

Олег Густера

к.е.н., асистент

КНУ ім. Тараса Шевченка, кафедра статистики, інформаційно-аналітичних систем та демографії, Київ.

ORCID ID 0000-0003-1010-6100

gusteraom87@gmail.com

СУЧАСНІ ПРОБЛЕМИ ВИКОРИСТАННЯ СТАТИСТИЧНИХ ДАНИХ В МОДЕЛЮВАННІ ОЦІНКИ РИЗИКУ

Анотація. Моделювання ризикових ситуацій та оцінка ризику може базуватись на декількох інформаційних джерелах серед яких одним з найбільш поширених виступають статистичні дані. Так, наприклад, оцінка ризику здійснення атак на комп'ютерну мережу залежить від загальної кількості атак та може оцінюватись з використанням статистичних даних та визначення тенденцій відносно загальної кількості кібератак. Ризик зменшення попиту на певний продукт чи послугу може базуватись на статистичних даних про зміни чисельності чи структури населення держави. При цьому якість отриманих статистичних даних, що включає в себе точність, надійність, відповідність форми напряму визначає доцільність їх використання та відповідно ефективність процесу оцінки ризику. В таких умовах формування бази знань для подальшого використання отриманих даних у процесі прийняття рішень відносно управління ризиковими ситуаціями та оцінки ризику можливе лише за умови використання відфільтрованих джерел та структурного аналізу отриманих даних.

Ключові слова: Моделювання оцінки ризику, ризикова ситуація, статистичні дані.

Моделювання ризикових ситуацій та оцінка ризику може базуватись на декількох інформаційних джерелах серед яких одним з найбільш поширених виступають статистичні дані. Так, наприклад, оцінка ризику здійснення атак на комп'ютерну мережу залежить від загальної кількості атак та може оцінюватись з використанням статистичних даних та визначення тенденцій відносно загальної кількості кібератак. Ризик зменшення попиту на певний продукт чи послугу може базуватись на статистичних даних про зміни чисельності чи структури населення держави. При цьому якість отриманих статистичних даних, що включає в себе точність, надійність, відповідність форми напряму визначає доцільність їх використання та відповідно ефективність процесу оцінки ризику.

До основних проблем що виникають при використанні статистичних даних можна віднести наступні:

- відсутність даних (за певним показником, за певний період),
- приховування даних,
- різні джерела з різними даними,
- псування або зміна даних.

Наявність великої кількості інструментів збору, передачі та візуалізації інформації, таких як Google Data Studio дозволяє вирішити питання кількісного наповнення бази даних але не завжди дозволяє адекватно оцінити якість отриманих даних та обирати лише джерела що надають якісні дані, обирати повні ряди даних, оцінювати якість даних після їх отримання.

В таких умовах формування бази знань для подальшого використання отриманих даних у процесі прийняття рішень відносно управління ризиковими ситуаціями та оцінки ризику можливе лише за умови використання відфільтрованих джерел. Так, наприклад, кожне джерело може отримувати вагу залежно від того, наскільки якісна інформація раніше отримувалась з цього джерела. Можливість отримання даних з різних джерел дозволяє усунути проблему повної відсутності даних – неможливо отримати абсолютно всю необхідну інформацію лише з одного джерела.

Виконувати ручну перевірку вхідних даних за допомогою засобів візуалізації чи перегляду не завжди раціонально та доцільно з економічної точки зору, перегляд всіх даних неможливий за умови великої кількості джерел інформації. Тому в першу чергу необхідно забезпечити перевірку вхідних даних в автоматичному режимі.

Так, наприклад, якщо з джерела даних вже було отримано один раз недостовірні дані, воно не може мати високий рівень довіри та відповідно всі дані що від нього надходять не можуть використовуватись напямую без перевірки. Визначити рівень довіри джерела можна за допомогою даних що були передані раніше якщо їх можна порівняти з еталонними значеннями. Так, наприклад, якщо найбільш точними результатами про чисельність вважати перепис то доцільно порівнювати дані що надходили з певного джерела з цим показником. Визначення еталонних даних може проводитись експертом або аналітиком що надалі проводить оцінку ризику згідно з отриманими вхідними даними. У якості еталонних значень можуть використовуватись дані державних органів статистики або інші джерела що є довіреними з точки зору експерта.

В той же час, даний підхід унеможливує перевірку фіктивних даних які були змінені навмисне, та при цьому їх неможливо порівняти з еталонними. В цьому випадку доцільно використовувати декілька методів перевірки одночасно та визначати ступінь довіри до джерела за сукупністю показників.

При цьому отримання точних та достовірних даних з офіційних джерел може призводити до використання застарілої інформації що вже не є актуальною та надходить із запізнення більше року.

Оцінка ризику настання ризикової події такої як дорожньо-транспортна пригода включає в себе такий параметр як кількість зареєстрованих транспортних засобів. Як правило, між цими показниками існує прямий зв'язок – збільшення кількості транспортних засобів призводить до збільшення дорожньо-транспортних пригод. Тому отримання точних даних про кількість транспортних засобів та побудування достовірних прогнозів на основі цих даних є однією з ключових задач для ринку страхових послуг.

Відслідкувати кількість транспортних засобів що фактично використовуються та приймають участь у дорожньому русі складніше – як правило даний показник буде наближений до кількості застрахованих автомобілів, тому що автомобілі без поліса майже не використовуються. Даний показник буде більш точним, але його отримання не завжди можливе через публічні джерела даних. Таким чином, модель оцінки ризику дорожньо-транспортної пригоди може використовувати одночасно як показники кількості зареєстрованих транспортних засобів, так і показники кількості проданих полісів страхування авто. При цьому прогноз кількості авто на ринку одночасно може використовуватись страховими компаніями для інших цілей – наприклад прогнозу попиту на ринку страхових послуг.

Окрім кількісної оцінки ринку авто для формування оцінки ризику в даному прикладі доцільно використовувати також структурний аналіз. Так, наприклад, збільшення кількості зареєстрованих авто може непропорційно збільшити ризик дорожньо-транспортних пригод. Тобто, якщо на 2021 рік в Україні нараховувалось 10,2 млн транспортних засобів і за 2022 рік цей показник збільшився на 0,5 млн, то кількість фактично використаних засобів зростає із більшим темпом ніж розрахунковий показник 4,9% та фактично може використовуватись тільки для оцінки наявності тенденції та напрямку зміни показника.

MINISTRY OF EDUCATION
AND SCIENCE OF UKRAINE

NATIONAL UNIVERSITY
OF LIFE AND ENVIRONMENTAL
SCIENCES OF UKRAINE

FACULTY OF INFORMATION
TECHNOLOGY

МІНІСТЕРСТВО ОСВІТИ
І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БІОРЕСУРСІВ І
ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ

PROCEEDINGS

XI International scientific
conference

**GLOBAL AND
REGIONAL PROBLEMS OF
INFORMATIZATION IN
SOCIETY AND
NATURE USING
'2023**

15-16 November 2023

Kyiv, NULES of Ukraine

Kyiv 2023

МАТЕРІАЛИ

XI Міжнародної науково-практичної
конференції

**ГЛОБАЛЬНІ ТА
РЕГІОНАЛЬНІ ПРОБЛЕМИ
ІНФОРМАТИЗАЦІЇ В
СУСПІЛЬСТВІ І
ПРИРОДОКОРИСТУВАННІ
'2023**

15-16 листопада 2023 року

Київ, НУБіП України

Київ 2023

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

МАТЕРІАЛИ

XI Міжнародної науково-практичної конференції

ГЛОБАЛЬНІ ТА РЕГІОНАЛЬНІ ПРОБЛЕМИ ІНФОРМАТИЗАЦІЇ В СУСПІЛЬСТВІ І ПРИРОДОКОРИСТУВАННІ '2023

15-16 листопада 2023 року

Київ, НУБіП України

Київ 2023

УДК 004

Рекомендовано до друку вченою радою факультету інформаційних технологій Національного університету біоресурсів і природокористування України (протокол № 4 від 20.11.2023)

Укладач: к.е.н., доцент Харченко В.В.

Збірник матеріалів XI Міжнародної науково-практичної конференції "Глобальні та регіональні проблеми інформатизації в суспільстві і природокористуванні '2023", 15-16 листопада 2023 року, НУБіП України, К. НУБіП України, 2023. 117 с.

Відповідальність за зміст публікацій несуть автори.

© Національний університет біоресурсів
і природокористування України, 2023