Міністерство освіти і науки України
Сумський державний університет
Навчально-науковий інститут бізнес-технологій «УАБС»
Кафедра економічної кібернетики

# КВАЛІФІКАЦІЙНА МАГІСТЕРСЬКА РОБОТА
# на тему «ПРОГНОЗУВАННЯ МОЖЛИВОСТІ КІБЕРАТАК ФІНАНСОВИХ УСТАНОВ»

Виконав студент 2 курсу, групи ЕК.м-81а
(номер курсу) (шифр групи)
Спеціальності 051 «Економіка»
(«Економічна кібернетика»)
Лосина Є.С.
(прізвище, ініціали студента)
Керівник к.ф.-м.н., доцент Братушка С.М.
(посада, науковий ступінь, прізвище, ініціали)

Суми – 2019 рік

РЕФЕРАТ

кваліфікаційної магістерської роботи на тему

«ПРОГНОЗУВАННЯ МОЖЛИВОСТІ КІБЕРАТАК ФІНАНСОВИХ УСТАНОВ»

студента Лосини Євгенія Сергійовича
(прізвище, ім’я, по батькові)

Актуальність теми, обраної для дослідження, визначається тим, що щороку банки та інші фінансові установи несуть збитки від діяльності кіберзлочинців. У тих випадках, коли кількість кібератак вдається спрогнозувати, то і відділ безпеки фінансової установи зможе прийняти правильне рішення для збільшення захисту в той чи інший момент.

Мета кваліфікаційної магістерської роботи полягає у є побудові моделі прогнозування можливості кібератак на фінансові установи.

Об’єктом дослідження є обсяги кібератак на фінансові установи.

Предметом дослідження є методи і моделі прогнозування часових рядів.

Для досягнення поставленої мети сформовано наступні задачі дослідження:

- розглянути напрямки кібератак і способи їх реалізації;
- проаналізувати методи і моделі прогнозування кібератак;
- сформувати вимоги до моделі;
- обрати програмне забезпечення для реалізації моделі;
- побудувати моделі кібератак;
- перевірити адекватність моделі;
- запропонувати шляхи вдосконалення побудованої моделі.

Для досягнення поставленої мети та задач дослідження були використані такі методи дослідження: аналіз часових рядів, нейромережеве прогнозування.

Інформаційною базою кваліфікаційної магістерської роботи є дані, зібрані на сайті hackmageddon.com .

Основний науковий результат кваліфікаційної магістерської роботи полягає у такому: була розроблена і перевірена на адеквантість модель прогнозування можливості кібератак, що дозволяє отримати інформацію про кількість кібератак на майбутній період.

Одержані результати можуть бути використані і відділом безпеки фінансової установи для прийняти правильного рішення для збільшення захисту в той чи інший момент.

Результати апробації основних положень кваліфікаційної магістерської роботи розглядалися на IV Всеукраїнській науково-практичній on-line конференції «Проблеми та перспективи розвитку фінансово-кредитної системи України» 21-22 листопада 2019 року.

Ключові слова: кібератака, прогнозування, фінансові установи, моделювання.

Зміст кваліфікаційної магістерської роботи викладено на 40 сторінках. Список використаних джерел із 53 найменувань, розміщений на 6 сторінках. Робота містить 6 таблиць, 19 рисунків, а також 3 додатків, розміщених на 3 сторінках.

Рік виконання кваліфікаційної роботи – 2019 рік.

Рік захисту роботи – 2019 рік.

Міністерство освіти і науки України

Сумський державний університет

Навчально-науковий інститут бізнес-технологій «УАБС»

Кафедра економічної кібернетики

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____
(науковий ступінь, вчене звання)

_____
(підпис) (ініціали, прізвище)
“___” _____20__ р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ МАГІСТЕРСЬКУ РОБОТУ
(спеціальність 051 «Економіка» («Економічна кібернетика»))
студенту 2 курсу, групи ЕК.м.-81а

Лосині Євгенію Сергійовичу
(прізвище, ім’я, по батькові студента)

1. Тема роботи: «Прогнозування можливості кібератак фінансових установ»
затверджена наказом по університету від «__» _____ 20___ року №____

2. Термін подання студентом закінченої роботи «__» грудня 2019 року

3. Мета кваліфікаційної роботи: *побудова моделі прогнозування можливих обсягів кібератак на фінансові установи.*

4. Об’єкт дослідження: *обсяги кібератак на фінансові установи.*

5. Предмет дослідження: *методи і моделі прогнозування часових рядів.*

6. Кваліфікаційна робота виконується на матеріалах: *результатів проходження переддипломної практики, а також даних про обсяги кібератак на банківські установи*

7. Орієнтовний план кваліфікаційної роботи, терміни подання розділів керівникові та зміст завдань для виконання поставленої мети
Розділ 1 ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ОСНОВИ ПРОГНОЗУВАННЯ МОЖЛИВОСТІ КІБЕРАТАК ФІНАНСОВИХ УСТАНОВ (до _____)
У розділі 1 необхідно визначити зміст та особливості кібератак на

фінансові установи, провести статистичний аналіз кібератак на фінансові установи, розглянути існуючі методи і моделі прогнозування кількості кібератак.

Розділ 2 РОЗРОБКА МОДЕЛІ ПРОГНОЗУВАННЯ МОЖЛИВИХ ОБСЯГІВ КІБЕРАТАК ФІНАНСОВИХ УСТАНОВ (до _____ р.)

У розділі 2 необхідно обгрунтувати вибір способу реалізації обраних моделей, сформувати вимоги до моделі, підготувати вхідні дані для роботи з ними.

Розділ 3 МОДЕЛЮВАННЯ МОЖЛИВИХ ОБСЯГІВ КІБЕРАТАК НА ФІНАНСОВІ УСТАНОВИ (до _____ р.)

У розділі 3 необхідно обгрунтувати вибір програмного забезпечення для проведення розрахунків, реалізувати запропоновані алгоритми моделей, перевірити адекватність розрахунків.

8. Консультації з роботи:

| Розділ | Прізвище, ініціали та посада консультанта | Підпис, дата | |
|---|---|---|---|
| | | завдання видав | Завдання прийняв |
| 1 | | | |
| 2 | | | |
| 3 | | | |

9. Дата видачі завдання: «___»_____ 20__ року

Керівник кваліфікаційної роботи _____ С.М. Братушка
(підпис) (ініціали, прізвище)

Завдання до виконання одержав _____ Є.С. Лосина
(підпис) (ініціали, прізвище)

ЗМІСТ

# ВСТУП

Ми живемо у такому світі, де злочини можна здійснювати не встаючи із-за комп′ютера. Незначний ризик бути виявленим та високі доходи стали основними причинами для збільшення кількості кіберзлочинів. Незважаючи на те, що деяких учасників кіберугрупувань вдається затримати, на їх місце приходять все нові і нові злочинці, які мають на озброєнні сучасніші методи кібератак.

Завдання прогнозування можливості кібератак на фінансові установи є досить актуальним, оскільки щороку банки та інші фінансові установи несуть збитки від діяльності кіберзлочинців. У тих випадках, коли кількість кібератак вдається спрогнозувати, то і відділ безпеки фінансової установи зможе прийняти правильне рішення для збільшення захисту в той чи інший момент.

Метою даного дослідження є побудова моделі прогнозування можливості кібератак на фінансові установи.

Предметом дослідження виступають методи і моделі прогнозування часових рядів.

Об’єктом є обсяги кібератак на фінансові установи.

Для досягнення поставленої мети сформовано наступні задачі дослідження:

- розглянути напрямки кібератак і способи їх реалізації;
- проаналізувати методи і моделі прогнозування кібератак;
- сформувати вимоги до моделі;
- обрати програмне забезпечення для реалізації моделі;
- побудувати моделі кібератак;
- перевірити адекватність моделі;
- запропонувати шляхи вдосконалення побудованої моделі.

Результатом виконаної роботи повинна стати модель прогнозу можливості кібератак, яка зможе допомогти фахівцю з кібербезпеки при збільшенні атак вчасно прийняти відповідні заходи для запобігання чи зменшення шкоди від атак.

# РОЗДІЛ 1. ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ОСНОВИ ПРОГНОЗУВАННЯ МОЖЛИВОСТІ КІБЕРАТАК ФІНАНСОВИХ УСТАНОВ

## 1.1 Поняття, види та засоби захисту від кібератак

Кібератакою можна назвати дії, які були вчинені в кіберпросторі навмисно, або були здійснені через засоби електронних комунікацій та які ставлять собі за мету досягнення однієї або кількох цілей. Цілями кібератак можуть бути порушення безпеки, штатної роботи комунікаційних або технологічних систем, несанкціонований доступ до них, використання їх для здійснення кібератак на інші об’єкти [1].

За даними дослідження Positive Technologies [2] фінансові організації у 2018 році входять в трійку найпопулярніших цілей кіберзлочинців. Основною метою кібератак на фінансові установи є отримання прямої фінансової вигоди, а також отримання даних, у тому числі даних облікових записів для доступу до фінансових додатків та даних банківських карт. Одним із способів монетизації облікових записів є їх продаж на «чорному ринку» (близько 80% інформації, що продається в «дарквебі» є різні облікові записи та дані банківських карт). При цьому дані для доступу в особисті кабінети в онлайн-банках продаються поштучно.

Середня ціна складає $22 і на рахунку може бути від декількох десятків до декількох тисяч доларів. Середня вартість даних однієї банківської карти з балансом в декілька сотень доларів становить всього $9. Дані банківських карт у майбутньому можуть бути використані для купівлі товарів в Інтернеті або для створення дублікатів банківських карт для зняття готівки в банкоматах [3].

Банкомати також стають ціллю кіберзлочинів. На чорному ринку зараз можна купити ПО для отримання коштів із касет банкомату. Ціна такого ПО висока від $1500 до $5000, але і потенційний прибуток значно перекриває витрати. Так як операції проводяться напряму із банкоматом, клієнти банку збитків не несуть, а страждає тільки власник банкомату [4, 5].

За способом розповсюдження кібератаки можна поділити на масові та цілеспрямовані.

Метою масових кібератак є глобальне розповсюдження шкідливих програм, що можуть завдати шкоди працездатності комп'ютера, пошкодити важливі файли, або, навіть, видалити їх. Прикладами подібних програм є: DDoS, шкідливі програми (malware): руткит, «троянський кінь», мережеві віруси [6].

DDoS (Distributed Denial of Service) – розподілена атака типу «відмова в обслуговуванні». Мережевий ресурс виходить з ладу в результаті багатьох запитів до нього, відправлених із різних точок. Зазвичай, атака організовується за допомогою бот-нетів.

Кіберзлочинці заражають комп'ютери ні про що не підозрюючих користувачів Інтернету. Такі «зомбі» і відправляють беззмістовні запити на сервер жертви. Чітко спланована атака може вивести з ладу практично будь-який незахищений ресурс: від сайту до великого корпоративного порталу. Оброблюючи мільйони запитів, сервер починає спочатку «гальмувати», а потім і зовсім перестає працювати [7].

Фінансова індустрія слугує ціллю для DDoS-атак на рівні додатків і мережевого рівня вже давно. Атаки із використанням бот-нетів призводять до зниження часу відгуку веб-сайтів і заважають клієнтам отримувати доступ до своїх онлайн-гаманців та банківських операцій. Атаки також слугують відволікаючою тактикою злочинців, які шукають способи компроментувати конфіденційні дані, вчиняти шахрайство та вкрасти приватні і фінансові дані [8,9].

Вірус типу Hacker Defender можна назвати класичним прикладом руткитів. Він замасковується від брандмауера та відкриває доступ хакерам через мережу Інтернет. Це дозволяє злочинцям отримати доступ до паролів, особистих даних та встановлювати інші шкідливі програми [7].

«Троянський кінь» – вірусна програма, яка є частиною нешкідливої програми. Він не копіює себе, а залишається в системі непомічено і виконує потенційно небезпечні дії, наприклад, відкриття портів для доступу хакерам, шпигунство за користувачем, видалення даних, виведення з ладу обладнання, крадіжка номерів кредитних карток та паролів [10].

Мережеві віруси розповсюджуються по комп'ютерній мережі через мережеві протоколи TCP, FTP, HTTP, UDP і протоколи передачі електронної пошти. Найчастіше вони не міняють системні файли, не видаляють важливу користувачу інформацію, але переповнюють мережу трафіком, що призводить до зниження швидкості або повної зупинки системи [6].

Відносно установ і організацій використовуються цілеспрямовані (таргетовані) атаки.

Таргетовані атаки – це завчасно сплановані дії, націлені на конкретну установу [11]. Найчастіше кіберзлочинці, які займаються таргетованими атаками є професіоналами. Метою цілевих атак є пряме викрадання грошей чи інформації. Щоб здійснити цілеспрямовану атаку потрібно знати схему комп'ютерної мережі жертви, засоби захисту, щоб потім успішно їх обійти. «Точкою входу» часто стає інсайдерська інформація від нелояльних співробітників компанії.

Недостатній захист комп'ютерної мережі може викликати збільшення кількості кібератак та завдану шкоду від них. Найпростіший спосіб – використання складних паролів і зменшення кількості пристроїв із доступом до мережі Інтернет. Найкращого захисту можна досягти частим

аудитом безпеки, оновлення уразливого програмного забезпечення, і використання антивірусних програм із актуальними базами даних [12].

Щоб не допустити зараження обладнання потрібно притримуватися певних правил [13]:

- не встановлювати програми від невідомого джерела;
- використовувати надійні паролі;
- перед початком користування необхідно сканувати зовнішні носії інформації,;
- регулярно сканувати обладнання з допомогою антивірусів на наявність шкідливих програм;
- робити резервні копії цінних даних.

Щоб виявити таргетовані атаки можливо скористатися сигнатурним аналізом, брандмауером, евристичним аналізом, білим списком [14].

Для того, щоб провести сигнатурний аналіз потрібно мати файл, заражений вірусом. Після дослідження шкідливої програми аналітик матиме можливість зняти з неї сигнатуру (цифровий відбиток). Після занесення відбитку в базу, порівнюючи сигнатури, з’являється можливість перевіряти файли на наявність цього вірусу в обладнанні. Даний спосіб дає можливість досить точно знаходити заражені файли, але потрібно постійно оновлювати сигнатурні бази.

Евристичний аналіз дозволяє перевіряти код на наявність властивостей, що є характерними для відомих вірусів. Даний спосіб гарний тим, що не залежить від актуальності баз. Недоліком є те, що відомі антивіруси є у відкритому доступі і хакери можуть тестувати своє ПО для обходу цього захисту.

Ще одним способом захисту можна назвати брандмауер. Він дозволяє виявити цілеспрямовані атаки методом фільтрації трафіку. Недоліком цього методу можна назвати його чутливість, тобто велику кількість хибних повідомлень, серед яких може загубитися необхідне попередження про атаку [15,16].

Використовувати кожен із цих методів окремо не кращий варіант, тільки комплексний захист може дати найбільший результат.

## 1.2 Аналіз статистичної інформації кібератак на фінансові установи

За даними дослідження Positive Technologies [2], сім з десяти кібератак в 2017 році були здійснені з метою отримання прямої фінансової вигоди, наприклад, для виведення грошей з банківських рахунків жертви. Більш того, протягом останніх років спостерігається зростання кібератак саме на банківські установи (рис. 1.1).

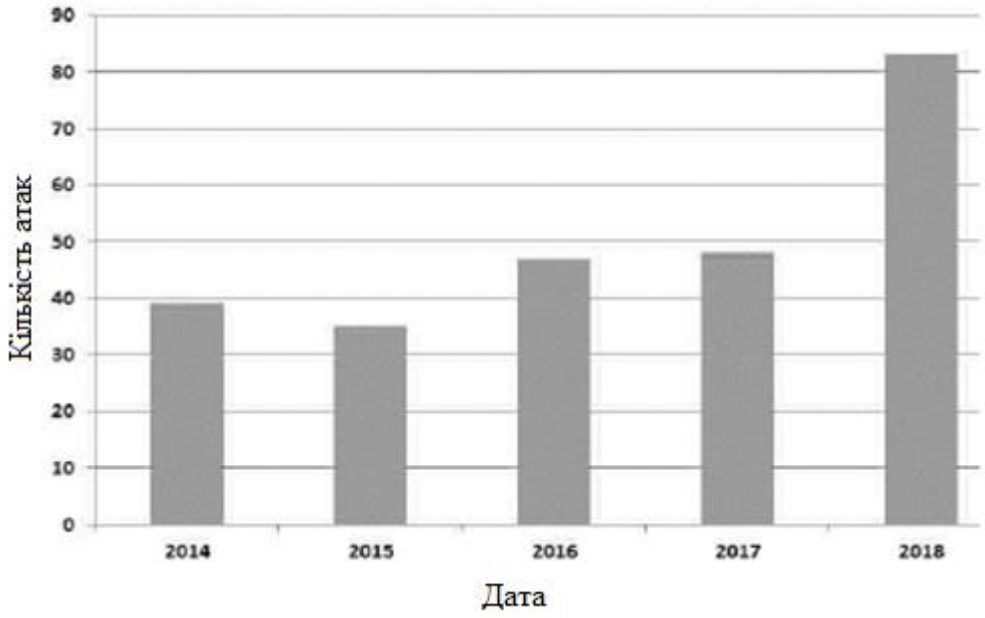



Рисунок 1.1 – Динаміка кібератак на банківські установи у світі, за даними [17]

У вересні 2019 року компанія Accenture представила результати дослідження [18], в якому виявила основні загрози інформаційній безпеці бізнесу в 2019 році. За оцінкою Accenture, ринок сервісів кібербезпеки

зростає темпами, аналогічними ринків Digital і IT. Accenture прогнозує, що до 2021 року обсяг світового ринку інтернет-бізнесу збільшиться на 66% і складе $202 млрд. При цьому сукупний світовий збиток від кібератак може вирости до 2021 року на 39% до $2,1 млрд.

Основні слабкі місця і недоліки механізмів захисту, які поширені на мережевому периметрі банків, можливо розділити на 4 категорії:

− уразливість веб-додатків (так звані SQL-ін’єкції - впровадження в запит до серверу SQL коду, міжсайтове виконання сценаріїв);

− недостатня мережева безпека (використання відкритих протоколів передачі даних, інтерфейси віддаленого доступу і управління доступом будь-якому інтернет користувачу);

− недоліки конфігурації серверів (використання старого ПЗ, зберігання важливих даних у відкритому вигляді);

− недоліки управління обліковими записами і паролями (використання словарних паролів).

Також потрібно враховувати, що наявність слабкого місця в периметрі системи ще не означає, що їх експлуатація дозволить потрапити у внутрішню мережу. В цілому, рівень захисту мережевого периметру в фінансовій сфері значно вищий, ніж у інших компаніях. За три роки у рамках внутрішнього тестування компанією Positive Technologies на проникнення доступ до внутрішньої мережі був отриманий у 58 % систем, а для банків цей показник склав лише 22 % [19]. У всіх випадках доступ був отриманий за допомогою недостатньому захисті веб-додатку. В одному банку було виявлено два вектори проникнення, причому обидва були основані на веб-додатку та недоліків конфігурації веб-серверу .

Таким чином можна зробити висновок, що злочинні угрупування, що планували б отримати доступ до внутрішньої мережі банку шляхом експлуатації уразливості на мережевому периметрі, мали б успіх у 22 %

банках. Схожі способи проникнення використовували у своїй діяльності, наприклад, угрупування ATMitch і Lazarus.

Указаний відсоток вразливості банківських установ може бути вищим. В рамках тестування не використовувалися уразливості, які можуть нанести шкоду інфраструктурі замовника. Наприклад, використання старого ПЗ в 67 % банків потенційно може дозволити подолати периметр, але використання цього способу може викликати відмову в обслуговуванні.

Більшість банків мають достатньо високий рівень захисту мережевого периметру, але персонал є найбільш уразливою ланкою в системі захисту будь-якої організації.

В процесі оцінки обізнаності в 75 % банків співробітники переходили по посиланню, зазначеного в фішинговому листі, в 25 % банків співробітники вводили свої облікові дані в помилкову форму аутентифікації, і ще в 25 % банків хоча б один співробітник запускав на своєму робочому комп'ютері шкідливе вкладення. В середньому в банках за фішинговим посиланням переходило близько 8 % користувачів, 2 % запускали вкладені файли, але свої облікові дані вводили менше 1 % користувачів.

Хоча рівень обізнаності у питаннях інформаційної безпеки серед банківських співробітників все ж вищий, ніж в інших галузях, але достатньо лише щоб один користувач виконав небажану дію, і кіберзлочинець отримає доступ до всієї корпоративної мережі. Таким чином, три чверті банків уразливі до атак методом соціальної інженерії, які використовуються майже всіма злочинними угрупуваннями в тому числі Cobalt, Lazaarus, Carbanak.

Поширеність способів кібератак на фінансові установи у світі за 2014/18 роки приведено на рис. 1.2.

Аналізуючи динаміку популярності різних типів атак за останні роки (рис. 1.2), можна зробити наступні висновки:

– найбільш поширеними типами кібератак на фінансові установи є атаки, спрямовані на знищення або заволодіння персональних даних клієнтів;

– останні рік-два спостерігається чітка тенденція, коли метою зловмисників є персональні дані клієнтів;

– останні три роки спостерігається збільшення кількості атак за допомогою malware.

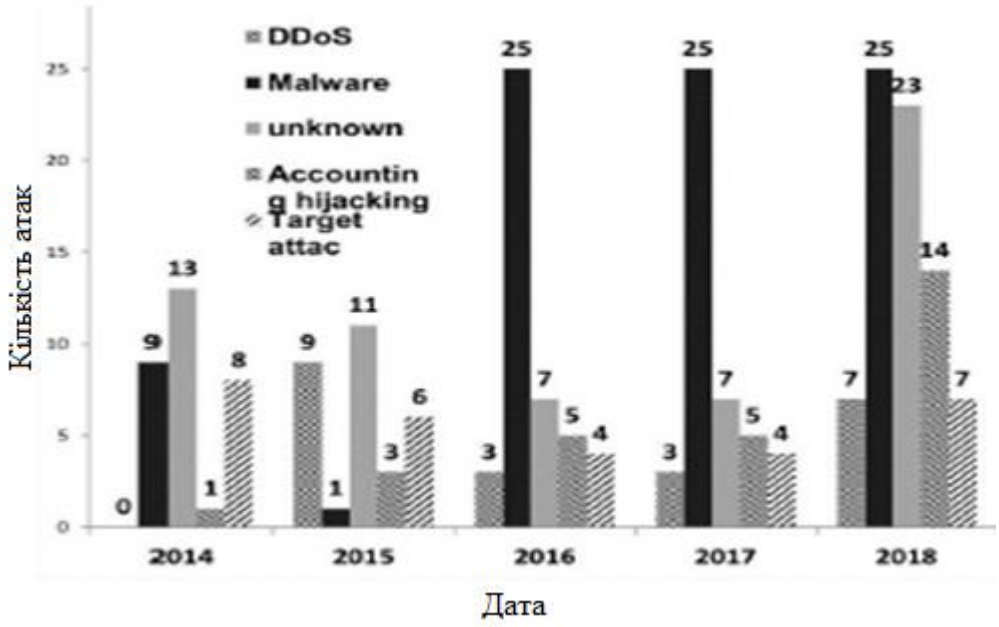



Рисунок 1.2. – Розподіл типів кібератак на банківські установи у світі по рокам, за даними [17]

Такі висновки дають можливість фахівцям з кібербезпеки обгрунтовано формувати стратегії та технології захисту від кібератак для забезпечення стабільної роботи, мінімізації втрат та збереження іміджу кредитно-фінансових установ.

## 1.3 Методи та моделі прогнозування кібератак

Для початку потрібно визначити, в чому різниця між методом і моделлю прогнозування.

Метод прогнозування складається з послідовності дій, які потрібно зробити для отримання моделі прогнозування.

Модель прогнозування – це формалізоване співвідношення, що адекватно описує досліджуваний процес і є базою для отримання прогнозованих значень.

Поєднання методу і моделі формує процес моделювання [20].

Недоліком більшості сучасних систем кібербезпеки є те, що при ідентифікації кібератак використовується уже відомі сигнатури або прототипи деяких процесів або подій. Ці способи корисні тільки для хакерів новачків, які можуть використовувати типові інструменти і прийоми для організації кібератак. Проти досвідчених злочинців ці системи є, як правило, малодієвими. Саме в цьому випадку корисним є направлення по прогнозування кількості кібератак. У випадку отримання прогнозу про те, що кількість кібератак буде вищим, ніж, наприклад, середній показник, то потрібно прийняти додаткові дії для забезпечення потрібного захисту (додатковий аналіз трафіку, обмеження доступу до інформації).

Всі сучасні методи прогнозування можна поділити на дві групи: імовірносні і точкові [21]. Точкові методи прогнозування дозволяють отримати чисельну інформацію в певний період часу, а імовірносні показують оцінку імовірності деякої події в майбутній момент часу, пов’язаний із вибраним показником. Прикладом точкового прогнозування можна назвати прогнозування курсу валюти на наступний день, а імовірносного – оцінка імовірності тих подій, що курс на наступний день буде вищим чи нищим ніж сьогодні.

В останній час все більше дослідників та науковців звертають свою увагу на імовірнісне прогнозування кібератак [22-26]. Це можна поясними тим, що імовірнісні прогнози дозволяють отримати результат не тільки саме майбутніх подій, а й оцінки їх здійснення. Різновидом імовірнісного прогнозування являється інтервальне прогнозування [27,28]. Головна ідея такого прогнозування полягає в тому, що прогнозується інтервал (із двох уже заданих інтервалів), у якому може знаходитись майбутнє значення показника на основі оцінок ймовірностей цих подій. Розмежувальна межа інтервалів задається розрахунковим способом, ґрунтуючись на статистичних характеристиках цього показника.

У праці «A Cyber Attack Modeling and Impact Assessment Component» [29], складено прототип CAMIAC, який допомагає побудувати атакуючий граф, що показує всі можливі послідовні дії зловмисника, що ведуть його до встановлених цілей, ці послідовні дії ще називають слідами атаки. Недоліком цього способу можна назвати його складність побудови, а для того, щоб ввести певні правки потрібно все потім робити спочатку.

Хоча прогнозування кібератак в основному спирається на дискретні моделі, існує безліч інших методів і моделей, що застосовуються для прогнозування кібератак, починаючи від дискретних моделей, наприклад, графів атак, до безперервних моделей, наприклад, часових рядів [30].

Наприклад, прогнозування може не починатися з уже спостереженої зловмисної події, а скоріше з імовірністю, що дана атака відбудеться, основуючись на вже відомих статистичних даних. Прикладом підходу, заснованого на безперервній моделі, є часові ряди, що представляють собою кількість атак на певну систему або мережу за певний проміжку часу. Потім часовий ряд може бути використаний для прогнозування того, чи станеться атака чи ні, або для прогнозування кількості можливих атак. Удосконалені методи можуть вирахувати типи нападів та характеристики нападників та жертв щоб фахівці з кібербезпеки могли оцінити, який тип

атаки буде найбільш ймовірним, хто стане нападником та хто має стати жертвою [31,32].

Також для прогнозування кібератак є можливість використовувати теорію ігор. Гра використовується як модель взаємодії між атакуючим (кіберзлочинцем) і захисником (адміністратором). Основні припущення в теорії ігор полягають в тому, що учасники раціональні (вони переслідують свої ролі) та міркують стратегічно (вони враховують свої знання чи сподівання інших учасників) [33].

Для кращої оцінки переваг і недоліків, методи наведені на таблиці 1.1.

Таблиця 1.1 – Порівняльна характеристика методів прогнозування

| Метод | Переваги | Недоліки |
|---|---|---|
| «A Cyber Attack Modeling and Impact Assessment Component» | Можливість визначення шляху (сліда) атаки | Складність побудови, після правок потрібно починати розрахунок спочатку |
| Дослідження часових рядів | Великий вибір способу побудови прогнозу, простота реалізації | Складний вибір моделі |
| Теорія ігор | Результат з високою точності | Кожен гравець повинен знати стратегію супротивника |

Отже, для побудови моделі прогнозування можливості кібератак було обрано дослідження часових рядів, через його простоту і способи побудови моделі.

# РОЗДІЛ 2. РОЗРОБКА МОДЕЛІ ПРОГНОЗУВАННЯ МОЖЛИВИХ ОБСЯГІВ КІБЕРАТАК ФІНАНСОВИХ УСТАНОВ

## 2.1 Моделі прогнозування часових рядів

Майже завжди ми зустрічаємося із явищами, які були б цікаві для вивчення їх змін і розвитку упродовж певного періоду часу. У повсякденному житті можуть представляти інтерес курси валют, зміна тиску і температури повітря, ціни на певний товар. Всі вони із плином часу змінюються. Сукупність змін однієї характеристики в часі являє собою часовий ряд.

Часовий ряд можна записати у короткому вигляді:

$$y_t, t = 1,2, \dots n, \tag{2.1}$$

де t – рівновіддалені моменти спостережень (година, доба, місяць рік).

Часовий ряд вірно відображає зміни певного показника, коли дані для нього відібрані за рівні проміжки часу, в одних одиницях виміру та мають достатню кількість спостережень.

При дослідженні часового ряду однією із цілей є визначення характеру і напрям змін даних у майбутньому, це дозволяє отримати нові значення, які будуть прогнозованими. О.В. Козьменко і О.В. Кузьменко пропонують визначення прогнозу як «науково обгрунтованого судження стосовно можливих станів об’єкта в майбутньому, альтернативні шляхи і терміни їх здійснення» [34, с.219].

Часовий ряд не повинен мати різких стрибків значень - такі значення створюють викривлення прогнозу і їх потрібно визначити і усунути.

Для виявлення аномальних значень в часовому ряді використовують методи, що аналізують статистичну сукупність (метод Ірвіна, модифікований метод Ірвіна).

Метод Ірвіна базується на порівнянні двох сусідніх значень і розрахунку значення λ, що дорівнює:

$$\lambda_t = \frac{|y_t - y_{t-1}|}{\hat{\sigma_y}}; t = 2,3, \dots, n; \quad (2.2)$$

де $\hat{\sigma_y}$ − оцінка середньоквадратичного відхилення вибіркового ряду $y_t$, яка розраховується з використанням формул:

$$\hat{\sigma}_y = \sqrt{\frac{\sum_{t=1}^{n}(y_t - y^-)^2}{n}}, \hat{y} = \frac{\sum_{t=1}^{n} y_t}{n}. \quad (2.3)$$

Розрахункові значення $\lambda_2$, $\lambda_3$, тощо порівнюють із критичним значенням $\lambda_\alpha$, і якщо вони не перевищують критичне, то відповідні рівні $y_t$, вважають нормальними. Критичні значення для рівня значущості $\alpha = 0,05$ (помилка 5 %) наведено в таблиці 2.1.

Таблиця 2.1 – Значення критерію Ірвіна

| n | 2 | 3 | 10 | 20 | 30 | 50 | 100 |
|---|---|---|---|---|---|---|---|
| $\lambda_\alpha$ | 2,8 | 2,3 | 1,6 | 1,3 | 1,2 | 1,1 | 1,0 |

Модифікований метод відрізняється від звичайного тим, що $\hat{\sigma_y}$ розраховується за трьома спостереженнями, а не всією сукупністю. Розраховувати можна для всього ряду, або тільки там, де підозрюється аномальність.

Середнє відхилення між двома сусідніми значеннями розраховується за формулою:

$$\overline{y_t} = \frac{y_{t-1} + y_{t+1}}{2} \tag{2.4}$$

$$\hat{\sigma}_y = \sqrt{\frac{\sum_{t=1}^{n}(y_{t-1} - \bar{y})^2 + (y_{t+1} - \bar{y})^2}{2}}. \tag{2.5}$$

Аномальні значення можна замінити на середнє між сусідніми рівнями ряду, коли причиною аномальності можна назвати помилки першого роду [35].

Моделі часових рядів можна розділити на дві групи: статистичні та структурні (рис. 2.1)

У статистичних моделях для прогнозування майбутніх значень використовується деякі рівняння залежності майбутніх значень від минулих.

У структурних моделях залежність прогнозованих даних від минулих задається певними правилами переходу.

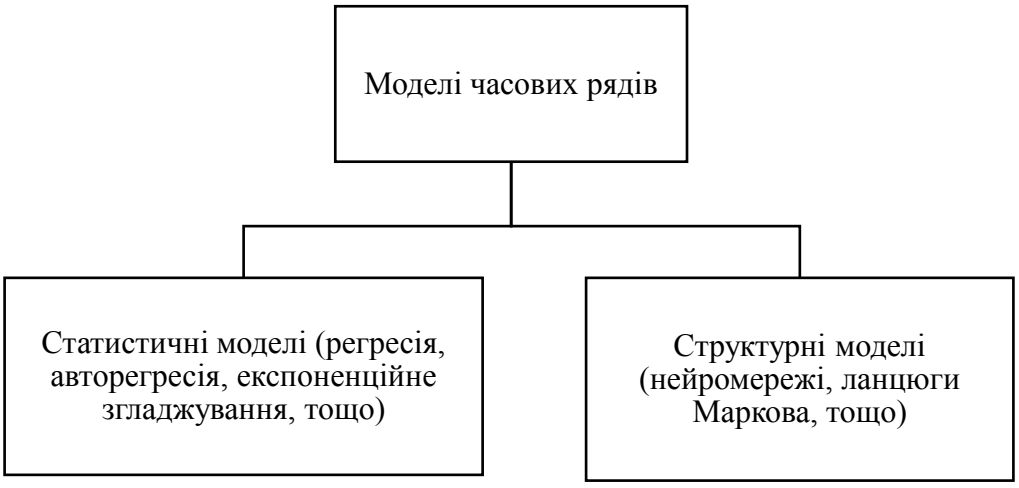



Рисунок 2.1 – Моделі часових рядів.

Основою даного дослідження є часовий ряд, то дуже ефективним методом можна назвати використання моделей авторегресії. Для побудови

та аналізу моделі прогнозу можна використовувати модель ARMA (авторегресійні моделі ковзної середньої). В основі лежать припущення лінійності процесу походження даних. ARMA описує стаціонарний процес, що складається з трьох ознак:

p– порядок авторегресії;

q – порядок ковзної середньої в моделі [36].

Необхідність в аналізі моделей з інтервенцією виникає, коли з деякого моменту різко змінюється поведінка ряду із зовнішніх причин. Зовнішній вплив на ряд може бути короткочасним (імпульсним) і довготривалим (стійким). В момент впливу траекторія різко змінюється, але далі знову описується моделлю ARMA. Існує можливість використовувати одночасно декілька різних інтервенцій (до 6).

До всіх перерваних рядів можуть бути побудовані прогнози, які можна вивести на графік (разом із початковим рядом) і, якщо потрібно, додати прогнози до початкового ряду. ARMA дозволяє аналізувати часові ряди із сезонністю [37].

До методів аналітичного згладжування входить регресійний аналіз в сукупності із МНК (методом найменших квадратів) та різними його модифікаціями. Методи прогнозування, що основуються на методах регресії, можуть бути використані для середньострокового і короткострокового прогнозування. Недоліком цього способу є відсутність адаптації – після отримання нових даних побудову нового прогнозу потрібно повторювати спочатку.

У наш час все частіше можна почути про системи штучного інтелекту, що базуються на використанні апарата штучних нейронних мереж. За допомогою них вирішуються велика кількість проблем: побудова моделі об’єктів при великій кількості шуму, недостатній кількості інформації, кластеризації [38].

Ідея нейронних мереж виникла в результаті спроб змоделювати поведінку живих істот, що відчувають вплив зовнішнього середовища і

навчаються на власному досвіді. Також штучні нейронні мережі використовуються і при прогнозуванні.

Існує безліч нейромережевих структур, що різняться кількістю і розміщенням нейронів і синаптичних зв’язків. Найвідомішою структурою є багатошаровий персептрон. Багатошаровий персептрон (MLP) являє собою повнозв’язну модель без зворотніх зв’язків. Кількість шарів і нейронів зазвичай обумовлено поставленою задачею. На рисунку 2.2 зображено схематичне розташування шарів і нейронів у персептроні.

Основою нейромереж слугує штучний нейрон, який імітує в першому наближенні властивості біологічного нейрону. На вхід штучного нейрону поступає велика кількість сигналів, які, в свою чергу, є виходами від інших нейронів. Потужність нейромережі залежить від кількості нейронів [39-41].

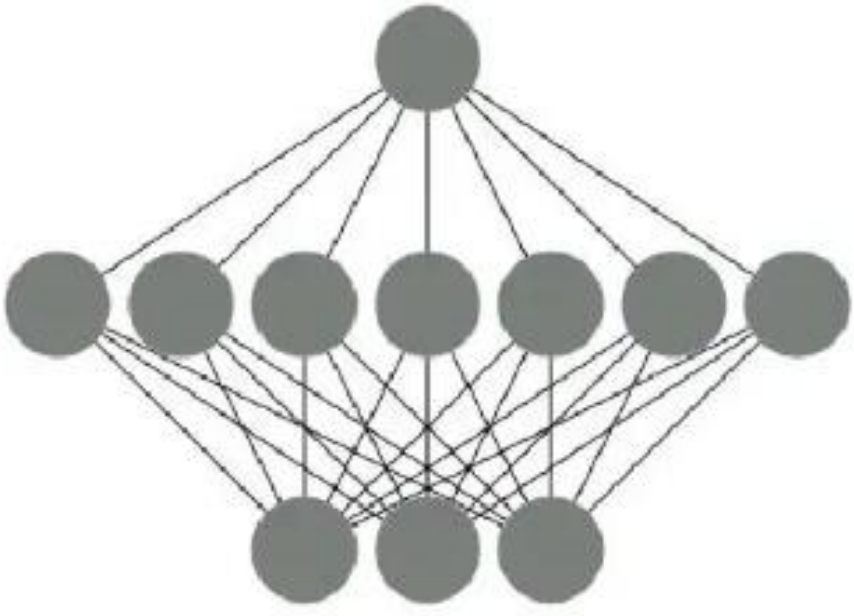


Рисунок 2.2. – Багатошаровий персептрон, що містить вхідний, вихідний та прихований шари

Скупчення нейронів можна поєднати в шари, на основі яких будуються одношарові або багатошарові нейромережі. Багатошарові

нейромережі дозволяють реалізувати більше можливостей, але вони можуть привести до збільшення потужності лише коли активаційна функція, що поєднує шари є не лінійною [42].

Якщо в часовому ряді спостерігається сезонність, то ще можна використати сезонну декомпозицію. Основна ідея сезонної декомпозиції проста. Часовий ряд складається із певних компонент: сезонна компонента, тренд, циклічна компонента, і випадкова , нерегулярна компонента. Різниця між циклічною і випадковою компонентою в тому, що остання має регулярну (сезонну) періодичність, тоді як циклічні фактори зазвичай мають більш довгий ефект, як причому, змінюється цикл від циклу.

В методі сезонної декомпозиції тренд і циклічну компоненту в основному поєднують в одну тренд-циклічну компоненту. Конкретні функціональні взаємозв'язки між цими компонентами бувають різноманітні. Але є можливість виділити два основних способу, за допомогою яких вони взаємодіють: адитивна та мультиплікативна. Складові адитивної моделі сумуються між собою, а в мультиплікативної перемножуються. Якщо існує інформація про існування циклічних факторів, що впливають на ряд, то потрібно використовувати оцінки для різних компонент для складання прогнозу майбутніх значень ряду.

Метод експоненціального згладжування входить до групи методів адаптивного прогнозування. Головною ідеєю цього методу можна назвати те, що кожен елемент часового ряду повинен бути згладженим за допомогою зваженої ковзної середньої, а її вага після кінця ряду зменшується [43].

Для розрахунку експоненційної середньої використовується наступна формула:

$$E_n = WY_n + (1 - W)E_{n-1} \tag{2.6}$$

де: n – параметр згладжування (0 < n < 1);

$E_n$ – експоненціальна середня в точці n.

Не всі способи прогнозування будуть однаково корисні при побудові короткострокового або довгострокового прогнозу. Для кращого візуального представлення використання моделі в залежності від терміну прогнозу можливо скористатися таблицею 2.2.

Таблиця 2.2 – Можливість використання моделі прогнозу.

| Метод | Тип прогнозу | | |
|---|---|---|---|
| | Короткостроковий | Середньостроковий | Довгостроковий |
| ARMA | + | +/- | - |
| Експоненціальне згладжування | + | + | - |
| Нейронні мережі | + | + | + |

Будь-яка модель будується для аналізу даних та прогнозування. При цьому точність прогнозів повинна бути як найвища.

## 2.2 Формування вимог до моделі

Для того, щоб побудувати модель прогнозу можливості кібератак, слід визначитися із вимогами до даної моделі. Основними вимогами можна назвати: адекватність, точність, простота.

Адекватність – це здатність моделі відповідати реальній системі і враховувати, найбільш важливі ознаки та якості. До початку побудови моделі оцінити адекватність важко, але є можливість орієнтуватися на раніше отриманий досвід у побудові моделей схожої тематики [44].

Точність показує, наскільки вірним є прогноз. Точність потрібно оцінювати по вже відомим даним, які не були задіяні при побудові моделі. Джерелом зменшення точності слугує неправильно підібрана модель та недостатня кількість спостережень.

Простота характеризується витратами обчислювальних ресурсів на побудову моделі прогнозу. Під обчислювальними ресурсами слід розуміти час і оперативну пам’ять. Чим вони менші, тим кращою буде модель, економічністю.

Розібравшись із загальними вимогами, необхідно перейти і до конкретних, ураховуючи специфіку прогнозування часових рядів.

До моделі прогнозування можливості кібератак на фінансові установи висуваються такі вимоги:

- результати, що будуть отримані після моделювання, повинні мати практичне застосування;
- модель повинна працювати швидко, забезпечуючи оперативне надання результатів;
- модель повинна бути адекватною;
- стійкість – при невеликій зміні вхідних значень, результат повинен незначно змінитися відповідно;
- вхідні дані повинні бути точними, повними і достовірними;

Таким чином, модель прогнозування можливості кібератак повинна задовільняти вищенаведеним вимогам для того, щоб результат був найточнішим та був корисним при обробці аналітиком із кібербезпеки.

## 2.3 Опис вхідних змінних

Вхідними даними для побудови моделі виступають статистичні дані, зібрані на сайті hackmageddon і представлені у додатку Б та додатку В.

Для початку було перевірено вхідні дані на однорідність модифікованим методом Ірвіна, результати перевірки наведені на рис 2.3.

| Модифікований метод Ірвіна | | | | | |
|---|---|---|---|---|---|
| Year | yi | yt | (yt-1-yt)^2 | (yt+1-yt)^2 | λ |
| 01.01.2014 | 3 | — | — | — | — |
| 01.02.2014 | 3 | 3,5 | 0,250 | 0,25 | 0 |
| 01.03.2014 | 4 | 2 | 1,000 | 1 | 0,10706 |
| 01.04.2014 | 1 | 2 | 4,000 | 4 | 0,32117 |
| 01.05.2014 | 0 | 2,5 | 2,250 | 2,25 | 0,10706 |
| 01.06.2014 | 4 | 2,5 | 6,250 | 6,25 | 0,42823 |
| 01.07.2014 | 5 | 3 | 1,000 | 1 | 0,10706 |
| 01.08.2014 | 2 | 5 | 0,000 | 0 | 0,32117 |
| 01.09.2014 | 5 | 3 | 1,000 | 1 | 0,32117 |
| 01.10.2014 | 4 | 3,5 | 2,250 | 2,25 | 0,10706 |
| 01.11.2014 | 2 | 5 | 1,000 | 1 | 0,21412 |
| 01.12.2014 | 6 | 2 | 0,000 | 0 | 0,42823 |
| 01.01.2015 | 2 | 5,5 | 0,250 | 0,25 | 0,42823 |
| 01.02.2015 | 5 | 1,5 | 0,250 | 0,25 | 0,32117 |
| 01.03.2015 | 1 | 3,5 | 2,250 | 2,25 | 0,42823 |
| 01.04.2015 | 2 | 2 | 1,000 | 1 | 0,10706 |
| 01.05.2015 | 3 | 2 | 0,000 | 0 | 0,10706 |
| 01.06.2015 | 2 | 3,5 | 0,250 | 0,25 | 0,10706 |
| 01.07.2015 | 4 | 2,5 | 0,250 | 0,25 | 0,21412 |
| 01.08.2015 | 3 | 3,5 | 0,250 | 0,25 | 0,10706 |
| 01.09.2015 | 3 | 2,5 | 0,250 | 0,25 | 0 |
| 01.10.2015 | 2 | 4,5 | 2,250 | 2,25 | 0,10706 |
| 01.11.2015 | 6 | 2 | 0,000 | 0 | 0,42823 |
| 01.12.2015 | 2 | 5,5 | 0,250 | 0,25 | 0,42823 |
| 01.01.2016 | 5 | 2 | 0,000 | 0 | 0,32117 |

Рисунок 2.3 – Фрагмент перевірка на однорідність

Отже, після розрахунку було виявлено, що λ розраховане є меншим від табличного значення (табл. 2.1), тому досліджуваний ряд є однорідним і використовувати даний часовий ряд для наступного аналізу можливо.

Також обов’язковою є перевірка на стаціонарність. Її було виконало за допомогою методу перевірки різниць середніх рівнів.

Вхідний часовий ряд був розподілений на дві однакові частини. Для кожної з цих частин було розраховано середні значення й дисперсії:

$$\bar{y}_{1(2)} = \frac{\sum_{t=1}^{n_1} y_t}{n} \tag{2.7}$$

де n – кількість значень у розподіленій частина;

y – складова ряду.

$$\hat{\sigma}^2_{1(2)} = \frac{\sum_{t=1}^{n_1}(y_t - \bar{y}_1)^2}{(n-1)} \quad (2.8)$$

Перевіряємо рівність (однорідність) обох частин ряду. Розраховуєм F- статистику за формулою:

$$F_{\text{розрах}} = \begin{Bmatrix} \hat{\sigma}_2^2/\hat{\sigma}_1^2, \text{якщо } \hat{\sigma}_2^2 > \hat{\sigma}_1^2 \\ \hat{\sigma}_1^2/\hat{\sigma}_2^2, \text{якщо } \hat{\sigma}_1^2 > \hat{\sigma}_2^2 \end{Bmatrix} \quad (2.9)$$

$F_{\text{розрах}} \geq F_\alpha$ – дисперсії не рівні між собою, ряд нестаціонарний;

$F_{\text{розрах}} < F_\alpha$ – дисперсії рівні між собою, ряд стаціонарний;

Результати наведені на рисунку 2.4. F-розраховене є меншим, ніж F-табличне, тому ряд можна вважати стаціонарним.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 01.12.2015 | 2 | 0,944 | | | 01.11.2018 | 9 | 14,224 | |
| 01.01.2016 | 5 | 4,115 | | | 01.12.2018 | 5 | 0,052 | |
| 01.02.2016 | 2 | 0,944 | | | 01.01.2019 | 5 | 0,052 | |
| 01.03.2016 | 2 | 0,944 | | | 01.02.2019 | 4 | 1,509 | |
| 01.04.2016 | 2 | 0,944 | | | 01.03.2019 | 6 | 0,595 | |
| 01.05.2016 | 2 | 0,944 | | | 01.04.2019 | 5 | 0,052 | |
| 01.06.2016 | 2 | 0,944 | | | 01.05.2019 | 4 | 1,509 | |
| 01.07.2016 | 2 | 0,944 | | | 01.06.2019 | 4 | 1,509 | |
| 01.08.2016 | 4 | 1,058 | | | 01.07.2019 | 2 | 10,424 | |
| 01.09.2016 | 2 | 0,944 | | | 01.08.2019 | 5 | 0,052 | |
| 01.10.2016 | 3 | 0,001 | | | 01.09.2019 | 5 | 0,052 | |
| 01.11.2016 | 4 | 1,058 | | | 01.10.2019 | 6 | 0,595 | |
| sum | 104,00 | 70,971 | | | sum | 183,00 | 115,315 | |
| n | 35 | | | | n | 35 | | |
| yn | 2,971 | | | | yn | 5,229 | | |
| σ1 | 2,087 | σ1^σ1 | 4,35722 | | σ2 | 3,392 | σ2^σ2 | 11,5031 |
| | | | | | | | | |
| F розрах | 2,640008231 | ряд стаціонарний | | | | | | |
| F табл | 3,127675601 | | | | | | | |

Рисунок 2.4 – Фрагмент перевірка на стаціонарність

Концептуальну модель прогнозування можливості кібератак, вхідні та вихідні дані можна представити у схематичному вигляді (рис. 2.5).


Рисунок 2.5 – Концептуальна модель прогнозування можливості кібератак способом ARMA

Алгоритм роботи нейронних мереж можна навести наступним способом:

- проводиться експоненційне згладжування ряду;
- обирається досліджувана змінна, кількість навчальної, контрольної, тестової підвибірок;
- визначається мінімальна і максимальна кількість нейронів у шарі;
- визначається кількість мереж для навчання і зберігання;
- здійснюється вибір кращої мережі і побудова прогнозу да допомогою неї.

# РОЗДІЛ 3. МОДЕЛЮВАННЯ МОЖЛИВИХ ОБСЯГІВ КІБЕРАТАК НА ФІНАНСОВІ УСТАНОВИ

## 3.1 Програмне забезпечення прогнозування часових рядів

Інформаційні технології використовуються будь-де, в тому числі і у сфері економіко-математичного моделювання, тож вибір програмних засобів для реалізації прогнозу є досить великим. Найкращими засобами для дослідження часових рядів можна назвати STATISTICA, MATLAB, пакет SAS (Statistical Analysis System), EViews [45].

STATISTICA – комплексний аналітичний інструмент, призначений для побудови точних прогнозів у будь-яких областях, використовуючи різні методи прогнозування. Система включає в себе модуль для обробки та аналізу часових рядів, який дозволяє побудувати прогноз без використання допоміжних факторів, тобто для прогнозування поведінки ряду на основі його власної історії. Модуль включає в себе найбільш ефективні та популярні методи для аналізу часових рядів: експоненційне згладжування, модель авторегресії і ковзного середнього, сезонна декомпозиція, спектральний аналіз Фур'є.

STATISTICA також має нейромережу, що місить в собі потужну збірку вбудованих інтелектуальних можливостей, які дозволяють вирішити реальні задачі, навіть користувачу, що ще не користувався нейронними мережами. В той самий час, досвідчений користувач зможе повністю керувати майже всіма аспектами нейромережевих структур і навчання.

При вивченні часових рядів часто дуже ефективними виявляються графічні і описові методи аналізу. Також модуль містить повний набір засобів для проведення будь-яких видозмін часового ряду, таких як взяття різниць різних порядків (вивчення мінливості ряду), згладжування ряду

(виявлення тенденцій в поведінці ряду), виділення тренду (виділення детермінованої систематичної складової ряду), обрахунок автокореляційних і кроскореляційних функцій, а також побудова їх графіків (корелограм) [46].

MATLAB – пакет прикладних програм що дозволяє провести багато різних розрахунків, в тому числі і з часовими рядами [47].

Інструменти MATLAB дозволяють отримувати доступ, візуалізувати і аналізувати історичні і поточні дані часових рядів, для того, щоб виявити і проаналізувати залежність. За допомогою пакету MATLAB у аналітика є можливість:

− отримувати доступ до даних із різних джерел (файли, електронні таблиці, бази даних);

− зберігати дані в об’єктах часових рядів, для того, щоб полегшити керування даними, обробки пропущених даних;

− виконувати технічний аналіз із різними фільтрами, стохастиками та індексами;

− створювати користувацькі процедури аналізу, візуалізації і анімації для демонстрації процесу аналізу;

− налаштування середовища аналізу із розширеною функціональністю для обробки сигналів, статистики чи економетрики [48].

Аналітик має можливість оцінити спектри часових рядів, які описують варіації часових рядів, використовуючи циклічні компоненти на різних частотах. Також є можливість проаналізувати авторегресію (AR), авторегресію з ковзним середнім (ARIMA) [49].

Пакет SAS (Statistical Analysis System) – комплексний статистичний пакет від компанії SAS Institute Inc. Основний додаток SAS – налаштовувана система Business Intelligence призначена для фінансово менеджменту, керування ризиками, маркетингу, прогнозування. Всі рішення базуються на загальній технологічній платформі (SAS Enterprise

Intelligence Platform), яка забезпечує базові функціональні можливості, необхідні всім додатків:

− ETL/ELT – процес добутку даних із різних джерел із наступною обробкою та очищенням;

− зберігання даних і спеціалізовуваному аналітичному банку даних;

− поглиблена аналітика – середовище для проведення поглибленого аналізу даних (data mining), описового і прогнозного моделювання, прогнозування часових рядів, оптимізації.

SAS включає в себе точні методи для невеликих наборів даних, потужні інструменти статистичного моделювання для задач із великими даними і сучасні методи аналізу даних, що містять відсутні значення [50]. Пакет SAS дозволяє:

− будувати дерева класифікації та регресії;

− відокремлювати дані на навчальні, контрольні і тестові ролі;

− використовувати сучасні методи підбору моделей, такі як еластична сітка і група LASSO;

Сотні вбудованих графіків і діаграм забезпечують подання чіткої та впорядкованої статистичної інформації, через це результати аналізу легко зрозуміти. Метадані зберігаються в централізованому сховищі, то є можливість включати моделі SAS/STAT в інші рішення SAS.

EViews – це сучасний пакет економетрики, статистики та прогнозування, який пропонує потужні аналітичні інструменти в гнучкому і зручному інтерфейсі. В порівнянні з конкурентами EViews немає модульної системи, але є доступним вікно робочого файлу, де є можливість зберігати ряд об’єктів. В додатку EViews для побудови моделі ARMA використовується розширений тест Дікі-Фулера, перевірка стаціонарності виконується автоматично після взяття різниць першого чи другого порядку. EViews включає в себе технологію електронних таблиць і реляційних баз даних із традиційними задачами, що використовуються в

статистичному програмному засобі і використовує графічний інтерфейс Windows [51,52].

Для точного уявлення переваг і недоліків було створено порівняльну таблицю 3.1.

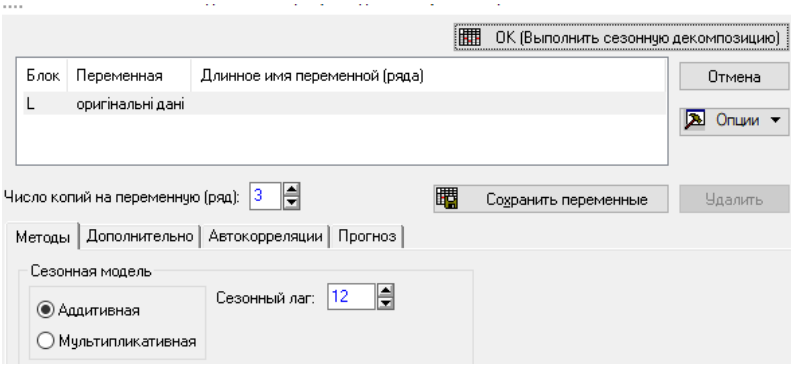
Таблиця 3.1 – Переваги і недоліки програм для побудови моделі

| ПЗ | Переваги | Недоліки |
|---|---|---|
| STATISTICA | - Можливість паралельної роботи в різних модулях;<br>- велика кількість довідкової літератури;<br>- зрозумілий інтерфейс;<br>- швидкодія;<br>- легкий імпорт/експорт даних в електронні і текстові процесори. | - Складно опанувати для не фахівця в області математичної статистики;<br>- висока ціна. |
| MATLAB | - Зручний інтерфейс;<br>- простота в роботі. | - Дорога ліцензія;<br>- заплутана інтеграція із JAVA додатками. |
| SAS | -Швидке оброблення дуже великих обсягів даних;<br>- можливість перетворювати формули у програмний код; - створення користувацьких модулів. | - Дорога ліцензія ;<br>- складність опанування. |
| EViews | - Швидкодія;<br>- можливість роботи із декількома файлами одночасно;<br>- великий вибір сучасних методів для обробки даних. | - Відсутність україномовної або русифікованої версії;<br>- невелика кількість україномовної літератури. |

Розглядаючи вищенаведені переваги і недоліки для прогнозування можливості кібератак було обрано пакет STATISTICA.

## 3.2 Проведення розрахунків

Перш, ніж почати до побудови моделі прогнозу потрібно проаналізувати часовий ряд на наявність сезонності. Якщо сезонність буде виявлена, то її потрібно буде врахувати при побудові моделі. Для цього використаємо сезонну декомпозицію. Для початку потрібно вибрати модель адитивна чи мультиплікативна. Знаючи їх різницю, обираємо адитивну модель, адже якщо сезонні коливання і присутні, то амплітуда їх буде однаковою. Результат сезонної декомпозиції зображено на рис. 3.1.




Сезонная декомпозиция: Аддитивн.сезон. (12) (Таблица данных1)
ПЕР1

| Набл | ПЕР1 | Скольз. средние | Разности | Сезонные составл. | Скоррек. ряд | Сглажен. тренд-ц. | Нерег. компон. |
|---|---|---|---|---|---|---|---|
| 1 | 3,000000 | | | 0,53576 | 2,464236 | 3,163310 | -0,69907 |
| 2 | 3,000000 | | | 0,56910 | 2,430903 | 2,925347 | -0,49444 |
| 3 | 4,000000 | | | 0,11910 | 3,880903 | 2,449421 | 1,43148 |
| 4 | 1,000000 | | | -0,48090 | 1,480903 | 2,319329 | -0,83843 |
| 5 | 0,000000 | | | -0,11424 | 0,114236 | 2,694792 | -2,58056 |
| 6 | 4,000000 | | | -2,01007 | 6,010069 | 3,568403 | 2,44167 |
| 7 | 5,000000 | 3,250000 | 1,75000 | -0,04757 | 5,047569 | 4,104051 | 0,94352 |
| 8 | 2,000000 | 3,166667 | -1,16667 | -0,28090 | 2,280903 | 4,119329 | -1,83843 |
| 9 | 5,000000 | 3,333333 | 1,66667 | -0,09757 | 5,097569 | 3,729051 | 1,36852 |
| 10 | 4,000000 | 3,083333 | 0,91667 | 0,06910 | 3,930903 | 3,480903 | 0,45000 |
| 11 | 2,000000 | 3,166667 | -1,16667 | 1,20243 | 0,797569 | 3,082755 | -2,28519 |
| 12 | 6,000000 | 3,416667 | 2,58333 | 0,53576 | 5,464236 | 3,253125 | 2,21111 |
| 13 | 2,000000 | 3,250000 | -1,25000 | 0,53576 | 1,464236 | 2,873495 | -1,40926 |
| 14 | 5,000000 | 3,166667 | 1,83333 | 0,56910 | 4,430903 | 2,880903 | 1,55000 |
| 15 | 1,000000 | 3,250000 | -2,25000 | 0,11910 | 0,880903 | 2,338310 | -1,45741 |
| 16 | 2,000000 | 3,083333 | -1,08333 | -0,48090 | 2,480903 | 2,652662 | -0,17176 |
| 17 | 3,000000 | 2,916667 | 0,08333 | -0,11424 | 3,114236 | 3,028125 | 0,08611 |
| 18 | 2,000000 | 3,250000 | -1,25000 | -2,01007 | 4,010069 | 3,568403 | 0,44167 |
| 19 | 4,000000 | 2,916667 | 1,08333 | -0,04757 | 4,047569 | 3,659606 | 0,38796 |
| 20 | 3,000000 | 3,166667 | -0,16667 | -0,28090 | 3,280903 | 3,341551 | -0,06065 |
| 21 | 3,000000 | 2,916667 | 0,08333 | -0,09757 | 3,097569 | 3,173495 | -0,07593 |
| 22 | 2,000000 | 3,000000 | -1,00000 | 0,06910 | 1,930903 | 2,925347 | -0,99444 |
| 23 | 6,000000 | 3,000000 | 3,00000 | 1,20243 | 4,797569 | 3,193866 | 1,60370 |
| 24 | 2,000000 | 2,916667 | -0,91667 | 0,53576 | 1,464236 | 2,919792 | -1,45556 |
| 25 | 5,000000 | 2,916667 | 2,08333 | 0,53576 | 4,464236 | 2,873495 | 1,59074 |
| 26 | 2,000000 | 2,750000 | -0,75000 | 0,56910 | 1,430903 | 2,325347 | -0,89444 |
| 27 | 2,000000 | 2,833333 | -0,83333 | 0,11910 | 1,880903 | 2,227199 | -0,34630 |

Рисунок 3.1 – Результат класичної сезонної декомпозиції

В такий спосіб була знайдена сезонна компонента. Побудуємо графік сезонної компоненти (рис. 3.2).

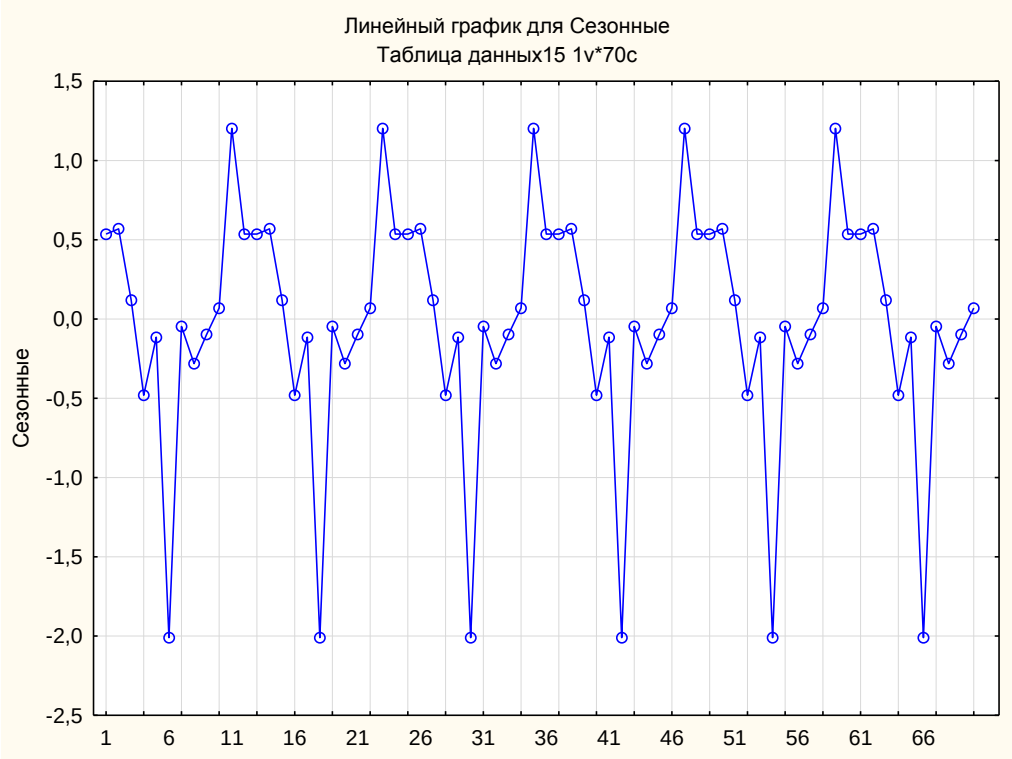



Рисунок 3.2 – Графік сезонної компоненти досліджуваного ряду

На рис. 3.2 можна побачити, що кожні 12 місяців спостерігаються піки атак, отже, потрібно їх урахувати при побудові прогнозу.

Для того, щоб зменшити вплив випадкової компоненти у часовому ряді, його можна згладити. Ця процедура допоможе отримати більш «чисті» дані, що складаються лише з детермінованих компонент. Для цього скористаємося експоненційним згладжуванням. Налаштування моделі та результат згладжування наведено на рис. 3.3 та рис. 3.4.

Поиск параметров на сетке (наименьшие абс.ошибки выделяю
Модель: Без тр., адд.сезон. (12); S0=4,017
оригінальні дані

| Модель Номер | Альфа | Дельта | Средняя ошибка | Ср. абс. ошибка | Сумма квадрат. | Средняя квадрат. |
|---|---|---|---|---|---|---|
| 19 | 0,300000 | 0,100000 | 0,054846 | 1,537177 | 238,8102 | 3,411575 |
| 10 | 0,200000 | 0,100000 | 0,080180 | 1,515967 | 239,7787 | 3,425410 |
| 28 | 0,400000 | 0,100000 | 0,045083 | 1,561029 | 245,5725 | 3,508179 |
| 20 | 0,300000 | 0,200000 | 0,053471 | 1,579664 | 253,0206 | 3,614580 |
| 11 | 0,200000 | 0,200000 | 0,076538 | 1,558182 | 254,8233 | 3,640332 |
| 37 | 0,500000 | 0,100000 | 0,039822 | 1,587282 | 257,4336 | 3,677623 |
| 1 | 0,100000 | 0,100000 | 0,172725 | 1,492956 | 258,3952 | 3,691360 |
| 29 | 0,400000 | 0,200000 | 0,044204 | 1,598319 | 258,8609 | 3,698014 |
| 21 | 0,300000 | 0,300000 | 0,052812 | 1,622073 | 268,1791 | 3,831130 |
| 38 | 0,500000 | 0,200000 | 0,039326 | 1,618021 | 269,6989 | 3,852842 |

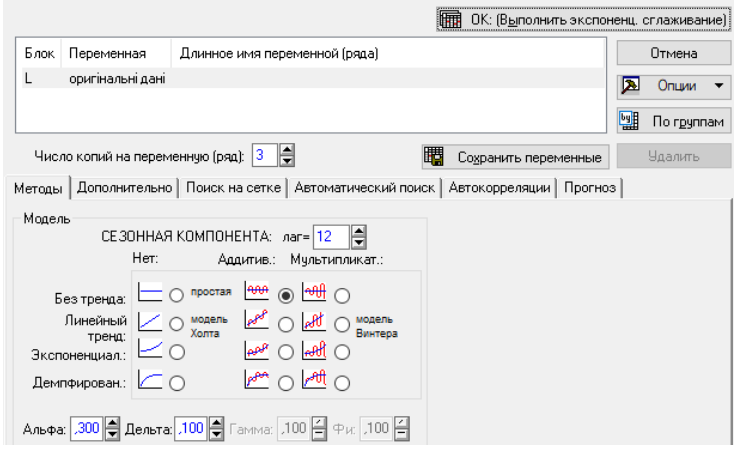



Рисунок 3.3 – Налаштування експоненційного згладжування

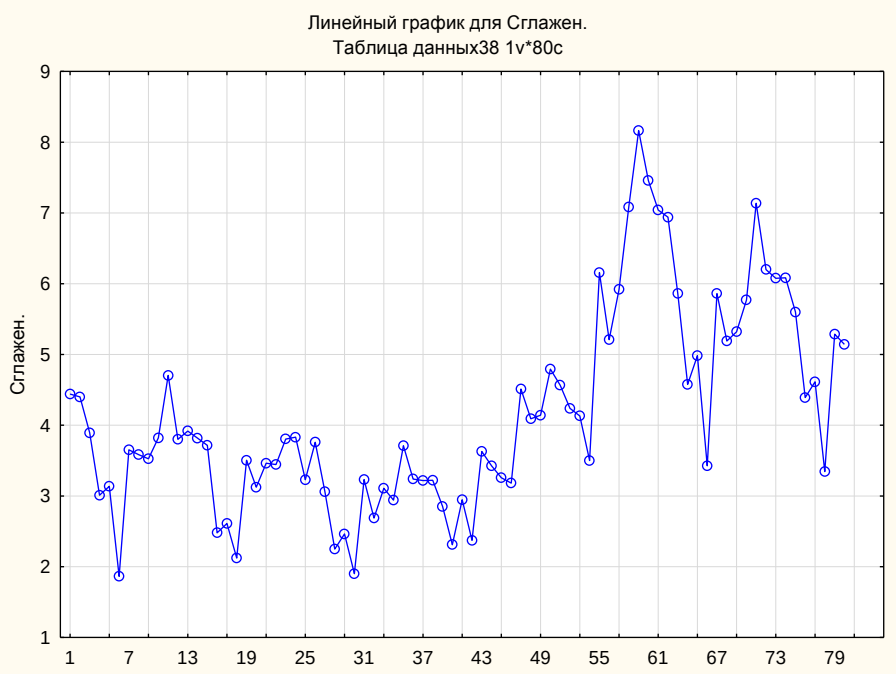



Рисунок 3.4 – Згладжений ряд

Будуємо прогноз за допомогою ARMA. Для цього підбираємо

p – порядок авторегресії,

q – порядок ковзної середньої в моделі, для того, щоб отримати найточніший результат.

При підборі показників p і q найкращий результат був досягнутий при p = 1, q = 0, налаштування моделі представлено на рисунку 3.5.

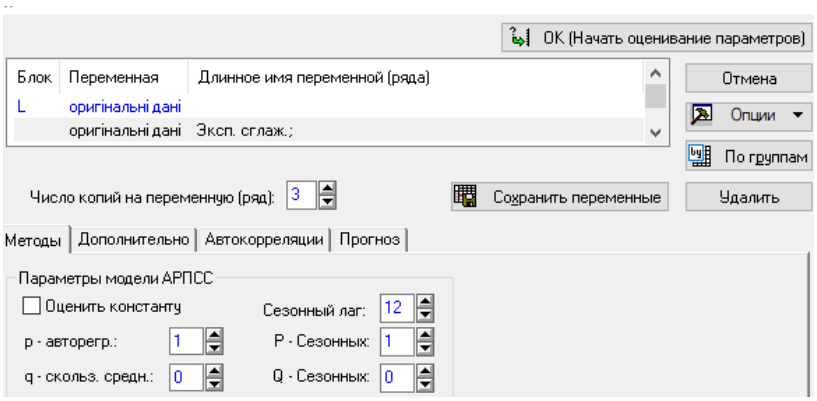



Рисунок 3.5 – Налаштування моделі ARMA

Оцінку параметрів зображено на рис. 3.6. Налаштування прогнозу зображено на рис. 3.7. Фрагмент розрахованих прогнозних значень зображено на рис. 3.8. Параметри виділені червоним кольором, це означає, що вони є статистично значущі, статистична помилка невисока.

Исход.:Зглажений ряд (Таблица данных2)
Преобразования: Нет
Модель(1,0,0)(1,0,0) Сезонный лаг: 12 MS Остаток= ,73737

| Параметр | Парам. | Асимпт. Ст.ошиб. | Асимпт. t( 68) | p | Нижняя 95% дов. | Верхняя 95% дов. |
|---|---|---|---|---|---|---|
| **p(1)** | 0,950314 | 0,041898 | 22,68155 | 0,000000 | 0,866707 | 1,033920 |
| Ps(1) | 0,564772 | 0,125678 | 4,49380 | 0,000028 | 0,313986 | 0,815559 |

Рисунок 3.6 – Оцінка параметрів ARMA

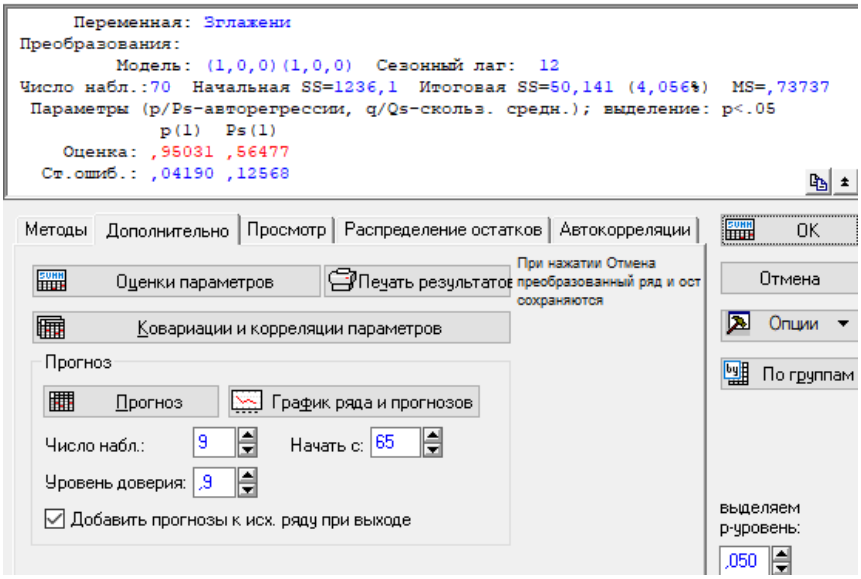



Рисунок 3.7 – Налаштування побудови прогнозу

Прогнозы; Модель:(1,0,0)(1,0,0) Сезонный лаг: 12 (Таблица данн
Исход.:Зглажений ряд
Начало исходных: 1 Конец исходн.: 64

| Набл. N | Прогноз | Нижний 90,0000% | Верхний 90,0000% | Ст.ошиб. | Наблюд. | Остатки |
|---|---|---|---|---|---|---|
| 65 | 4,408868 | 2,947514 | 5,870222 | 0,876336 | 4,983648 | 0,574781 |
| 66 | 3,950780 | 1,934802 | 5,966759 | 1,208930 | 3,426621 | -0,524160 |
| 67 | 5,351185 | 2,941644 | 7,760727 | 1,444940 | 5,859817 | 0,508632 |
| 68 | 4,723577 | 2,007177 | 7,439978 | 1,628955 | 5,189958 | 0,466381 |
| 69 | 5,036020 | 2,069651 | 8,002390 | 1,778855 | 5,319739 | 0,283718 |
| 70 | 5,609330 | 2,434081 | 8,784580 | 1,904115 | 5,769725 | 0,160395 |
| 71 | 6,139712 | 2,786988 | 9,492436 | 2,010542 | | |
| 72 | 5,666017 | 2,160729 | 9,171305 | 2,102030 | | |

Рисунок 3.8 – Фрагмент прогнозу за допомогою ARMA

При виборі архітектури нейронної мережі орієнтувалися на значення продуктивності навчання, контрольної і тестової продуктивності. Найкращі результати були у нейромережах з арітектурами 1-19-1. Трішки гірші у 1-15-1, тому для побудови прогнозу вона не використовувалася.

Наступним етапом є прогнозування за допомогою нейромережі. Налаштування та найкращі мережі зображено на рис. 3.9.

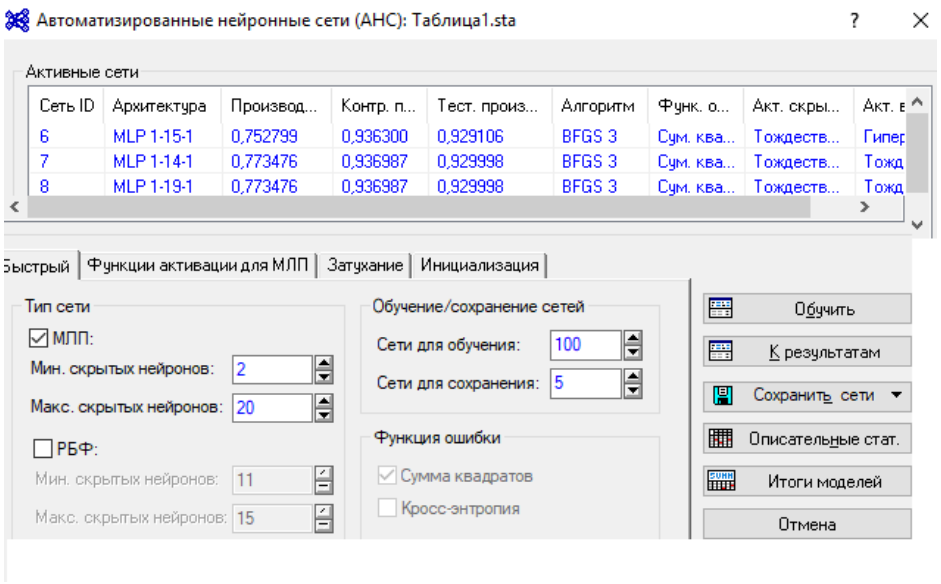



Рисунок 3.9 – Найкращі нейромережі

Для навчальної підвибірки було обрано 70%, для контрольної 15%, для тестової 15% від загальної кількості часового ряду.

Результат прогнозу за допомогою нейромережі зображено на рисунку 3.10.

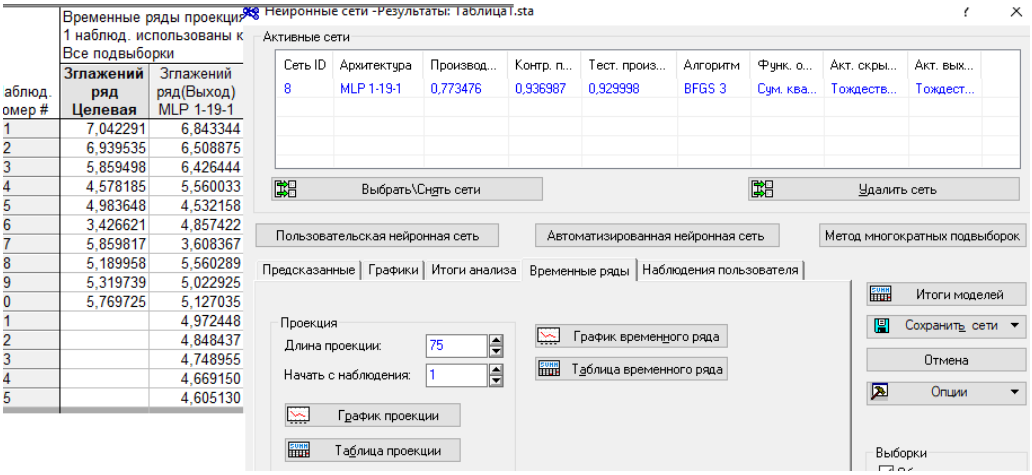



Рисунок 3.10 – Результат прознозування за допомогою нейромережі

Таким чином, на рис. 3.11 та рис. 3.12 можна побачити порівняння результатів прогнозування різними способами.

| | 1 оригінальні дані | 2 MLP 1-19-1 | 3 Згладжений ряд | 4 APMA |
|---|---|---|---|---|
| 01.05.2019 | 4 | 4,542131 | 4,578185 | 4,408868 |
| 01.06.2019 | 4 | 4,871446 | 4,983648 | 3,950780 |
| 01.07.2019 | 3 | 3,606836 | 3,426621 | 5,351185 |
| 01.08.2019 | 5 | 5,583066 | 5,859817 | 4,723577 |
| 01.09.2019 | 5 | 5,039010 | 5,189958 | 5,036020 |
| 01.10.2019 | 6 | 5,144417 | 5,319739 | 5,609330 |

Рисунок 3.11 – Порівняння вхідних даних із прогнозом нейромережею

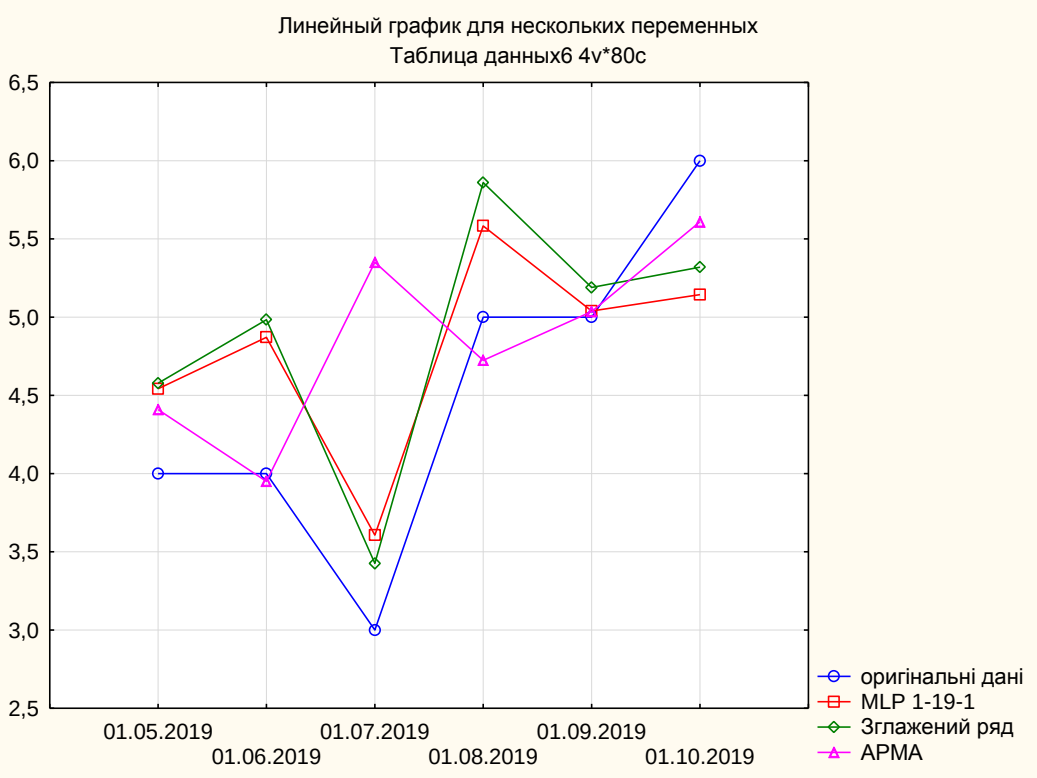



Рисунок 3.12 – Порівняння моделей прогнозування

Отже, після візуального огляду прогнозу можна стверджувати, що модель досить точно відображає можливість кібератак на фінансові установи.

## 3.3 Оцінка точності отриманих результатів

Важливим моментом прогнозування є оцінка точності прогнозу. На етапі перевірки використовують декілька критеріїв, що дають можливість поставити оцінку побудованому прогнозу [53].

Найбільш поширеними методами верифікації можна назвати показники, за допомогою яких є можливість кількісно визначити розмір помилки прогнозу у відсотках чи числовому показнику. До цих показників відносяться: абсолютна помилка прогнозу, середньоквадратична помилка, середня відносна помилка прогнозу.

Абсолютна помилка прогнозу розраховується як різниця між поточним значенням і прогнозом :

$$\Delta_{ПР} = y_{-t} - y^* \tag{3.1}$$

де $\Delta_{np}$ – абсолютна помилка прогнозу;

$y_t$ – поточне значення;

$y*$ – прогнозне значення.

Середнє абсолютне значення помилки розраховується:

$$\bar{\Delta}_{Пр} = \frac{\sum|y_t - y_t^*|}{n} \tag{3.2}$$

де $\overline{\Delta}_{np}$ – абсолютна помилка прогнозу;

$n$– кількість спостережень.

Середньоквадратична помилка прогнозу визначається за формулою:

$$\sigma_t = \sqrt{\frac{\sum(y_\tau - y_t^*)^2}{n}} \tag{3.3}$$

де $\sigma$– абсолютна помилка прогнозу.

А середня відносна помилка розраховується наступним чином:

$$\bar{\varepsilon}_{Пр} = \frac{\sum \frac{1^{|y_t - y_t^*|}}{y_t}}{n} \cdot 100 \quad (3.4)$$

де $\bar{\varepsilon}_{пр}$ – абсолютна помилка прогнозу.

Визначення значень середніх відносних помилок наведена у таблиці 3.2.

Таблиця 3.2 – Визначення значень середніх відносних помилок

| Значення | Пояснення |
|---|---|
| <10 | Висока точність |
| 10–20 | Середня точність |
| 20–40 | Задовільна точність |
| >40 | Погана точність |

Також для оцінки імовірності отриманих результатів можна використати формулу:

$$p = 1 - (\sum \frac{|\text{Поточне} - \text{Прогнозоване}|}{\text{Поточне}})/n \quad (3.5)$$

Таблиця 3.3 – Порівняльна оцінка точності прогнозу і його імовірності

| | Нейромережа | ARMA |
|---|---|---|
| Середньоквадратична помилка | 0,63 | 0,30 |
| Середня відносна помилка | 7,52 | 3,62 |
| Імовірність | 0,92 | 0,96 |

Отже, із таблиці 3.3 бачимо, що найкращі показники у ARMA, тому можемо вважати, що модель, побудована правильно і дає імовірність здійснення прогнозу в короткостроковій перспективі 96 %. У підсумку,

моделі дають схожі результати, які близькі до реальних даних. Це у свою чергу також свідчить про адекватність моделей та коректність їх використання.

# ВИСНОВКИ

Під час виконання дипломної роботи були виконані наступні завдання:

− розглянуто напрямки кібератак і способи їх реалізації. Визначено способи захисту від них. Визначено, що в останній час найпопулярнішим способом атак є malware.

− проаналізовано методи і моделі кібератак. Для побудови було обрано часові ряди.

− сформовано основні вимоги, які повинна мати модель.

− розглянуто наступне програмне забезпечення для побудови моделі: STATISTICA, MATLAB, пакет SAS, EViews. Для побудови моделі було обрано пакет STATISTICA через його простоту та можливості.

Для створення прогнозної моделі було обрано наступні:

– модель ARMA;

– нейронні мережі.

Для підвищення точності та якості прогнозу було додатково було обрано методи сезонної декомпозиції та згладження для обробки початкових даних.

Другий розділ дипломної роботи містить детальний аналіз процесу побудови моделей. Зокрема, в даному розділі проведений математичний опис моделей, використаних при моделюванні прогнозів. Другий розділ було завершено описом підготовкою часового ряду для побудови моделей.

В третьому розділі роботи була описана реалізація моделей прогнозування можливості кібератак на фінансові установи. Був представлений алгоритм побудови моделі прогнозування та описані його етапи. Після побудови моделей була проаналізована їх якість на даних, що не брали участь у прогнозуванні. Найкращі результати показала модель ARMA. Усі обрані моделі виявилися придатними для застосування, тому

всі вони були використані для прогнозу майбутніх значень. В кінці третього розділу був побудований прогноз обсягів продажів на необхідний період. Візуальний огляд показав, що прогнозні значення досить якісно описують тенденцію зміни кількості кібератак. Отже, даним прогнозом була досягнена мета дослідження дипломної роботи.

# СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Про основні засади забезпечення кібербезпеки України: Закон України від 05.10.2019. URL: https://zakon.rada.gov.ua/laws/show/2163-19/ed20171005#n10 (дата звернення 10.12.2019)

2. Захищеність кредитно-фінансової сфери, результати 2018 року / Positive Technologies. URL: https://www.ptsecurity.com/ru-ru/research/analytics/credit-and-financial-security-2019/?sphrase_id=67849 (дата звернення 10.12.2019)

3. Кібербезпека – 2018 – 2019: висновки і прогнози. URL: https://www.ptsecurity.com/ru-ru/research/analytics/cybersecurity-2018-2019/?sphrase_id=67744 (дата звернення 10.12.2019)

4. ATM malware is being sold on Darknet market. URL: https://securelist.com/atm-malware-is-being-sold-on-darknet-market/81871/ (дата звернення 10.12.2019)

5. Secret Service Warns of Sophisticated ATM Jackpotting Attack. URL: https://www.secretservice.gov/data/press/releases/2018/18-JAN/GPA_01-18_ATM_Jackpotting_Attack.pdf (дата звернення 10.12.2019)

6. Віруси і шкідливі програми / TREND MICRO 2017. URL: http://docs.trendmicro.com/ru-ru/smb/worry-free-business-security-90-sp1-agent-help/about/understanding-threat/viruses-and-malware.aspx (дата звернення 10.12.2019)

7. Що таке руткити. Програми для видалення руткитів. / Налаштування і оптимізація операційних систем. URL: https://www.windxp.com.ru/rotdel.htm – Назва з екрана.

8. Protecting Financial Institutions from DDoS Attacks URL: https://www.imperva.com/blog/protecting-financial-institutions-from-ddos-attacks/ (дата звернення 10.12.2019)

9. Keunsoo, L. DDoS attack detection method using cluster analysis [Text] / L. Keunsoo, J. Kim, K. Hoon Kwon, Y. Han, S. Kim // Expert Systems with Applications. – 2009. – Vol. 4, Issue 3. – p. 1659–1665.

10. Raiyn, J. A survey of Cyber Attack Detection Strategies [Text] / J. Raiyn // International Journal of Security and Its Applications. – 2014. – Vol. 8, Issue 1 –P. 247–256. doi: 10.14257/ijsia.2014.8.1.23

11. Таргетовані атаки: нове слово світі загроз / High-Tech Club. URL: https://www.kv.by/content/340248-targetirovannye-ataki-novoe-slovo-v-mire-ugroz (дата звернення 10.12.2019)

12. Дудикевич, В. Б. Проблеми оцінки ефективності систем захисту [Текст] / В. Б. Дудикевич, І. А. Прокопишин, В. Ф. Чекурін // Вісник Національного університету «Львівська політехніка». Сер.: Автоматика, вимірювання та керування. – 2012. – № 741. – С. 118–122.

13. Бурячок, В. Л. Інформаційна та кібербезпека: соціотехнічний аспект: підручник / [В. Л. Бурячок, В. Б. Толубко, В. О. Хорошко, С. В. Толюпа]; За заг. ред. В. Б. Толубка. - К.: ДУТ, 2015. - 288 с.

14. Targeted Attacks / Trend micro .URL: https://www.trendmicro.com/vinfo/us/security/definition/targeted-attacks (дата звернення 10.12.2019)

15. Варлатая С.К. Защита информационных процессов в компьютерных сетях / С.К. Варлатая, М.В. Шаханова. – М.: Проспект, 2015. – 216 с.

16. Шаньгин В.Ф. Защита информации в компьютерных системах и сетях / В.Ф. Шаньгин. – М.: ДМК Пресс, 2012. – 592 с.

17. Сайт терміни та статистика безпеки інформації: HACKMAGEDDON. URL: https://www.hackmageddon.com/ (дата звернення 10.12.2019)

18. Accenture: тренди кібербезпеки 2019. URL: https://www.accenture.com/ru-ru/company-news-release-cybersecurity-trends-2019 (дата звернення 10.12.2019)

19. Вектори атак на банки / Positive Technologies URL: https://www.ptsecurity.com/ru-ru/research/analytics/banks-attacks-2018/ (дата звернення 10.12.2019)

20. Armstrong, J.S. Forecasting for Marketing Quantitative Methods in Marketing [Text] / J. S. Armstrong. – London : International Thompson Business Press. – 1999. – P. 92 – 119.

21. Elliott G. Handbook of Economic Forecasting / G. Elliott, C Granger , A. Timmermann — 2013. Vol 2.

22. Kim Y. A probabilistic approach to estimate the damage propagation of cyber attacks / Y. Kim, T. Lee, H. In, Y. Chung, I. Kim, D. Baik // Proceeding of the 8th International Conference on Information Security and Cryptology. – 2005. – P. 175–185.

23. Wu J. Cyber Attacks Prediction Model Based on Bayesian Network / J. Wu // Proceedings of the 2012 IEEE 18th International Conference on Parallel and Distributed Systems. – 2012. – P. 730–731.

24. Haitao D. Probabilistic Modeling and Inference for Obfuscated Network Attack Sequences URL: http://scholarworks.rit.edu/theses/8331 (дата звернення 10.12.2019)

25. Zhan Z. Predicting cyber attack rates with extreme values / Z. Zhan, M. Xu, S. Xu. // IEEE Transactions on Information Forensics and Security. – 2015. – №. 10 (8).

26. Werner G. Time series forecasting of cyber attack intensity / G. Werner, S. Yang, K. McConky // Proceedings of the 12th Annual Conference on Cyber and Information Security. – 2017. – P. 224–240.

27. Krakovsky Y. M Applied aspects of application of interval forecasting of dynamic indicators in system analysis / Y. M. Krakovsky , A. N. Luzgin –Modern technology. System analysis. Modeling, 2017, no. 2/

28. Краковский Ю. М Программное обеспечения интервального прогнозирования нестационарных динамических показателей / Ю.М.

Краковский, А.Н. Лузгин // Вестник ИрГТУ. – 2015. – Т.1. – № 4. – С.12–16.

29. 2013 5th International Conference on Cyber Conflict K. Podins, J. Stinissen, M. Maybaum (Eds.) 2013 © NATO CCD COE Publications, Tallinn, A Cyber Attack Modeling and Impact Assessment Framework, Igor Kotenko, Andrey Chechulin. – p. 24.

30. Методы прогнозирования. URL: http://statsoft.ru/solutions/tasks/forecast (дата звернення 10.12.2019)

31. Time Series Analysis Solution for Business. URL: https://www.gmdhshell.com/time-series-analysis-software (дата звернення 10.12.2019)

32. Time Series Modeler. URL: https://www.ibm.com/support/knowledgecenter/en/SSLVMB_22.0.0/com.ibm.spss.statis tics.help/spss/trends/idh_idd_tab_vars.htm (дата звернення 10.12.2019)

33. Husak M. Survey of Attack Projection, Prediction, and Forecasting in Cyber Security URL: https://www.researchgate.net/publication/327449459_Survey_of_Attack_Projection_Prediction_and_Forecasting_in_Cyber_Security (дата звернення 10.12.2019)

34. Козьменко О. В. Економко-математичні методи та моделі (економетрика): навчальний посібник / О. В. Козьменко, О. В. Кузьменко. – Суми : Університетська книга, 2014 – 406 с.

35. Присенко Г. В. Прогнозування соціально-економічних процесів: навчальний посібник / Г. В. Приенко, Є. І Равікович — К.: КНЕУ, 2005. — 378 с.

36. Прогнозування та аналіз часових рядів. Методичні вказівки до практичних занять та самостійної роботи студентів спеціальності 051 «Економіка» освітня програма «Економічна кібернетика», «Економічна аналітика» / Укл.: Юрченко М. Є. – Чернігів: ЧНТУ, 2018. – 88 с.

37. Халафян А. А. STATISTICA 6.0 Статистичний аналіз даних. 3-вид. Підручник – М.: ООО «Бином-Прес», 2007 р.

38. Нейронні мережі. URL: http://www.aiportal.ru/articles/neural-networks/neural-networks.html (дата звернення 10.12.2019)

39. Zanab, H. Osman Mohamed L.Awad Neural network based approach for short term load forecasting [Text] / H. Osman Zanab // Power systems conference and exposition. – 2009. – P. 1-8.

40. Kargapoltsev S.K. A dynamic updating algorithm of smoothing parameter values of probabilistic neural networks / S.K. Kargapoltsev, Y.M. Krakovsky, A.V. Lukyanov, A.N. Luzgin // Far East Journal of Electronics and Communications. – 2017. – vol. 17. – №. 4. – P. 909–914.

41. Ethem A. Introduction to Machine Learning: Massachusetts Institute of Technology. – The MIT Press Cambridge. – 2014. – 616 p.

42. Howard Demuth, Neural Network ToolboxTM 6 User's Guide. URL: https://www.mathworks.com/ (дата звернення 10.12.2019)

43. Грабовецький Б.Є. Основи економічного прогнозування: навчальний посібник. / Б.Є. Грабовецький. – Вінниця: ВФ ТАНГ, 2000. – 209 с.

44. Злепко С. М. Математичні моделі: означення, характеристика, етапи побудови URL: https://web.posibnyky.vntu.edu.ua/firen/3zlepko_osnovy_biomedychnogo_radioelektronnogo_aparatobuduvannya/6.html (дата звернення 10.12.2019)

45. Порівняння програмних продуктів для аналізу даних. URL: https://www.mbureau.ru/blog/sravnenie-programmnyh-produktov-dlya-analiza-dannyh-r-matlab-scipy-ms-excel-sas-spss-stata (дата звернення 10.12.2019)

46. Методы прогнозирования. [URL: http://statsoft.ru/solutions/tasks/forecast/ (дата звернення 10.12.2019)

47. Математика. Графіка. Програмування. URL: https://exponenta.ru/products/matlab#description (дата звернення 10.12.2019)

48. Аналіз часових рядів в MATLAB. URL: https://matlab.ru/solutions/application/computational-finance/time-series-analysis (дата звернення 10.12.2019)

49. System identification toolbox. URL: https://www.mathworks.com/products/sysid.html (дата звернення 10.12.2019)

50. SAS/STAT Software. URL: https://www.sas.com/ru_ua/software/stat.html (дата звернення 10.12.2019)

51. What is EVIEWS. URL: https://www.eviews.com/home.html (дата звернення 10.12.2019)

52. Центр статистичного аналізу. URL: https://www.statmethods.ru/trainings/eviews-treningi/ (дата звернення 10.12.2019 )

53. Kyung-Bin Song et al., Hybrid load forecasting method with analysis of temperature sensitivities] / Kyung-Bin Song et al // Trans. Power Syst. – 2006. – vol. 21, №2

# ДОДАТКИ

Додаток А

## SUMMARY

Losyna Y.S. Predicting the possibility of cyber-attacks by financial institutions. – Masters-level Qualification Thesis. Sumy State University, Sumy, 2019.

The work examines the nature of cyber attacks, methods and models of forecasting cyber-attacks. The main factors of bank instability and risks arising from cyberattacks have been analyzed. The main purpose of the study is to build a model for predicting the potential volume of cyberattacks on financial institutions.

Keywords: cyberattack, forecasting, financial institutions, modeling, neural network

## АНОТАЦІЯ

Лосина Є.С. Прогнозування можливості кібератак фінансових установ. – Кваліфікаційна магістерська робота. Сумський державний університет, Суми, 2019 р.

У роботі досліджено сутність кібератак, методи і моделі прогнозування обсягів кібератак. Проведено аналіз основних факторів нестабільності роботи банків та ризиків, що виникають при цьому в результаті кібератак. Основною метою дослідження є побудова моделі прогнозування можливих обсягів кібератак на фінансові установи.

Ключові слова: кібератака, прогнозування, фінансові установи, моделювання, нейромережа.

# Додаток Б

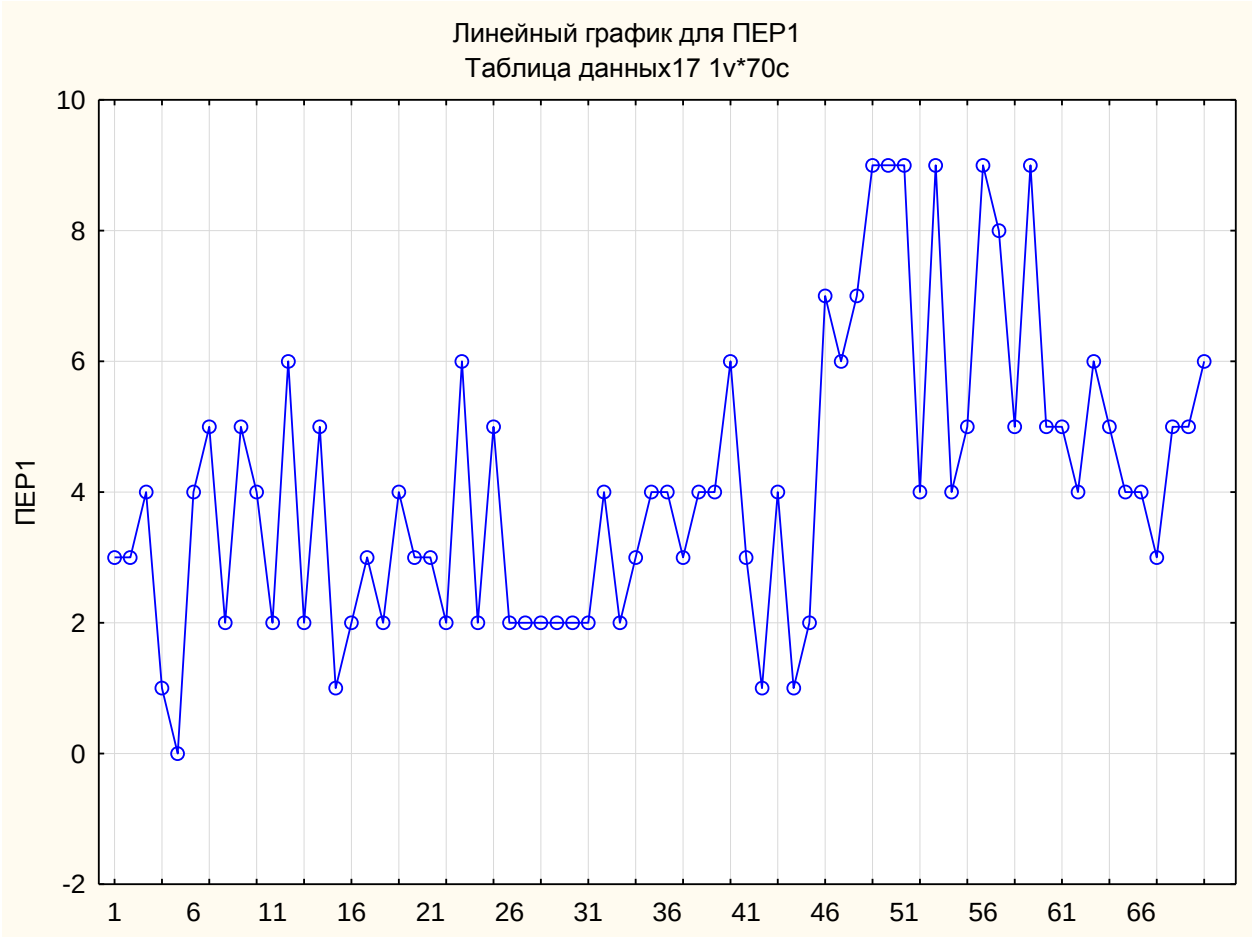



Рисунок А.1 – Кількість кібератак на всьому досліджуваному діапазоні

# Додаток В

Таблиця Б.1 – Кількість кібератак на фінансові установи

| Кількість атак | Дата | Кількість атак | Дата | Кількість атак | Дата |
|---|---|---|---|---|---|
| 3 | 01.01.2014 | 2 | 01.12.2015 | 6 | 01.11.2017 |
| 3 | 01.02.2014 | 5 | 01.01.2016 | 7 | 01.12.2017 |
| 4 | 01.03.2014 | 2 | 01.02.2016 | 9 | 01.01.2018 |
| 1 | 01.04.2014 | 2 | 01.03.2016 | 9 | 01.02.2018 |
| 0 | 01.05.2014 | 2 | 01.04.2016 | 9 | 01.03.2018 |
| 4 | 01.06.2014 | 2 | 01.05.2016 | 4 | 01.04.2018 |
| 5 | 01.07.2014 | 2 | 01.06.2016 | 9 | 01.05.2018 |
| 2 | 01.08.2014 | 2 | 01.07.2016 | 4 | 01.06.2018 |
| 5 | 01.09.2014 | 4 | 01.08.2016 | 5 | 01.07.2018 |
| 4 | 01.10.2014 | 2 | 01.09.2016 | 9 | 01.08.2018 |
| 2 | 01.11.2014 | 3 | 01.10.2016 | 8 | 01.09.2018 |
| 6 | 01.12.2014 | 4 | 01.11.2016 | 5 | 01.10.2018 |
| 2 | 01.01.2015 | 4 | 01.12.2016 | 9 | 01.11.2018 |
| 5 | 01.02.2015 | 3 | 01.01.2017 | 5 | 01.12.2018 |
| 1 | 01.03.2015 | 4 | 01.02.2017 | 5 | 01.01.2019 |
| 2 | 01.04.2015 | 4 | 01.03.2017 | 4 | 01.02.2019 |
| 3 | 01.05.2015 | 6 | 01.04.2017 | 6 | 01.03.2019 |
| 2 | 01.06.2015 | 3 | 01.05.2017 | 5 | 01.04.2019 |
| 4 | 01.07.2015 | 1 | 01.06.2017 | 4 | 01.05.2019 |
| 3 | 01.08.2015 | 4 | 01.07.2017 | 4 | 01.06.2019 |
| 3 | 01.09.2015 | 1 | 01.08.2017 | 3 | 01.07.2019 |
| 2 | 01.10.2015 | 2 | 01.09.2017 | 5 | 01.08.2019 |
| 6 | 01.11.2015 | 7 | 01.10.2017 | 5 | 01.09.2019 |
| | | | | 6 | 01.10.2019 |