Фізико-математичні науки

УДК 65.51

Савчук Олександр Васильович

аспірант

Національного університету «Острозька академія»

Savchuk Oleksandr

Postgraduate Student of the

National University of Ostroh Academy

МОДЕЛЮВАННЯ РИЗИКІВ КІБЕРЗАГРОЗ ІЗ ВИКОРИСТАННЯМ МОДЕЛЕЙ ПОШИРЕННЯ ЕПІДЕМІЙ
MODELING CYBER THREAT RISKS USING EPIDEMIC SPREAD MODELS

Анотація. *У статті представлено опис поняття моделі поширення епідемій, розглянуто їх основні види та детально описано SIR-модель. Також виділено особливості застосування моделей поширення епідемій у сфері інформаційної безпеки банку, адаптовано SIR-модель для її застосування у моделюванні поширення кіберзагроз і розглянуто переваги та недоліки її застосування.*

Ключові слова: *модель поширення епідемії, SIR-модель, кібербезпека, інформаційна безпека банку, математичне моделювання.*

Summary. *The article describes the concept of an epidemic spread model, considers their main types and describes the SIR model in detail. The article also highlights the peculiarities of applying epidemic spread models in the field of information security of banks, adapts the SIR model for its use in modeling the*

spread of cyber threats, and considers the advantages and disadvantages of its application.

Key words: *epidemic spread model, SIR model, cybersecurity, bank information security, mathematical modeling.*

Постановка проблеми. За останні роки кількість і складність загроз кібербезпеки в банківському секторі значно зросла. Конфіденційність, цілісність і доступність чутливої фінансової інформації перебувають під серйозною загрозою через ці небезпеки, які також можуть завдати серйозної шкоди фінансам і репутації постраждалих банків.

Моделі поширення епідемій, такі як модель SIR, використовуються в епідеміології як одна зі стратегій для розуміння та моделювання передачі інфекційних захворювань. Базові припущення та керівні принципи цих моделей можуть бути модифіковані з урахуванням поширення кіберзагроз у банках.

У цьому дослідженні розглядаються особливості використання моделей поширення епідемій для моделювання інформаційної безпеки банків, особливо у світлі поширення кіберзагроз. У дослідженні розглядається, як ці моделі використовуються в епідеміології та як вони можуть бути адаптовані для моделювання поширення кіберзагроз у банках. У звіті також проаналізовано недоліки та труднощі застосування цих моделей у сфері кібербезпеки. Зрештою, звіт прагне запропонувати розуміння того, як моделі поширення епідемій можуть допомогти банкам краще зрозуміти та зменшити кіберзагрози.

Аналіз останніх досліджень та публікацій. У різні періоди досліджень вчені виділяли особливості застосування моделей поширення епідемій. Зазвичай, вони використовуються у медичній галузі, що описано Г. Ґарґом та

А. Насіром [7] і Г. Чаквунонсо [4]. Ці дослідники описують практичне застосування моделей поширення епідемій, зокрема, під час пандемії COVID-19. Застосування моделей у сфері інформаційної безпеки поверхнево описано у М. Лорен та О. Плантефіва [6], вони окреслюють особливості банківських моделей.

Виклад основного матеріалу. Моделі поширення епідемій - це математичні моделі, які описують поширення інфекційного захворювання серед населення в часі. Ці моделі є важливими інструментами для прогнозування та розуміння того, як поширюються хвороби, а також для обґрунтування заходів громадського здоров'я, спрямованих на боротьбу зі спалахами [1]. Існує кілька типів епідемічних моделей, до яких відносять компартмент-моделі та моделі, засновані на даних.

Компартмент-моделі, такі як модель SIR (S – здорові, I – інфіковані R – ті, що одужали), поділяють населення на різні групи залежно від їхнього статусу захворювання. Ці моделі використовують диференціальні рівняння для опису потоку людей між різними відділеннями і для прогнозування того, як хвороба буде поширюватися з часом [2]. Модель SIR припускає, що люди можуть переміщатися між групами у фіксований спосіб, і не враховує гетерогенність всередині популяції. Це проста, але ефективна модель для прогнозування поширення інфекційних захворювань в однорідній популяції.

Моделі, керовані даними, використовують дані часових рядів для оцінки параметрів моделі та прогнозування того, як хвороба буде поширюватися з часом. Ці моделі не покладаються на попередні знання про хворобу і можуть відображати більш складну динаміку, наприклад, вплив таких заходів, як локдауни або кампанії з вакцинації [3]. Однак вони можуть бути більш трудомісткими в обчисленнях і вимагати більше даних, ніж окремі моделі.

Прикладами епідемічних моделей є моделі SIR, SEIR (здорові – контактні – інфіковані – ті, що одужали) та SIRD (здорові – інфіковані – ті, що одужали – померлі) [3]. Ці моделі були використані для вивчення ряду інфекційних захворювань, включаючи COVID-19, лихоманку Ебола та грип. Наприклад, дослідники використовували модель SIR для прогнозування поширення COVID-19 у різних регіонах та оцінки впливу різних заходів, таких як соціальне дистанціювання та кампанії з вакцинації. Моделі, засновані на даних, також використовувалися для прогнозування поширення COVID-19 та оцінки ефективності різних втручань у сфері охорони здоров'я.

Фахівці з мережевої безпеки часто використовують концепцію застосування моделей епідемій для імітації передачі вірусу через мережу. У цьому методі мережа розбивається на секції, і набір диференціальних рівнянь використовується для представлення руху людей між цими секціями. Модель може бути використана для визначення найслабших місць мережі та створення засобів захисту для зменшення ризиків.

Наприклад, в контексті інформаційної безпеки банку одним з підходів до використання моделей епідемій є моделювання поширення вірусу в мережі банку. Мережа банку може бути розділена на частини, такі як робочі станції, сервери та маршрутизатори. Потім модель може імітувати поширення вірусу через мережу, описуючи потоки людей між цими частинами. Модель може бути використана для виявлення найбільш вразливих частин мережі та розробки контрзаходів для зменшення ризиків. Цей підхід можна використовувати в поєднанні з іншими методами, такими як моделювання загроз, для розробки комплексної стратегії захисту інформаційної безпеки банку [4].

Інший підхід полягає у використанні моделі SEIR для моделювання поширення кібератаки. У цьому підході модель SEIR адаптується для

моделювання поширення кібератаки шляхом поділу сукупності користувачів на чотири групи: здорові, контактні, інфіковані та ті, що одужали. Здорова частина представляє частини мережі, які є вразливими до атаки, контактна частина представляє частини мережі, які були уражені атакою, інфікована частина представляє частини мережі, які були інфіковані атакою, а відновлена частина представляє частини мережі, які були відновлені після атаки [5].

Розглянемо для прикладу модель SIR – просту математичну модель, яка використовується для розуміння поширення інфекційних захворювань. Диференціальні рівняння моделі SIR мають наступний вигляд:

$$\frac{dS}{dt} = -\frac{\beta IS}{N} \ (1),$$

$$\frac{dI}{dT} = \frac{\beta IS}{N} - \gamma I \ (2),$$

$$\frac{dR}{dt} = \gamma I \ (3)$$

де S – кількість здорових особин, які можуть захворіти, I – кількість інфікованих особин, R - кількість особин, що одужали, N – загальна кількість особин (тобто, S+I+R), β – рівень передачі, або швидкість, з якою інфікуються здорові особи, γ – рівень одужання або швидкість, з якою інфіковані особи одужують і стають несприйнятливими до хвороби.

Рівняння (1) відображає швидкість зміни здорових особин у часі, яка пропорційна кількості інфікованих особин (I) та швидкості передачі (β) і обернено пропорційна загальній кількості особин (N). Знак мінус тут вказує на те, що кількість здорових особин зменшується в міру того, як вони інфікуються.

Рівняння (2) представляє швидкість зміни кількості інфікованих особин з часом, яка пропорційна кількості здорових особин (S), швидкості передачі (β) та кількості інфікованих особин (I) і обернено пропорційна загальній

кількості особин (N). Величина γI відображає швидкість, з якою інфіковані особини одужують і переходять у категорію тих, хто одужав.

Рівняння (3) представляє швидкість зміни кількості тих, хто одужав з часом, яка пропорційна кількості тих, хто одужав (R) і швидкості одужання (γ).

Щоб адаптувати модель SIR для моделювання поширення кібератаки, слід змінити припущення та параметри моделі, аби врахувати унікальні характеристики кіберзагроз. Так, у традиційній моделі особини поділяються на здорових, інфікованих та тих, хто одужав. В контексті кіберзагроз можна змінити ці категорії наступним чином: здорові (S) – особи або пристрої, які є вразливими до кіберзагрози; інфіковані (I) – особи або пристрої, які були скомпрометовані або заражені кіберзагрозою; ті, хто одужав (R) – особи або пристрої, які відновилися після кіберзагрози або були успішно відновлені.

Також у традиційній моделі швидкість передачі та відновлення представляє ймовірність переходу особи з однієї категорії в іншу. У контексті кіберзагроз можна модифікувати ці показники, щоб відобразити ймовірність зараження або відновлення пристрою на основі таких факторів, як виправлення, сканування вразливостей і оновлення програмного забезпечення для підтримки безпеки. Тому, швидкість передачі (β) – ймовірність того, що пристрій буде заражений кіберзагрозою; швидкість одужання (γ) – ймовірність того, що заражений пристрій буде успішно відновлений.

На поширення кіберзагроз впливає широкий спектр факторів, таких як поведінка користувачів, структура мережі та заходи безпеки. Ці фактори можна включити в модель, відповідно змінивши швидкість передачі та відновлення. Наприклад, швидкість передачі може змінюватись залежно від кількості вразливих пристроїв у мережі та наявних заходів безпеки, які застосовуються при роботі з ними.

За допомогою модифікованих параметрів моделі можна здійснити моделювання поширення кіберзагрози у часі. Для створення найпростішої моделі скористаємось середовищем python. Код моделі виглядатиме наступним чином:

```
import numpy as np
import matplotlib.pyplot as plt

# Встановлюємо початковий розмір популяції
population_size = 1000
# Встановлюємо початкову кількість заражених пристроїв
initial_infected = 5
# Встановлюємо швидкість передачі та швидкість відновлення
transmission_rate = 0.2
recovery_rate = 0.1
# Встановлюємо кількість часових проміжків
simulation_time = 100

# Ініціалізуємо масиви для здорових, інфікованих та відновлених пристроїв
susceptible = np.zeros(simulation_time)
infected = np.zeros(simulation_time)
recovered = np.zeros(simulation_time)
infected[0] = initial_infected
susceptible[0] = population_size - initial_infected

# Змоделюємо поширення кіберзагрози у часі
for i in range(1, simulation_time):
  # Розрахунок кількості пристроїв, які можуть бути зараженими
  infected_chance = transmission_rate * infected[i-1] * susceptible[i-1] /
population_size
  # Розрахунок заражених пристроїв, які відновилися
  recovered_chance = recovery_rate * infected[i-1]

  susceptible[i] = susceptible[i-1] - infected_chance
  infected[i] = infected[i-1] + infected_chance - recovered_chance
  recovered[i] = recovered[i-1] + recovered_chance

plt.plot(susceptible, label="Здорові")
plt.plot(infected, label="Заражені")
plt.plot(recovered, label="Відновлені")
plt.legend()
plt.xlabel("Часовий інтервал")
plt.ylabel("Кількість пристроїв")
plt.title("Моделювання поширення кіберзагрози")
plt.show()
```

І, відповідно, візуальне відображення поширення буде наступним:

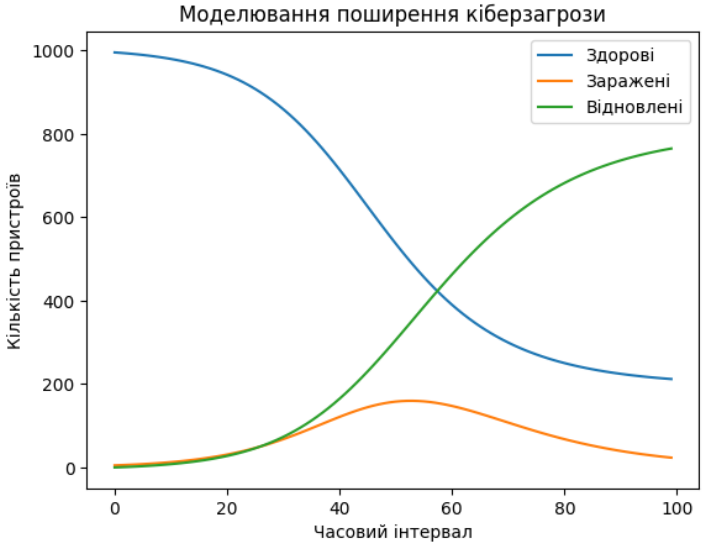

Рис. 1. Результат виконання коду

У цьому коді була модифікована SIR-модель, де було визначено початковий розмір популяції як 1000 одиниць пристроїв; початкову кількість пристроїв, які були інфіковані, встановили 5 одиниць; а також встановили швидкість передачі та швидкість відновлення на рівні 0,2 та 0,1 відповідно, що відповідає досить низькому рівню зараженості кіберзагрози та дуже низькому рівню відновлюваності пристроїв. Ці показники є індивідуальними для кожного середовища. Також було визначено час моделювання, протягом якого і відбувається весь процес, у 100 часових одиниць. Код є спрощеним прикладом та для подальшого використання потребує модифікацій та калібрування, що враховуватиме максимальну кількість факторів як діють в реальних умовах, для точного моделювання.

Загалом, модель SIR, як модель поширення епідемій має кілька особливостей, які роблять її корисною для моделювання поширення кіберзагроз у часі:

- Модель можна використовувати для вивчення динаміки поширення кіберзагроз, а також динаміки розвитку хвороб або інфекцій. Відстежуючи прогрес переходу особин у групах, можна зрозуміти довгострокову поведінку явища, тобто як загроза буде поширюватись із часом.

- Модель можна використовувати для оцінки ефективності різних заходів контролю, таких як посилення заходів кібербезпеки або запровадження правил, що обмежують доступ до вразливих систем. Це дає можливість визначити найкращу тактику для зменшення поширення кіберзагрози.

- Модель є багатофункціональною і легко адаптується до нових змінних або факторів, які можуть мати відношення до конкретної кіберзагрози, що моделюється. Це дозволяє адаптувати модель до конкретних обставин і характеристик загрози.

Щоб визначити вплив різних припущень або параметрів на поширення кіберзагрози, модель можна використовувати для проведення аналізу чутливості. Це може допомогти знайти потенційні слабкі місця в моделі або області невизначеності і внести необхідні корективи.

Моделі поширення епідемій широко використовуються в контексті захворюваності для розуміння та прогнозування поширення вірусів та інфекцій. Однак, коли йдеться про моделювання інформаційної безпеки банку, їх використання має низку переваг і недоліків.

До переваг можна віднести наступні [6]:

- Моделі поширення епідемій можуть забезпечити основу для розуміння взаємозалежності різних систем безпеки в банку. На основі моделювання, банки можуть розробити надійну основу для прийняття стратегічних рішень.

- Моделі поширення епідемій можуть допомогти банкам передбачати та управляти ризиком порушень безпеки. Розробивши план дій на випадок непередбачених обставин, банки можуть переглянути свої політики щодо ризиків, пов'язаних з моделлю, і посилити обмеження моделі з чіткими рівнями толерантності для конкретних сценаріїв.

До недоліків можна віднести наступні [7]:

- Моделі поширення епідемій призначені для прогнозування стабільного майбутнього і можуть бути неефективними при моделюванні непередбачуваних подій, таких як кібератаки.

- Моделі сильно залежать від зовнішніх факторів і можуть бути неефективними для відображення довгострокової динаміки порушень безпеки.

- Точність моделей епідемій значною мірою залежить від припущень і меж, визначених на етапі розробки.

Висновки. Моделювання та прогнозування поширення кіберзагроз у банківській сфері можна здійснити за допомогою моделей поширення епідемій, таких як SIR-модель. Ці моделі мають низку переваг, серед яких їхня простота, здатність висвітлювати динаміку поширення загрози та оцінювати ефективність заходів контролю. Однак слід розуміти обмеження цих моделей, особливо в контексті інформаційної безпеки банків. Нестача даних, необхідність коригування моделі з урахуванням особливостей загрози або банківського сектору, а також можливість того, що припущення не виправдаються, - ось кілька прикладів цих обмежень.

Загалом, використання моделей поширення епідемії може бути корисним інструментом для науковців і практиків банківського сектору для розуміння поширення кіберзагроз і пошуку ефективних методів зменшення їхнього впливу. Однак ці моделі слід використовувати в поєднанні з іншими

стратегіями і ретельно адаптувати до унікальних обставин, пов'язаних із загрозою і сектором, що моделюється. Банкам слід розглянути різні підходи, включаючи теорію ігор і машинне навчання, щоб ефективно управляти ризиком порушень безпеки.

Література

1. Epidemic, Endemic, Pandemic: What are the Differences? // Columbia University Mailman School of Public Health. URL: https://www.publichealth.columbia.edu/news/epidemic-endemic-pandemic-what-are-differences (дата звернення: 25.05.2023).
2. Mollison D. Epidemic Models: Their Structure and Relation to Data. Edinburgh: Heriot-Watt University, 1995. 444 p.
3. Lee G., Yoon S-e., Shin K. Simple epidemic models with segmentation can be better than complex ones // Plos One. 2022. №1. P. 1-18.
4. Nwokoye Ch.H., Madhusudanan V. Epidemic Models of Malicious-Code Propagation and Control in Wireless Sensor Networks: An Indepth Review. Wireless Personal Communications. 2022. №125. P. 1827-1856.
5. Rao Y.S., Rauta A.K., Saini H., Panda T.C. Mathematical Model for Cyber Attack in Computer Network. International Journal of Business Data Communications and Networking. 2017. №13(1). P. 58-65.
6. Laurent M.-P., Plantefève O., Tejada M., Weyenbergh F. Banking models after COVID-19: Taking model-risk management to the next level // McKinsey & Company. URL: https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/banking-models-after-covid-19-taking-model-risk-management-to-the-next-level (Дата звернення: 25.05.2023).
7. Garg H., Nasir A., Jan N., Khan S. U. Mathematical analysis of COVID-19 pandemic by using the concept of SIR model // National Library of Medicine.

URL: https://pubmed.ncbi.nlm.nih.gov/34483720/ (дата звернення: 25.05.2023).