,9178	5,9115	5,9130	5,8966*	5,9721	5,9706	5,9709	5,9698
Sums of Squares	14991,5	14869,8	14784,5	14545,7*	15997,2	15907,2	15828,7	15713,2
Mean Square	57,4386	56,9725	56,6455	55,7306*	61,2920	60,9472	60,6463	60,2038
Mean Percentage Error	-1,1104	-1,1033	-0,8663*	-1,2791	-1,3870	-1,3739	-1,1162	-1,5270
Mean Absolute Percentage Error	11,3017	11,2936	11,2782*	11,2887	11,3923	11,3905	11,3712	11,4027

* Найменші значення похибок виділені сірим кольором

Таблиця 2.11 – Показники точності прогнозів показника «Кібератаки на хмарну інфраструктуру фінансової установи»

Назва помилки	Модель1	Модель2	Модель3	Модель4	Модель5	Модель6
Mean Error	0,0812	0,0332*	0,0915	0,0430	0,0052*	0,0692
Mean Absolute Error	4,4419	4,4221	4,4229	4,6343	4,6063	4,3387*
Sums of Squares	10845,971	10828,033	10827,360*	11894,42	11833,363	11525,863
Mean Square	41,5554	41,4867	41,4841*	45,5725	45,3386	44,1604
Mean Percentage Error	-1,6617	-1,7840	-1,5980*	-1,6119	-1,7041	-1,6788
Mean Absolute Percentage Error	13,4378	13,3832	13,3663	13,9018	13,8019	12,9143*

* Найменші значення похибок виділені сірим кольором

Аналіз розрахованих показників точності дозволив обрати відповідні моделі для досліджуваних часових рядів. Для ряду «Кібератаки на комп’ютерні системи фінансової установи» найбільш точною за більшою кількістю показників виявилася модель 3 (табл. 2.9) – тренд-циклічна адитивна модель з експоненційним трендом. Тренд-циклічна адитивна модель з затухаючим трендом є точною для ряду «Кібератаки на мережеву інфраструктуру фінансової установи» (табл. 2.10). Тренд-циклічна адитивна модель з експоненційним трендом показала найкращі результати для «Кібератак на хмарну інфраструктуру фінансової установи» (табл. 2.11). Отримані результати також підтвердили, що досліджувані ряди слідують адитивному процесу та мають трендову і сезонну складові.

Досліджувана тема щодо прогнозування інформаційних трендів кіберзлочинів набуває актуальності у зв’язку із стрімким зростанням їх рівня за останнє десятиліття. Наслідки кіберзлочинності відчуваються по всьому світу, що пов’язано із збільшенням фінансових втрат від викрадення, втрати та відновлення персональної інформації, даних суб’єктів господарювання, урядових організації, тощо. Особливо відчутною дана проблема є в умовах війн та світових пандемій, оскільки вони формують сприятливі умови для кіберзлочинців та кібершахраїв. Саме тому їх попередження та завчасне виявлення є стратегічним завданням у боротьбі з цим явищем.

В роботі встановлено, що наукова спільнота активно досліджує проблематику кіберзлочинності. Вони приділяють увагу макроекономічним проблемам, а саме її впливу на макроекономічну стабільність, інноваційні можливості країни, її імідж, а також зростання тіньового сектору. Також науковці досліджують вплив інформаційних технологій на розвиток бізнесу, питання реорганізації бізнес-процесів в умовах впровадження хмарних технологій, умови зростання кіберризиків та заходи організації кібербезпеки. Актуальним є науковий напрямок, пов'язаний із питаннями кіберзлочинів по відношенню до користувачів інформаційних систем та комп’ютерних технологій, які можуть

відбуватися через соціальні мережі, мобільні та Інтернет додатки. Досліджуються психологічні причини кіберзлочинів, мотивація злочинців, та інші фактори.

В роботі запропоновано методологію дослідження, яка передбачала дослідження початкового набору даних на наявність аномальних спостережень за допомогою Z-score, проведення QS-тест для виявлення періоду циклічності рядів, застосування тесту різниць середніх рівнів для доведення гіпотези щодо відсутності тренду, моделювання і прогнозування рядів динаміки на основі методу експоненційного згладжування та побудови адитивних та мультиплікативних моделей циклічності, тренд-циклічних з лінійним, експоненційним та затухаючим трендами, а також проведення оцінки якості побудованих моделей. Вхідними даними було обрано інформаційні тренди запитів Google-користувачів щодо кібератак на комп’ютерні системи, мережеву та хмарну інфраструктуру фінансової установи. Вибір даних був пов'язаний із тими міркуваннями, що в Інтернет-мережі реакція на будь-яку подію є швидшою ніж у практичній діяльності, тому відповідне зростання запитів користувачів ідентифікується як відклик на кіберзлочини.

В результаті проведення декомпозиції обраних часових рядів встановлено, що вони слідують адитивному процесу, мають сезонну та трендову компоненти. Проведення аналізу рядів на предмет наявності аномальних спостережень дозволило встановити, що інформаційний тренд запитів щодо кібератак на комп’ютерні системи містить одне аномальне спостереження, тренд запитів щодо кібератак на мережеву інфраструктуру – 5, на хмарну інфраструктуру – 3. Їх значення були замінені на середньоарифметичні значення спостережень, що передують та слідують ним. Проведення QS-тест визначило, що період циклічності дорівнює 48 для всіх трьох рядів, що також підтверджується візуалізацією їх сезонної компоненти. За результатами тесту перевірки різниць середніх рівнів було виявлено, що інформаційні тренди запитів щодо кібератак на комп’ютерні системи та хмарну інфраструктуру мають неоднорідні дисперсії, а ряд запитів щодо кібератак на мережеву інфраструктуру має однорідні. Але дослідження значень критерія Стьюдента дозволило встановити, що ряди

нестаціонарні та мають трендову складову, тому для їх моделювання та прогнозування можна використовувати моделі експоненційного згладжування. В результаті їх побудови та проведення оцінки якості визначено, що адитивна тренд-циклічна модель з експоненційним трендом гарно моделює та прогнозує ряди запитів щодо кібератак на комп'ютерні системи та хмарну інфраструктуру, а адитивна тренд-циклічна модель із затухаючим трендом – ряд запитів щодо кібератак на мережеву інфраструктуру.

Запропоновану в роботі методологію та результати прогнозування інформаційних трендів запитів користувачів щодо кіберзлочинів доцільно використовувати для удосконалення стратегії боротьби із кіберзлочинами як на рівні держави, так і на рівні фінансових установ. Дослідження подібних тенденцій та їх прогнозування дозволить у майбутньому попереджати масові кібератаки, які сьогодні є поширеними в рамках ведення кібервійн та кібертероризму. Отримання подібних прогнозів сприятиме активізації заходів кібербезпеки на прогнозовані періоди у більш активному режимі, а також швидше реагувати в ситуаціях, пов'язаних із кібератаками на різні об'єкти комп'ютерної та мережевої інфраструктури.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**].

2.4.2 Розроблення ударно-хвильової моделі впливу кібершахрайських атак на рівень фінансової безпеки

Статистичну базу даних дослідження кібершахрайських атак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи, а також рівня фінансової безпеки сформуємо на основі застосування інструментарію Google Trends. В якості індикаторів розглянемо кількість запитів інтернет-користувачів до доданих понять за період 16.04.2017 по 10.04.2022 в розрізі потижневих рівнів. Для проведення дескриптивного аналізу отриманих часових рядів, побудуємо графіки (рисунки 2.146-2.149) за допомогою програми Statistica.

Для опису інформаційних війн були обрані кількісні критерії: кількість запитів інтернет-користувачів до понять кібершахрайські атаки на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи. На основі рівнів обраних для дослідження часових рядів за допомогою щотижневих даних за проміжок часу з 16.04.2017 по 10.04.2022 були ідентифіковані «інформаційні бульбашки» та прояви їх розриву: computer system – 10.09.2017, 23.09.2018, 15.09.2019, 11.10.2020, 19.09.2021 (рисунок 2.147), cyber fraud – 26.11.2017, 25.11.2018, 01.12.2019, 12.07.2020, 09.01.2022 (рисунок 2.146), network infrastructure – 22.10.2017, 13.10.2019, 13.09.2020, 12.09.2021, 10.04.2022 (рисунок 2.148), cloud infrastructure – 22.10.2017, 26.04.2020, 21.02.2021, 12.12.2021 (рисунок 2.149).

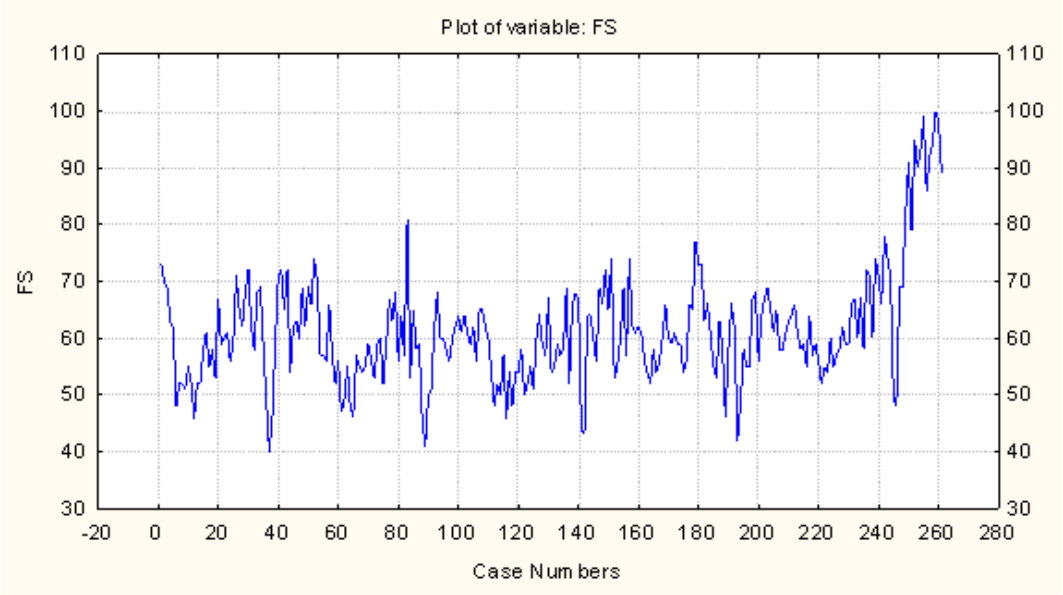


Рисунок 2.146 – Потижнева динаміка показника кількості запитів інтернет-користувачів до поняття фінансова безпека

Кожна із виділених дат виступає проявом розриву «інформаційної бульбашки», коли зазначені кількісні індикатори характеризують варіацію «маси», що має стрибкоподібний характер та визначає «енергію» відповідної ударної хвилі.

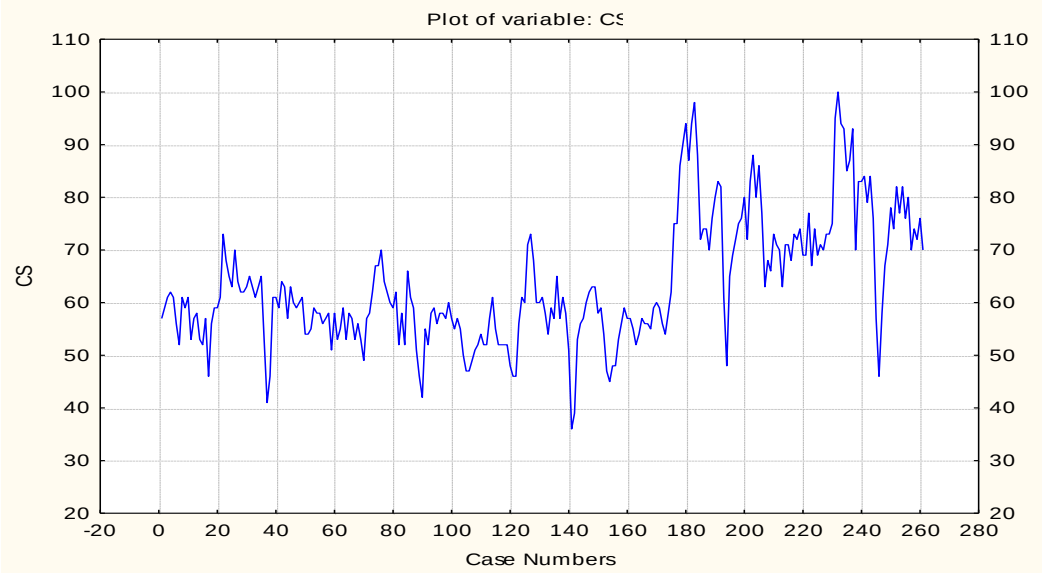


Рисунок 2.147 – Потижнева динаміка показника кількості запитів інтернет-користувачів до поняття кібернетичних атак на комп'ютерні системи фінансової установи

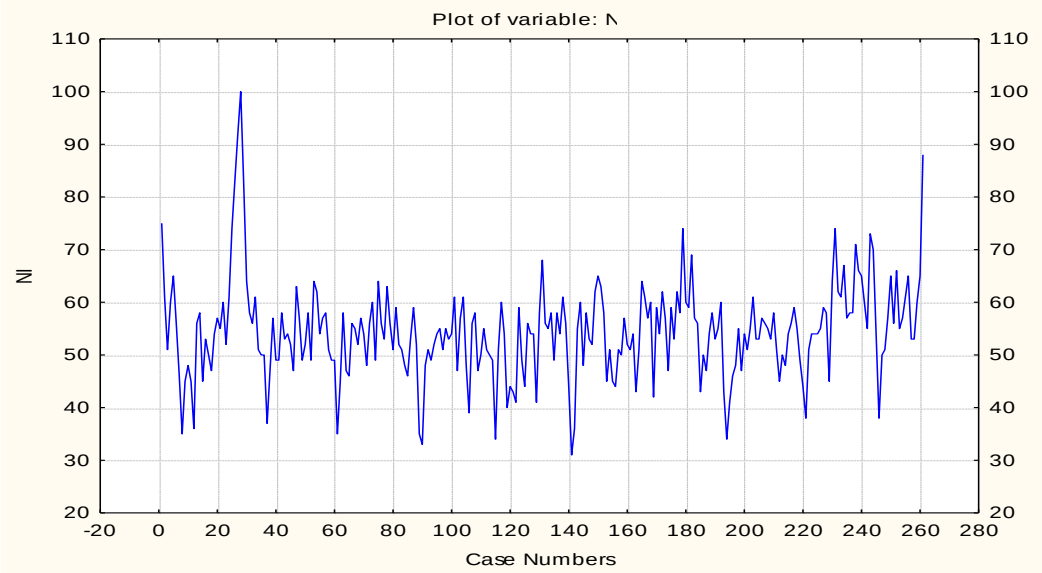


Рисунок 2.148 – Потижнева динаміка показника кількості запитів інтернет-користувачів до поняття кібернетичних атак на мережеву інфраструктуру фінансової установи

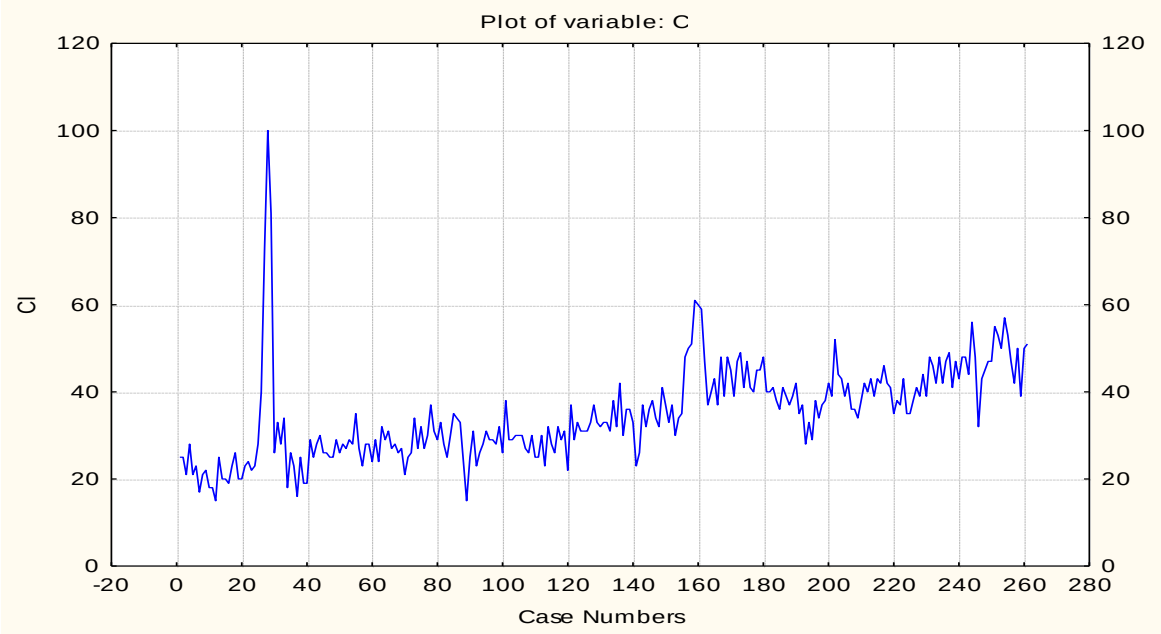


Рисунок 2.149 – Потижнева динаміка показника кількості запитів інтернет-користувачів до поняття кібернетичних атак на хмарну інфраструктуру фінансової установи

На рис. 2.146 – 2.149 представлена динаміка показників фінансової бузпеки та кібершахрайських атак на комп’ютерні системи, мережеву та хмарну інфраструктуру фінансової установи, де зазначені вище дати чітко в розрізі кожного часового ряду ідентифікуються як аномальні рівні у вигляді певного несподіваного стрімкого стрибка з подальшим поверненням до попереднього рівня часового ряду.

Розглянемо теоретичну сутність моделі, яка виступає основною формалізації залежності фінансової безпеки від кібершахрайських атак на комп’ютерні системи, мережеву та хмарну інфраструктуру фінансової установи. Так, модель Сєдова-Тейлора для опису моделі ударної хвилі розповсюдження наслідків «інформаційних війн» (кібершахрайських атак на комп’ютерні системи, мережеву та хмарну інфраструктуру фінансової установи) на фінансову безпеку набуває вигляду:

$$\Delta F\ \ (t) = \frac{C}{t^4} + k_1\left(\frac{C}{t^4}\right)^{3/4} + k_2\left(\frac{C}{t^4}\right)^{2/4} + k_3\left(\frac{C}{t^4}\right)^{1/4} \quad (2.38)$$

де C – енергія в початковий момент після розриву бульбашки, що в нашому випадку пропонується інтерпретувати як показник «інформаційних війн» (кібершахрайських атак на систем/мережевої/хмарної інфраструктури фінансової установи);

$\Delta F\ \ (t)$ – як абсолютний приріст показника фінансової безпеки за період t;

t – період часу після розриву «інформаційної бульбашки» (кількість тижнів);

k_1, k_2, k_3 – характеристики середовищ поширення ударних хвиль наслідків «інформаційних війн» (впливу кібершахрайських атак на фінансову безпеку).

Нелінійна модель Сєдова-Тейлора як задача оптимізації нелінійного програмування набуває наступного вигляду:

$$\begin{cases} K_i \to min \\ K_i = \sum_{j=1}^{V} K_{ij}^{t_j} = \\ = \sum_{j=1}^{A}\left(\frac{e_{t_j}^i}{\tau_j^i} + a_1\left(\frac{e_{t_j}^i}{(\tau_j^i)^2}\right)^{3/4} + a_2\left(\frac{e_{t_j}^i}{(\tau_j^i)^3}\right)^{2/4} + a_3\left(\frac{e_{t_j}^i}{(\tau_j^i)^4}\right)^{1/4}\right) \end{cases} \quad (2.39)$$

де K_i – розсіювання енергії ударної хвилі в і-му каналі поширення кібернетичних атак;

A – кількість інформаційних війн (кібернетичних атак);

$e_{t_j}^i$ – значення початкової енергії в момент часу t_j;

t_j – момент часу початку j-ої інформаційної війни в і-му напрямку здійснення кібернетичних атак;

τ_j^i - тривалість j-ої інформаційної війни в і-му напрямку здійснення кібернетичних атак.

Переходячи до практичного впровадження моделі Сєдова-Тейлора на прикладі залежності фінансової безпеки від кібершахрайських атак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи, зазначимо, що починаючи із ідентифікованих дат розриву «інформаційної бульбашки» в розрізі розглянутих індикаторів ударна хвиля буде розповсюджуватися, однак може не дійти до фінансової установи і не завдати значних негативних наслідків чи дійшовши не призвести до суттєвого впливу. Початкові значення енергій розривів у моменти кризових явищ відображені в табл. 2.12-2.14 для кібершахрайських атак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи.

Таблиця 2.12 – Початкові енергії розривів у моменти криз для показника computer cyber fraud (computer system)

data	financial security	cyber fraud (computer system)	ΔF	C/t4	(C/t4)^3/4	(C/t4)^2/4	(C/t4)^1/4
10.09.2017	60	73	1	0,0304	0,0728	0,1744	0,4176
23.09.2018	59	70	7	0,0292	0,0706	0,1707	0,4132
15.09.2019	64	73	4	0,0304	0,0728	0,1744	0,4176
11.10.2020	66	98	3	0,0408	0,0908	0,2020	0,4495
19.09.2021	67	100	1	0,0416	0,0922	0,2041	0,4518

Таблиця 2.13 – Початкові енергії розривів у моменти криз для показника cyber fraud (network infrastructure)

data	financial security	cyber fraud (network infrastructure)	ΔF	C/t4	(C/t4)^3/4	(C/t4)^2/4	(C/t4)^1/4
22.10.2017	62	100	-3	0,0416	0,0922	0,2041	0,4518
13.10.2019	54	68	-13	0,0283	0,0690	0,1683	0,4102
13.09.2020	77	74	12	0,0308	0,0736	0,1756	0,4190
12.09.2021	66	74	7	0,0308	0,0736	0,1756	0,4190
10.04.2022	89	88	-9	0,0367	0,0838	0,1914	0,4375

Таблиця 2.14 – Початкові енергії розривів у моменти криз для показника cyber fraud (cloud infrastructure)

data	financial security	cyber fraud (cloud infrastructure)	ΔF	C/t4	(C/t4)^3/4	(C/t4)^2/4	(C/t4)^1/4
22.10.2017	62	100	-3	0,0416	0,0922	0,2041	0,4518
26.04.2020	61	61	-1	0,0254	0,0636	0,1594	0,3992
21.02.2021	67	52	3	0,0217	0,0565	0,1472	0,3836
12.12.2021	71	56	-3	0,0233	0,0597	0,1527	0,3908

З метою побудови моделі Сєдова-Тейлора формалізації ударної хвилі наслідків розповсюдження інформаційних війн в процесі здійснення впливу кібершахрайських атак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи на рівень фінансової безпеки виникає необхідність ідентифікації проміжку часу після моменту розриву «інформаційної бульбашки» (кількість тижнів) щодо її поширення. З метою розв'язання зазначеного питання необхідно провести автокореляційний аналіз (рисунки 2.150-157), який дозволяє сформувати графіки корелограм нульових різниць для розглянутих часових рядів як фінансової безпеки, так і кібершахрайських атак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи.

Autocorrelation Function (Spreadsheet1 FS.FS
(Standard errors are white-noise estimates)

Lag	Auto-Corr.	Std.Err.	Box & Ljung Q	p
1	0,68689	0,06154	124,567	0,00000
2	0,56569	0,06142	209,379	0,00000
3	0,45783	0,06130	265,149	0,00000
4	0,40837	0,06118	309,693	0,00000
5	0,33245	0,06106	339,329	0,00000
6	0,30150	0,06095	363,800	0,00000
7	0,27804	0,06083	384,693	0,00000
8	0,23711	0,06071	399,947	0,00000
9	0,16269	0,06059	407,157	0,00000
10	0,16454	0,06047	414,561	0,00000
11	0,13572	0,06034	419,619	0,00000
12	0,07335	0,06022	421,102	0,00000
13	0,03699	0,06010	421,481	0,00000
14	-0,01019	0,05998	421,510	0,00000
15	0,00102	0,05986	421,510	0,00000

Рисунок 2.150 – Значення автокорляційної функції нульових різниць часового ряду кібершахрайських атак

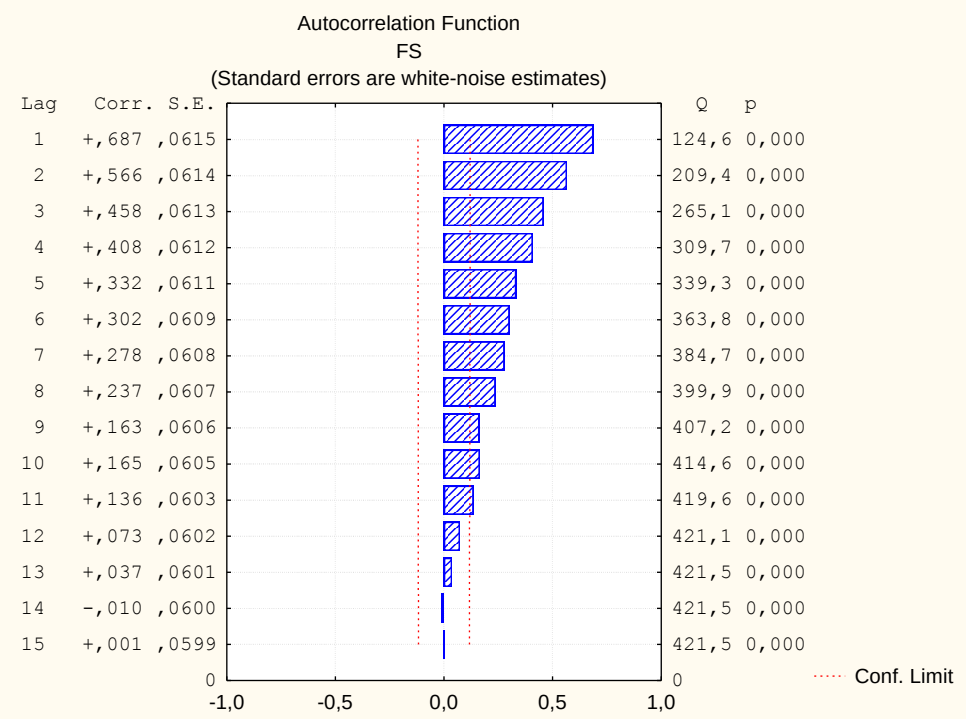


Рисунок 2.151 – Корелограма нульових різниць часового ряду кібершахрайських атак

Lag	Autocorrelation Function (Spreadsheet3.: CS (Standard errors are white-noise estimate Auto-Corr.	Std.Err.	Box & Ljung Q	p
1	0,86421	0,06154	197,18	0,00
2	0,75191	0,06142	347,02	0,00
3	0,65812	0,06130	462,26	0,00
4	0,58955	0,06118	555,09	0,00
5	0,52455	0,06106	628,87	0,00
6	0,48948	0,06095	693,37	0,00
7	0,46760	0,06083	752,46	0,00
8	0,46503	0,06071	811,13	0,00
9	0,43636	0,06059	863,00	0,00
10	0,41992	0,06047	911,22	0,00
11	0,40700	0,06034	956,71	0,00
12	0,40872	0,06022	1002,76	0,00
13	0,39974	0,06010	1046,99	0,00
14	0,35372	0,05998	1081,76	0,00
15	0,34255	0,05986	1114,50	0,00

Рисунок 2.152 – Значення автокорляційної функції нульових різниць часового ряду кібершахрайських атак на комп’ютерні системи

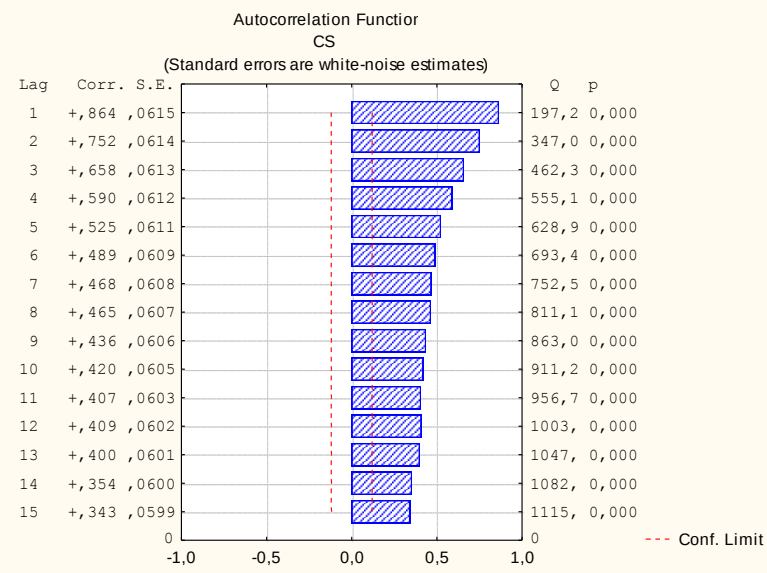


Рисунок 2.153 – Корелограма нульових різниць часового ряду кібершахрайських атак на комп’ютерні системи

	Autocorrelation Function (Spreadsheet3.s NI (Standard errors are white-noise estimate			
Lag	Auto-Corr.	Std.Err.	Box & Ljung Q	p
1	0,53193	0,06154	74,703	0,00000
2	0,26924	0,06142	93,916	0,00000
3	0,24074	0,06130	109,336	0,00000
4	0,14934	0,06118	115,293	0,00000
5	0,11865	0,06106	119,068	0,00000
6	0,09744	0,06095	121,624	0,00000
7	0,03297	0,06083	121,918	0,00000
8	-0,03992	0,06071	122,350	0,00000
9	-0,02590	0,06059	122,533	0,00000
10	-0,10795	0,06047	125,721	0,00000
11	-0,11355	0,06034	129,261	0,00000
12	-0,04596	0,06022	129,843	0,00000
13	-0,02616	0,06010	130,033	0,00000
14	-0,09446	0,05998	132,512	0,00000
15	-0,10024	0,05986	135,317	0,00000

Рисунок 2.154 – Значення автокорляційної функції нульових різниць часового ряду кібершахрайських атак на мережеву інфраструктуру фінансової установи

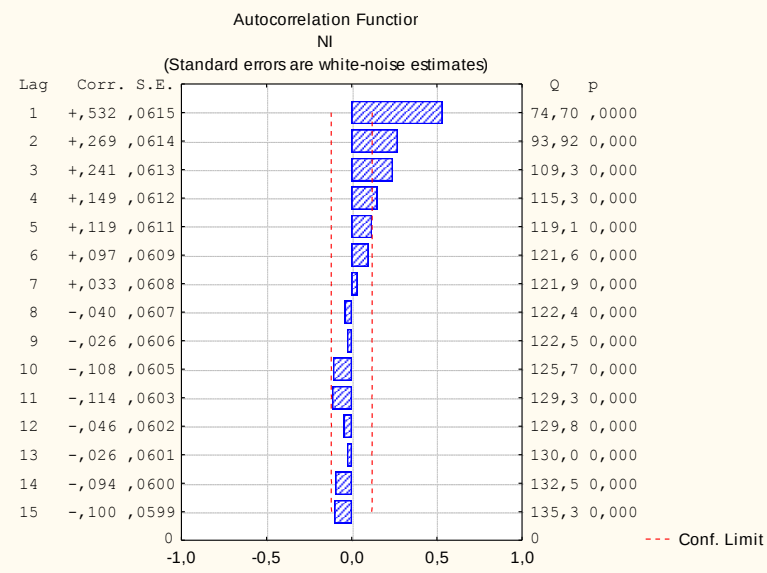


Рисунок 2.155 – Корелограма нульових різниць часового ряду кібершахрайських атак на мережеву інфраструктуру фінансової установи

Lag	Autocorrelation Function (Spreadsheet3.s CI (Standard errors are white-noise estimate Auto-Corr.	Std.Err.	Box & Ljung Q	p
1	0,77739	0,06154	159,551	0,00
2	0,62684	0,06142	263,691	0,00
3	0,50649	0,06130	331,944	0,00
4	0,44977	0,06118	385,977	0,00
5	0,39864	0,06106	428,589	0,00
6	0,37002	0,06095	465,446	0,00
7	0,33746	0,06083	496,222	0,00
8	0,31626	0,06071	523,360	0,00
9	0,30472	0,06059	548,654	0,00
10	0,32584	0,06047	577,691	0,00
11	0,30401	0,06034	603,069	0,00
12	0,28265	0,06022	625,094	0,00
13	0,30584	0,06010	650,985	0,00
14	0,29088	0,05998	674,500	0,00
15	0,27846	0,05986	696,137	0,00

Рисунок 2.156 – Значення автокорляційної функції нульових різниць часового ряду кібершахрайських атак на хмарну інфраструктуру фінансової установи

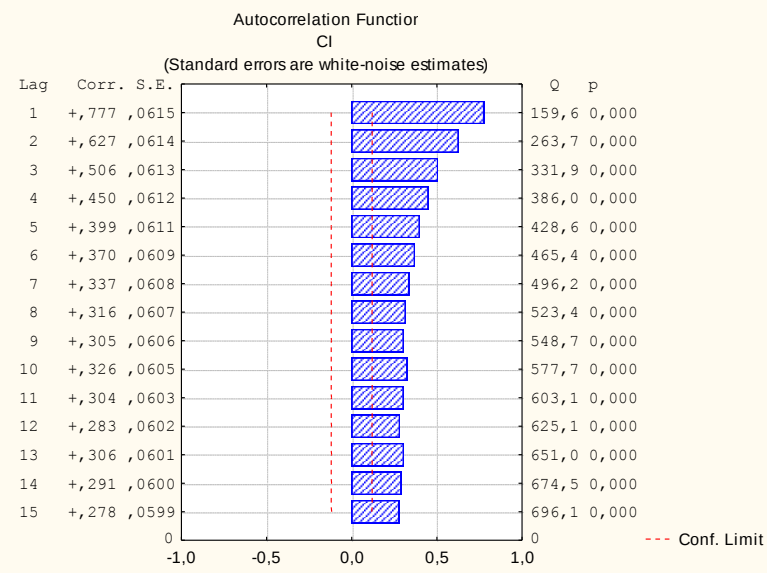


Рисунок 2.157 – Корелограма нульових різниць часового ряду кібершахрайських атак на хмарну інфраструктуру фінансової установи

Аналіз рисунків 2.150 і 2.151 дозволяє стверджувати, що для часового ряду кібершахрайських атак (нульові різниці) значення коефіцієнтів автокореляції для різних часових лагів є статистично значущими, постійно варіюються. Аналогічна тенденція спостерігається і для корелограм кібершахрайських атак на комп'ютерні системи, та хмарну інфраструктуру фінансової установи. Винятком виступають лише значення коефіцієнтів автокореляції для мережевої інфраструктури, оскільки статистично значущими є лише перші три рівня, але вони мають коливальну тенденцію. Тому пропонується за часовий проміжок моделі Сєдова-Тейлора взяти тиждень (тобто 7 денний інтервал).

Переходячи до наступного кроку побудови моделі Сєдова-Тейлора для формалізації моделі ударної хвилі розповсюдження наслідків інформаційних війн у вигляді залежності фінансової безпеки від кібершахрайських атак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи виникає необхідність обчислення коефіцієнтів економіко-математичної моделі розсіювання енергії ударних хвиль. З цією метою, базуючись на даних таблиць 1 – 3 в розрізі кібершахрайських атак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи відповідно, які відображують початкові

рівні енергії у кризові моменти часу, а також беручи до уваги обґрунтований на основі результатів проведення автокореляційного аналізу тижневий інтервал розповсюдження наслідків в розрізі фінансової безпеки після моменту розриву так званої «інформаційнох бульбашки», оптимізаційну модель загального вигляду (4.19) запишемо окремо в розрізі кожного напрямку кібершахрайських атак на основі наявних статистичних даних у вигляді формул (2.40) – (2.42). Крім того, для формування правої частини системи обмежень оптимізаційної задачі обчислимо значення абсолютних прирості поточного та попереднього рівнів відповідних часових рядів:

- для кібершахрайських атак на комп'ютерні системи фінансової установи:

$$\begin{cases} \frac{73}{7^4} + a_1\left(\frac{73}{7^4}\right)^{\frac{3}{4}} + a_2\left(\frac{73}{7^4}\right)^{\frac{2}{4}} + a_3\left(\frac{73}{7^4}\right)^{\frac{1}{4}} \to min \\ \frac{73}{7^4} + a_1\left(\frac{73}{7^4}\right)^{\frac{3}{4}} + a_2\left(\frac{73}{7^4}\right)^{\frac{2}{4}} + a_3\left(\frac{73}{7^4}\right)^{\frac{1}{4}} = 1 \\ \frac{70}{7^4} + a_1\left(\frac{70}{7^4}\right)^{\frac{3}{4}} + a_2\left(\frac{70}{7^4}\right)^{\frac{2}{4}} + a_3\left(\frac{70}{7^4}\right)^{\frac{1}{4}} = 7 \\ \frac{73}{7^4} + a_1\left(\frac{73}{7^4}\right)^{\frac{3}{4}} + a_2\left(\frac{73}{7^4}\right)^{\frac{2}{4}} + a_3\left(\frac{73}{7^4}\right)^{\frac{1}{4}} = 4 \\ \frac{98}{7^4} + a_1\left(\frac{98}{7^4}\right)^{\frac{3}{4}} + a_2\left(\frac{98}{7^4}\right)^{\frac{2}{4}} + a_3\left(\frac{98}{7^4}\right)^{\frac{1}{4}} = 3 \\ \frac{100}{7^4} + a_1\left(\frac{100}{7^4}\right)^{\frac{3}{4}} + a_2\left(\frac{100}{7^4}\right)^{\frac{2}{4}} + a_3\left(\frac{100}{7^4}\right)^{\frac{1}{4}} = 1 \end{cases} \quad (2.40)$$

- для кібершахрайських атак на мережеву інфраструктуру фінансової установи:

$$\begin{cases} \frac{100}{7^4}+a_1\left(\frac{100}{7^4}\right)^{\frac{3}{4}}+a_2\left(\frac{100}{7^4}\right)^{\frac{2}{4}}+a_3\left(\frac{100}{7^4}\right)^{\frac{1}{4}}\to min \\ \frac{100}{7^4}+a_1\left(\frac{100}{7^4}\right)^{\frac{3}{4}}+a_2\left(\frac{100}{7^4}\right)^{\frac{2}{4}}+a_3\left(\frac{100}{7^4}\right)^{\frac{1}{4}}=-3 \\ \frac{68}{7^4}+a_1\left(\frac{68}{7^4}\right)^{\frac{3}{4}}+a_2\left(\frac{68}{7^4}\right)^{\frac{2}{4}}+a_3\left(\frac{68}{7^4}\right)^{\frac{1}{4}}=-13 \\ \frac{74}{7^4}+a_1\left(\frac{74}{7^4}\right)^{\frac{3}{4}}+a_2\left(\frac{74}{7^4}\right)^{\frac{2}{4}}+a_3\left(\frac{74}{7^4}\right)^{\frac{1}{4}}=12 \\ \frac{74}{7^4}+a_1\left(\frac{74}{7^4}\right)^{\frac{3}{4}}+a_2\left(\frac{74}{7^4}\right)^{\frac{2}{4}}+a_3\left(\frac{74}{7^4}\right)^{\frac{1}{4}}=7 \\ \frac{88}{7^4}+a_1\left(\frac{88}{7^4}\right)^{\frac{3}{4}}+a_2\left(\frac{88}{7^4}\right)^{\frac{2}{4}}+a_3\left(\frac{88}{7^4}\right)^{\frac{1}{4}}=-9 \end{cases} \quad (2.41)$$

- кібершахрайських атак на хмарну інфраструктуру фінансової установи:

$$\begin{cases} \frac{100}{7^4}+a_1\left(\frac{100}{7^4}\right)^{\frac{3}{4}}+a_2\left(\frac{100}{7^4}\right)^{\frac{2}{4}}+a_3\left(\frac{100}{7^4}\right)^{\frac{1}{4}}\to min \\ \frac{100}{7^4}+a_1\left(\frac{100}{7^4}\right)^{\frac{3}{4}}+a_2\left(\frac{100}{7^4}\right)^{\frac{2}{4}}+a_3\left(\frac{100}{7^4}\right)^{\frac{1}{4}}=-3 \\ \frac{61}{7^4}+a_1\left(\frac{61}{7^4}\right)^{\frac{3}{4}}+a_2\left(\frac{61}{7^4}\right)^{\frac{2}{4}}+a_3\left(\frac{61}{7^4}\right)^{\frac{1}{4}}=-1 \\ \frac{52}{7^4}+a_1\left(\frac{52}{7^4}\right)^{\frac{3}{4}}+a_2\left(\frac{52}{7^4}\right)^{\frac{2}{4}}+a_3\left(\frac{52}{7^4}\right)^{\frac{1}{4}}=3 \\ \frac{56}{7^4}+a_1\left(\frac{56}{7^4}\right)^{\frac{3}{4}}+a_2\left(\frac{56}{7^4}\right)^{\frac{2}{4}}+a_3\left(\frac{56}{7^4}\right)^{\frac{1}{4}}=-3 \end{cases} \quad (2.42)$$

З метою розв’язання оптимізаційних задач, які передбачають розрахунок коефіцієнтів моделі Сєдова-Тейлора формалізації моделі ударної хвилі поширення негативних наслідків інформаційних війн у вигляді залежності фінансової безпеки від кібершахрайських атак на комп’ютерні системи, мережеву та хмарну інфраструктуру фінансової установи, пропонується

скористатись можливостями інструментарію MS Excel «Пошук рішення», зокрема методом ОПГ (метод узагальненого приведеного градієнта).

Таким чином, вирішення систем (2.40) – (2.42) шляхом пошуку відповідних коефіцієнтів надає можливість формалізувати шукану модель Сєдова-Тейлора для опису ударної хвилі поширення негативних наслідків інформаційних війн у вигляді залежності фінансової безпеки від кібершахрайських атак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи в наступному вигляді:

- в розрізі кібершахрайських атак на комп'ютерні системи фінансової установи:

$$\Delta F \quad (t) = \frac{C}{t^4} - 0.7187 \cdot \left(\frac{C}{t^4}\right)^{\frac{3}{4}} + 0.0601 \cdot \left(\frac{C}{t^4}\right)^{\frac{2}{4}} + 0.0273 \cdot \left(\frac{C}{t^4}\right)^{1/4} \qquad (2.43)$$

- у вигляді кібершахрайських атак на мережеву інфраструктуру фінансової установи:

$$\Delta F \quad (t) = \frac{C}{t^4} - 0.2510 \cdot \left(\frac{C}{t^4}\right)^{\frac{3}{4}} - 0.3076 \cdot \left(\frac{C}{t^4}\right)^{\frac{2}{4}} + 0.0994 \cdot \left(\frac{C}{t^4}\right)^{1/4} \qquad (2.44)$$

- у вигляді кібершахрайських атак на хмарну інфраструктуру фінансової установи:

$$\Delta F \quad (t) = \frac{C}{t^4} - 0.2214 \cdot \left(\frac{C}{t^4}\right)^{\frac{3}{4}} - 0.2784 \cdot \left(\frac{C}{t^4}\right)^{\frac{2}{4}} + 0.0829 \cdot \left(\frac{C}{t^4}\right)^{1/4} \qquad (2.45)$$

Візуалізація поширення наслідків ударної хвилі після моменту розриву «інформаційної бульбашки» набуває вигляду, представленому на рисунках 2.158 – 2.160 відповідно для кібершахрайських атак на комп'ютерні системи, мережеву та хмарну інфраструктуру фінансової установи.

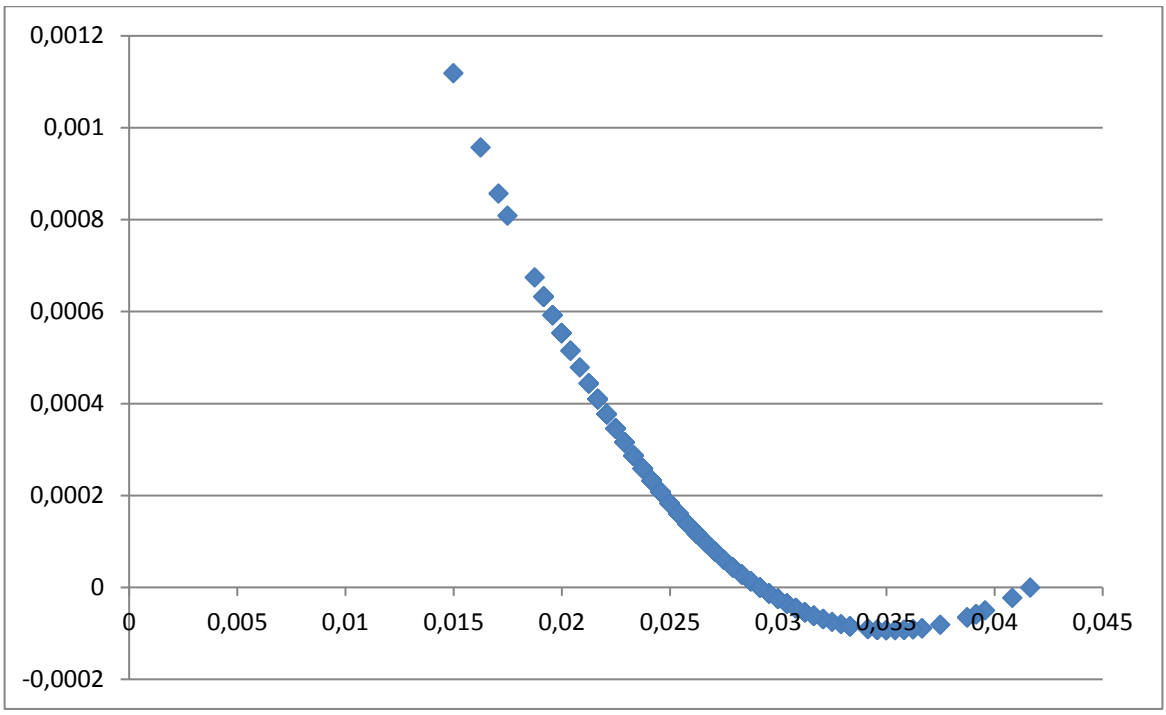


Рисунок 2.158 – Візуалізація поширення наслідків ударної хвилі після моменту розриву «інформаційної бульбашки» для впливу кібершахрайських атак на комп'ютерні системи фінансової установи на рівень фінансової безпеки

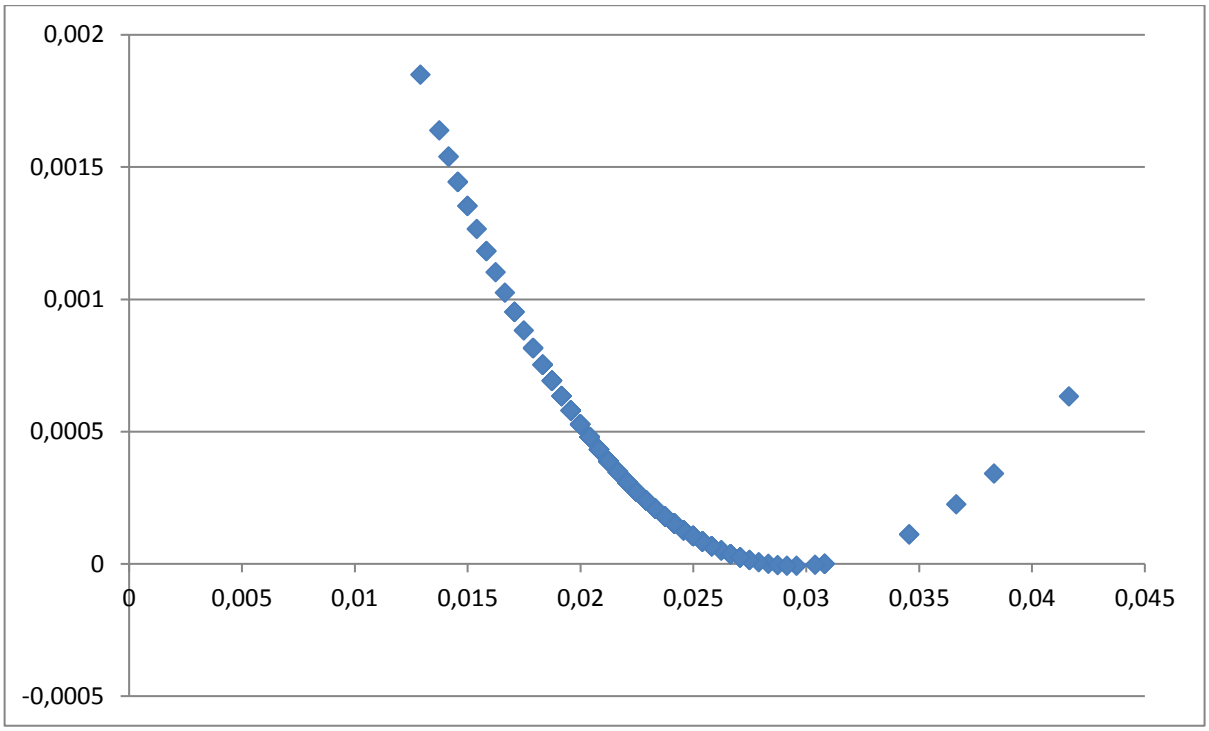


Рисунок 2.159 – Візуалізація поширення наслідків ударної хвилі після моменту розриву «інформаційної бульбашки» для впливу кібершахрайських атак на мережеву інфраструктуру фінансової установи на рівень фінансової безпеки

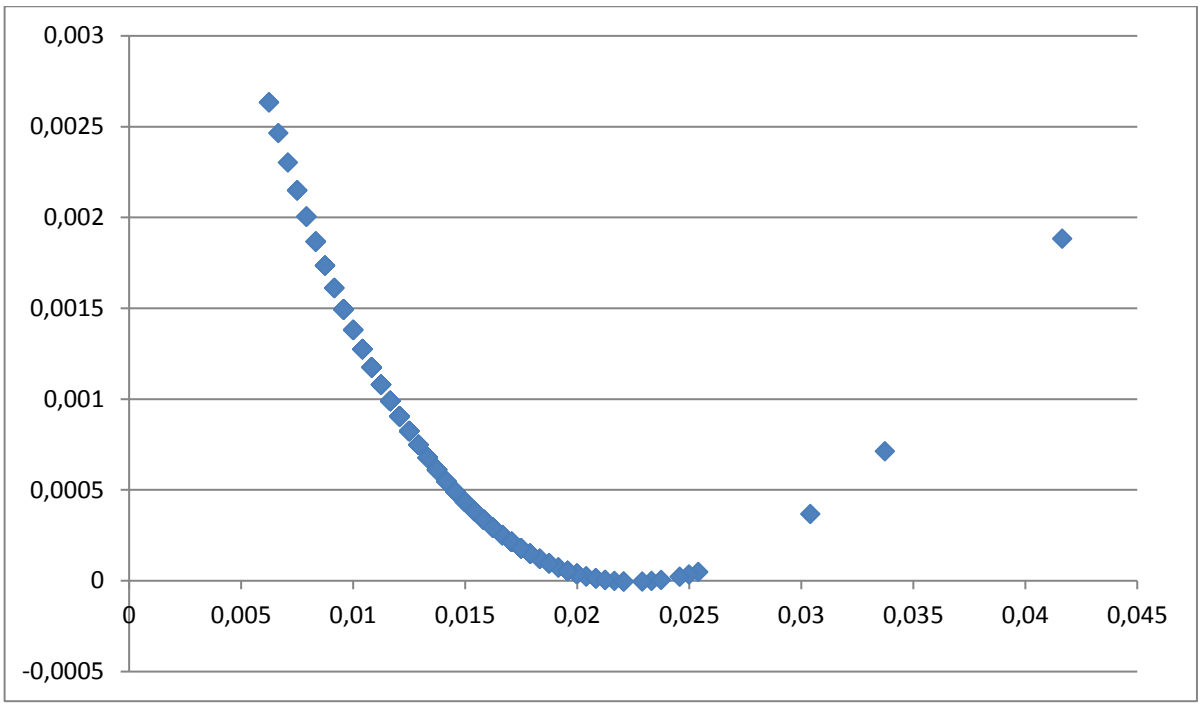


Рисунок 2.160 – Візуалізація поширення наслідків ударної хвилі після моменту розриву «інформаційної бульбашки» для впливу кібершахрайських атак на хмарну інфраструктуру фінансової установи на рівень фінансової безпеки

Аналіз рисунків 2.158–2.160 дозволяє констатувати наявність точки розриву «інформаційної бульбашки» (здійснення кібершахрайських атак) з подальшою адаптацією комп'ютерних систем, мережевої та хмарної інфраструктури фінансової установи і боротьбою з негативними наслідками поширення ударної хвилі на рівень фінансової безпеки, свідченням чого виступає зростаюча права гілка параболи.

З метою визначення критичних рівнів фінансової безпеки, кібершахрайських атак в розрізі комп'ютерних систем, мережевої та хмарної інфраструктури фінансової установи, перевищення яких супроводжується розривом «інформаційної бульбашки», визначимо такі рівні зазначених характеристик, за яких функція залежності фінансової безпеки від кібершахрайських атак має точку перегину. Такими значеннями виступають: 0,036, 0,030, 0,022. Для визначених рівнів обчислимо значення функції залежності фінансової безпеки від кібершахрайських атак другої похідної (рисунки 2.161–2.163).

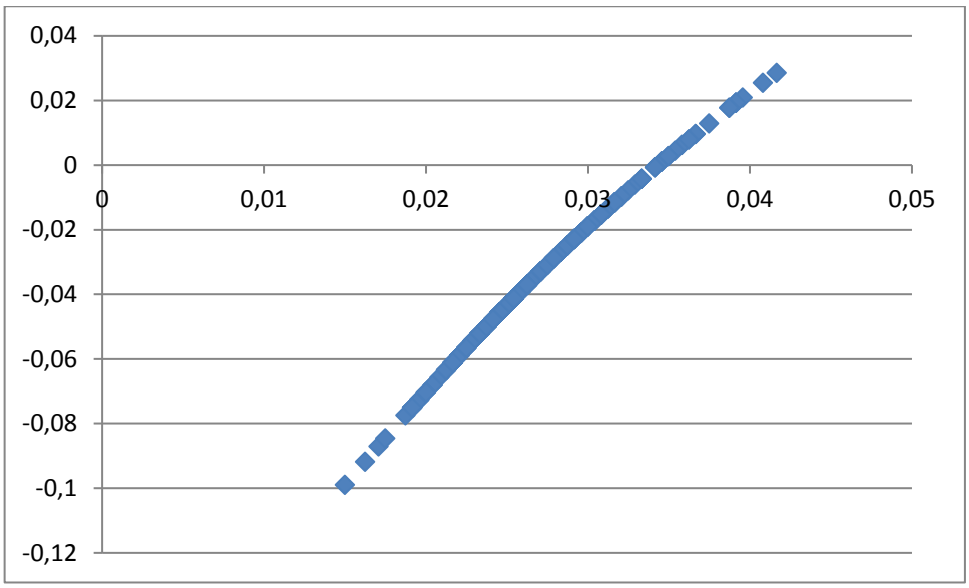


Рисунок 2.161 - Візуалізація функції першої похідної моделі Сєдова-Тейлора для опису ударної хвилі розповсюдження наслідків інформаційних війн у вигляді впливу кібершахрайських атак на комп'ютерні системи фінансової установи на рівень фінансової безпеки

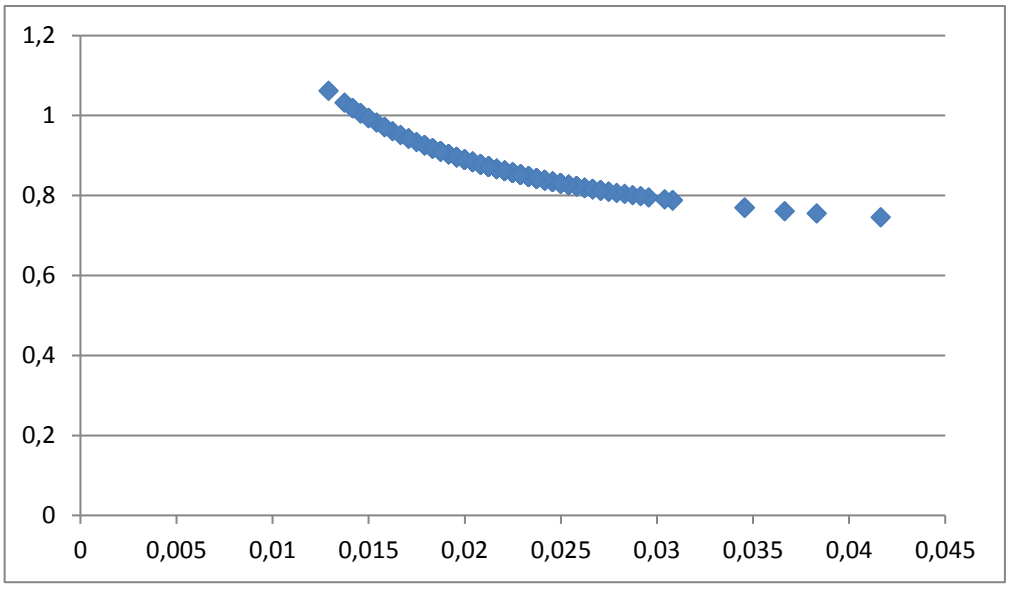


Рисунок 2.162 - Візуалізація функції першої похідної моделі Сєдова-Тейлора для опису ударної хвилі розповсюдження наслідків інформаційних війн у вигляді впливу кібершахрайських атак на мережеву інфраструктуру фінансової установи на рівень фінансової безпеки

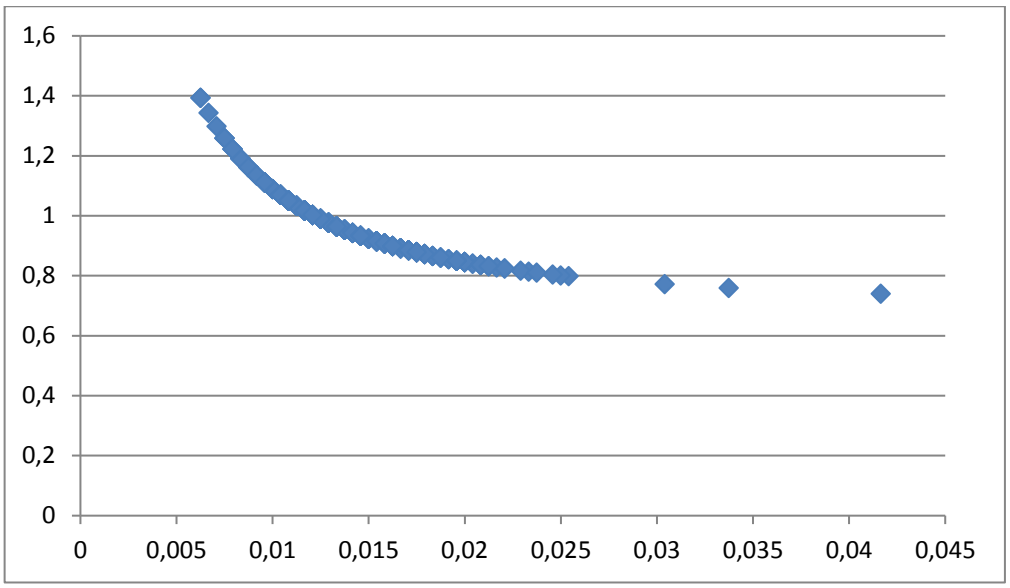


Рисунок 2.163 - Візуалізація функції першої похідної моделі Сєдова-Тейлора для опису ударної хвилі розповсюдження наслідків інформаційних війн у вигляді впливу кібершахрайських атак на хмарну інфраструктуру фінансової установи на рівень фінансової безпеки

Аналіз рисунків 2.161-2.163 дозволяє констатувати, що в момент розриву інформаційної бульбашки найбільший впливу кібершахрайських атак на рівень фінансової безпеки спостерігається в розрізі атак на комп'ютерні системи фінансової установи (0,797%), хоча в розрізі мережевої інфраструктури та хмарної інфраструктури фінансової установи рівень фінансової безпеки під впливом кібершахрайських атак відповідає рівням 82,86% та 78,82% відповідно. Після досягнення зазначених рівнів активність здійснення кібершахрайських атак знижується.

Даний пункт було виконано із використанням матеріалів публікацій виконавців [**Ошибка! Источник ссылки не найден.**].