DOI 10.31891/2307-5732-2021-303-6-115-123
УДК 519.857.3

ДЬОГТЄВА І. О.
Вінницький національний технічний університет
ORCID ID: 0000-0002-8567-6952
e-mail: iryna.djogtjeva@gmail.com
ШИЯН А. А.
Вінницький національний технічний університет
ORCID ID: 0000-0002-5418-1498
e-mail: anatoliy.a.shiyan@gmail.com

ІМІТАЦІЙНЕ МОДЕЛЮВАННЯ РОБОТИ ГРУПИ РЕАГУВАННЯ НА ІНЦИДЕНТИ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ ПРИ КІБЕРАТАКАХ

Метою роботи є статистичний аналіз результатів імітаційного моделювання діяльності групою реагування на інциденти інформаційної безпеки (ГРІІБ) залежно від параметру підвищення інтенсивності потоку інцидентів інформаційної безпеки. ГРІІБ розглядається як система масового обслуговування в умовах зростання інтенсивності навантаження. В якості вхідних параметрів для імітаційного моделювання використано інтенсивність надходження заявок, характеристики наростання інтенсивності, характеристики обслуговування заявок ГРІІБ. Розроблена комп'ютерна програма задає випадкові значення цих параметрів, розподілені за показниковим законом. Для кожного набору значень розраховується набір стаціонарних значень, що характеризує ефективність роботи ГРІІБ, в якості яких вибрано ймовірності режимів роботи та очікування. Програма здійснює статистичний аналіз отриманих значень. В результаті проведеного імітаційного моделювання отримані кількісні значення як для середніх значень, так і для показників варіативності характеристик, які описують ефективність роботи ГРІІБ. Проведене імітаційне моделювання засвідчило можливість прогнозувати статистичних закономірностей кібератак на ефективність роботи ГРІІБ. Відслідковуючи в процесі розгортання поточної кібератаки статистичні характеристики інцидентів, можна вже в процесі проведення кіберзахисту замінювати одну ГРІІБ на іншу, статистичні характеристики якої будуть забезпечувати потрібний рівень захисту. Для реалізації такого підходу може бути використано розроблений програмний засіб.

Ключові слова: кібератака, інцидент інформаційної безпеки, група реагування, стаціонарний режим, показники варіації.

IRYNA DOHTIEVA, ANATOII SHYIAN
Vinnytsia National Technical University

SIMULATION OF THE WORK OF THE INFORMATION SECURITY INCIDENT RESPONSE TEAM DURING CYBERATTACKS

The aim of the work is statistical analysis of the results of simulation modeling of the information security incident response team (ISIRT) depending on the parameter of increasing the intensity of the flow of information security incidents. ISIRT is considered as a queuing system in conditions of increasing load intensity. As input parameters for simulation the intensity of receipt of applications, characteristics of increase of intensity, characteristics of service of applications of ISIRT are used. The developed computer program sets random values of these parameters, distributed according to the indicator law. For each set of values, a set of stationary values is calculated, which characterizes the efficiency of ISIRT, as selected probabilities of operating modes and expectations. The program performs statistical analysis of the obtained values. As a result of the simulation, quantitative values were obtained both for the average values and for the indicators of variability of characteristics that describe the effectiveness of ISIRT. The simulation showed the ability to predict the statistical patterns of cyber-attacks on the effectiveness of ISIRT. Tracking the statistical characteristics of incidents in the process of deploying the current cyber-attack, it is possible in the process of cyber defense to replace one ISIRT with another, the statistical characteristics of which will provide the required level of protection. To implement this approach, the developed software can be used.

Key words: cyber-attack, information security incident, response group, stationary mode, indicators of variation.

Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями

Стрімке збільшення кількості користувачів Інтернет, зростання потужності серверів та розвиток програмного забезпечення створює умови для зростання насиченості та комплексності кібератак, розширення номенклатури кіберінцидентів. Мають місце також складні умови функціонування кіберпростору, нормою стає не стабільність, а свого роду турбулентність інформаційного простору.

У звіті [1] наведено дані чисельного співвідношення представлення категорій кіберінцидентів: біля 30 % займає категорія шкідливого (образливого) змісту (спам), майже стільки ж відводиться для шкідливого програмного коду, менше 20 % складає збір інформації зловмисником (сканування, сніфінг, фішинг), на інші категорії (спроби втручання, порушення доступності, властивостей інформації, шахрайства, вразливості тощо) припадає більше 20 %.

В результаті необхідності реагування на такі кіберінциденти спостерігається зростання кількості груп реагування на інциденти інформаційної безпеки (ГРІІБ).

Загалом серед основних категорій викликів, які визначені в Стратегіях кібербезпеки країн ЄС, США, Великої Британії, Ізраїлю виділяють: кіберзлочинність, використання нових технологій, цифровізація, недостатній рівень цифрових навичок, питання пов'язані з функціонуванням та використанням сервісів [2]. В Україні даний перелік [3] доповнюється: використанням кіберзасобів у міжнародній конкуренції, мілітаризація кіберпростору, введення дистанційних режимів функціонування. Також в Стратегіях активно

ведеться мова і про функціонування, розбудову систем виявлення вразливостей і реагування на кіберінциденти та кібератаки. В Україні за результатами звіту Оперативного центру реагування на кіберінциденти [1] наведені рекомендації щодо формування з фахівців інформаційної безпеки команд реагування на комп'ютерні надзвичайні події та визначення кола відповідальних осіб за виконання процесу та процедур реагування на кіберінциденти.

Тому сьогодні зростає актуальність проблеми комплектації ГРІБ, причому для забезпечення ефективної діяльності такої групи доцільно досліджувати як показники ефективності самої групи, так і показники окремих інцидентів під час кібератаки.

Аналіз досліджень та публікацій

Зростання кількості кібератак на соціальне середовище вимагає від дослідників порушити питання про створення нового наукового напрямку, який може носити назву «соціальної кібербезпеки» [4]. Необхідність у цьому виникла внаслідок того, що створення та розповсюдження спеціальних наративів здатне спонукати велику кількість людей до таких дій, які підривають стабільність держави. Подібні дії можуть носити як позитивний характер (наприклад, екологічного характеру), так і явно виражене екстремістське спрямування.

При цьому безпосередньо «інформаційний привід» може бути не стільки важливим, як організація його обговорення через соціальні мережі. Вперше на це звернули увагу в [5, 6], розглядаючи феномен так званої «Арабської весни». Автори підкреслюють, що спочатку відбулася інформаційна подія, і тільки після цього з'явилась постійно зростаюча лавина реакції в соціальних мережах, яка в результаті переросла в соціальні заворушення.

Нещодавно цілком аналогічна ситуація спостерігалася у США, коли відносно невелика подія із затриманням порушника переросла в серію народних виступів під гаслами «Black Lives Matter – BLM», де аналогічно головну роль відігравали якраз соціальні мережі. Детальне дослідження здійснено в [7], де автори зазначають (переклад авторів статті): «Люди брали участь в обговоренні з кількох причин; включаючи протести та пропаганду, а також розповсюдження спаму та дезінформації. Згодом багато облікових записів користувачів видаляються їх власниками або призупиняються через порушення правил соціальних мереж.».

Для отримання кількісних показників, що необхідні для ефективної діяльності ГРІБ в таких умовах застосовують імітаційне моделювання. Наприклад, в [8] досліджувалися варіанти впливу ботів на формування суспільної думки. Було виявлено, що «в дуже поляризованих обставинах, залежно від загальної щільності мережі, участь лише 2–4% ботів може бути достатньою, щоб змінити клімат думок … Дані результати демонструють механізм, за допомогою якого боти можуть формувати норми, які потім приймаються користувачами соціальних мереж.».

Виділення невирішених раніше частин загальної проблеми

Таким чином, виникає нове поле для діяльності ГРІБ, яке має досить специфічні характеристики. По-перше, розгортання кібератаки здійснюється протягом досить тривалого часу. По-друге, виникає потреба в аналізі текстів лінгвістичного характеру. По-третє, розробка системи протидії кіберінцидентам такого характеру та усунення їх негативного впливу на соціальне середовище вимагає тривалої роботи ГРІБ. По-четверте, інформаційний привід для подальшої кібератаки, як правило, виникає раптово.

Все це призводить до того, що характеристики як окремих фрагментів кібератаки, так і діяльності ГРІБ мають значну випадкову складову. До того ж, як свідчать експериментальні дані [4 – 8], інтенсивність кібератаки досить стрімко зростає із часом.

Формулювання цілей статті

Метою роботи є розрахунок характеристик статистичних відхилень показників ефективності діяльності ГРІБ в залежності від параметру підвищення інтенсивності потоку інцидентів інформаційної безпеки на основі імітаційного моделювання.

Виклад основного матеріалу

Для моделювання роботи ГРІБ буде використано математичний апарат систем масового обслуговування (СМО) та марковських процесів. В даному випадку специфічність моделі, яка застосовується для моделювання особливостей роботи ГРІБ, полягає в наявності параметру підвищення інтенсивності ідентифікації подій інформаційної безпеки.

В нормальному режимі розгортання кібератаки в часі пропонується розглядати в якості рекурентного потоку заявок (інцидентів) на реагування з параметром λ. В умовах наростання інтенсивності кібератак потік заявок на реагування описується законом Пуассона з параметром інтенсивності $\alpha\lambda$, де $\alpha > 0$ – складова, пов'язана з наростанням кібератаки, тобто кількісна характеристики збільшення її інтенсивності (враховується випадок комплексної кібератаки, підвищення рівня складності атаки тощо). Тривалість обробки інцидентів ГРІБ задається показниковим законом розподілу, який має параметр μ.

Для опису діяльності ГРІІБ в умовах навантаження з причин наростання інтенсивності кібератак пропонується розглянути M/M/1/0 з прискоренням надходження заявок (вхідний потік зі змінною інтенсивністю) за умови незайнятості системи. Процес діяльності даної ГРІІБ можна описати марковським процесом $\xi(t)$, який характеризується множиною станів $\{e_0, e_1\}$: у стані e_0 система перебуває $\tau^{\alpha\lambda}$ – показниково розподілена випадкова величина з параметром $\alpha\lambda$ (інтервали часу між надходженнями заявок з підвищеною інтенсивністю) і з ймовірністю 1 переходить у стан e_1; у стані e_1 система перебуває час $\min(\tau^{\lambda}, \eta^{\mu})$, де τ^{λ} – показниково розподілена випадкова величина з параметром λ (інтервали часу між надходженнями заявок в нормальному режимі), η^{μ} – показниково розподілена випадкова величина з параметром μ (час обслуговування). Із стану e_1 система може перейти назад у стан e_0 з ймовірністю:

$$p_{10} = P(\eta^{\mu} < \tau^{\lambda}), \tag{1}$$

або ж залишитись у стані e_1 з імовірністю:

$$p_{11} = P(\tau^{\lambda} < \eta^{\mu}). \tag{2}$$

Отже, марковський процес $\xi(t)$ у заданих станах перебуває відповідно час $\zeta_0 = \tau^{\alpha\lambda}$ та $\zeta_1 = \min(\tau^{\lambda}, \eta^{\mu})$, причому випадкові величини ζ_0, ζ_1 мають показникові розподіли відповідно з параметрами $\alpha\lambda$, $\lambda + \mu$.

Використання марковського представлення статистичних процесів дало можливість знайти перехідні та стаціонарні характеристики системи. Відмітимо, що в загальному випадку це, як правило, не вдається зробити [9].

Перехідний процес обслуговування у даній системі масового обслуговування у випадку збільшення робочого навантаження здійснюється згідно з вкладеним ланцюгом Маркова [10], який задається матрицею:

$$P = \begin{pmatrix} 0 & 1 \\ \frac{\mu}{\lambda + \mu} & \frac{\lambda}{\lambda + \mu} \end{pmatrix} \tag{3}$$

Для нерозкладного, аперіодичного, з поверненням і додатними елементами ланцюга Маркова існують стаціонарні розподіли ймовірностей. В умовах однорідності для таких ланцюгів існують фінальні ймовірності [11]. Можливість отримати граничні ймовірності дозволяє з'ясувати середній відносний час перебування ГРІІБ в даному *i*-му стані, зокрема, в умовах її функціонування під час протидії кібератакам. Дані ймовірності (3) дозволяють встановити в системі прогнозований стаціонарний режим функціонування. Для досліджуваної ГРІІБ, тобто M/M/1/0 з підвищенням інтенсивності надходження вимоги за умови незайнятості системи, такий режим встановлюється при зростанні $t \to +\infty$ (збільшенні кількості кроків). Результати такого процесу, знайдені для системи, яка досліджується:

$$\begin{aligned} \pi_0 &= \frac{\mu}{\alpha\lambda + \mu} \\ \pi_1 &= \frac{\alpha\lambda}{\alpha\lambda + \mu} \end{aligned} \tag{4}$$

дозволяють описати протягом певного проміжку часу $(0;T)$ процес $\xi(t)$, який оцінюється числами $\pi_0 T$, $\pi_1 T$, де ГРІІБ під час протидії кібератакам, у середньому буде незайнятою приблизно час $\pi_0 T$, і зайнятою обслуговуванням у середньому час $\pi_1 T$.

Моделювання випадкових величин в рамках імітаційного експерименту. З метою проведення імітаційних експериментів для дослідження поведінки математичної моделі ГРІІБ використано програмний пакет реалізації імітаційної моделі досліджуваної ГРІІБ, тобто M/M/1/0 із підвищенням інтенсивності надходження заявки за умови незайнятості системи (модель СМО з навантаженням). Для здійснення моделювання обрано мову програмування Python, яка застосовується для рішень широкого спектру задач, зокрема, для роботи з даними в наукових дослідженнях, DataMining, DataScience, тощо. Використано ряд бібліотечних та сторонніх модулів, серед яких: пакети NumPy, Pandas, бібліотека Matplotlib зі стеку SciPy; Statsmodels, Plotly тощо [12, 13].

Програмний продукт розбито на модулі: динаміки вхідних даних; попередніх вихідних даних з прогнозованими аналітично розрахованими показниками ефективності; для проведення експерименту; серії

експериментів; отримання графічних даних, в тому числі 3D графіка; статистичної обробки за переліком метрик.

До вхідних даних моделювання належать: інтенсивність потоку заявок λ (інтенсивність надходження заявок), параметр навантаження α (параметр підвищення інтенсивності надходження заявок), продуктивність каналу (сервера) μ обслуговування заявок (інтенсивність обслуговування заявок), кількість реалізацій n_r (кількість циклів для проведення сценаріїв в рамках одного експерименту) та кількість проведених експериментів N_e.

З використанням вхідної інформації програма забезпечує генерацію послідовностей випадкових чисел: $\tau^{\alpha\lambda}$, τ^{λ}, η^{μ} – показниково розподілені випадкові величини відповідно з параметрами $\alpha\lambda$, λ, μ. На рис. 1 вказані вхідні значення параметрів для моделювання роботи системи та представлені гістограми щільності розподілів інтервалів часу надходження заявок і розподілу часу обслуговування заявок для різних вхідних параметрів та графіки значень інтервалів між надходженням заявок для вхідного потоку чи інтервалу часу для обслуговування в межах кількості циклів експерименту.

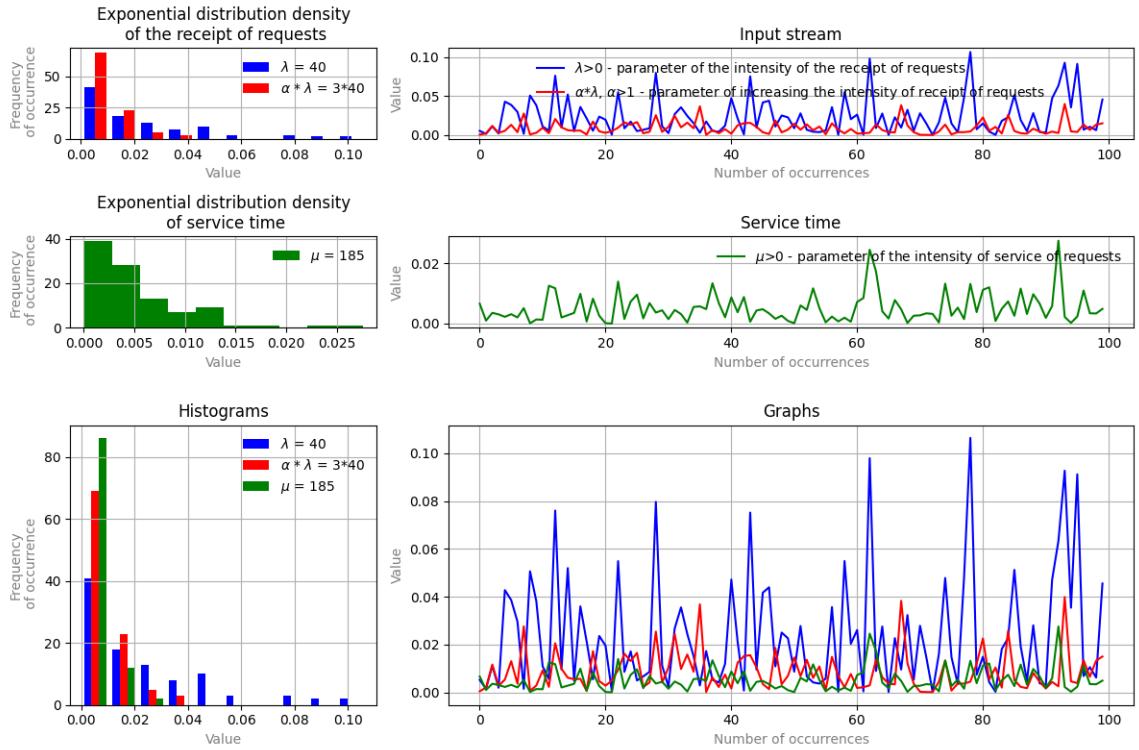

Рис. 1. Приклади гістограм щільності розподілів та графіків послідовностей випадкових значень для характеристик ГРІІБ

Рис. 1 демонструє представлення Пуассонівського вхідного потоку у варіаціях: інтенсивності надходження заявок у звичайному режимі функціонування системи, виділено синім кольором, та з застосуванням параметру підвищення даної інтенсивності – виділено червоним кольором. Оскільки параметр, використаний для генерації інтервалів часу обслуговування – виділено зеленим кольором, найбільший, то спостерігається фіксація відповідних пропорційних значень.

Закон розподілу станів системи (1), (2), на відміну від середньої кількості заявок в ній (завантаженості), більш повно характеризує функціонування СМО під впливом випадкових факторів. Таким чином в ланцюзі Маркова потік випадкових величин визначається лише імовірністю переходу (3) від попереднього значення випадкової величини (стану системи) до наступного [14].

Приклад масиву даних, що моделює поведінку системи, яка досліджується, продемонстровано на рис. 2.

Таблиця (рис. 2) демонструє проведення кожного окремого експерименту, зокрема генерацію послідовностей випадкових чисел необхідних для забезпечення проведення експерименту, значення стану в якому перебуває система та інтервали часу перебування у відповідних станах, що відповідає аналітичному опису діяльності ГРІІБ в умовах навантаження в межах марковського процесу $\xi(t)$, який характеризується множиною станів $\{e_0, e_1\}$ та умовами переходу до відповідних станів (1) та (2).

Для СМО граничний розподіл ймовірностей станів ланцюга Маркова не повинен залежати від початкового розподілу і визначається перехідною матрицею, ймовірність станів у міру збільшення переходів

практично перестають змінюватись і система переходить у стаціонарний режим функціонування [10]. За результатами проведеного експерименту отримуємо (рис. 3) значення експериментальних практичних статистичних характеристик та теоретичних, аналітично розрахованих за формулами для M/M/1/0 з підвищенням інтенсивності надходження вимоги за умови незайнятості системи (4).

lambda	alpha * lambda	mu	from_state1_to01	system_in_state0_intensive	system_in_state1
0.019834775862926236	0.028755848548413416	0.003946903136970928	0	0.028755848548413416	0.003946903136970928
0.03863564658508987	0.0024761127253955808	0.002216281963840574	0	0.0024761127253955808	0.002216281963840574
0.015536290861032327	0.003088377881424763	0.002015935296157324	0	0.003088377881424763	0.002015935296157324
0.003781496997519042	0.0023840602354755053	0.002202640388163617	0	0.0023840602354755053	0.002202640388163617
0.0024651177601278623	0.010271570538348842	0.006024260823362612	1	0.010271570538348842	0.0024651177601278623
0.011196225608147063	0.00713635495061559	0.0030586442251805735	0	0.00713635495061559	0.0030586442251805735
0.015959951560756157	0.009043218809090483	0.0022043268449363686	0	0.009043218809090483	0.0022043268449363686
0.0454970613786084	0.014576436099344718	0.002133863981164457	0	0.014576436099344718	0.002133863981164457
0.024679951316641087	0.008899182210496659	0.00707429450535314	0	0.008899182210496659	0.00707429450535314
0.03368252511098172	0.006981755879556672	0.013030878316272606	0	0.006981755879556672	0.013030878316272606
0.004589970393518583	0.004327952193881813	0.006125574016003025	1	0.004327952193881813	0.004589970393518583
0.031413755593235226	0.0033272547168088197	0.004746763319601698	0	0.0033272547168088197	0.004746763319601698
0.04135329789677394	0.00355012104665229	0.0021129884470656394	0	0.00355012104665229	0.0021129884470656394
0.007969852663897641	0.011860365369183052	0.010801548849876259	1	0.011860365369183052	0.007969852663897641
0.016326839833657236	0.015606040339828217	0.0001641463807734206	0	0.015606040339828217	0.0001641463807734206
0.012594111452934481	0.004144727055692858	0.003964226736402503	0	0.004144727055692858	0.003964226736402503
0.02499284630305134	0.003251202902945514	0.0021031242293622237	0	0.003251202902945514	0.0021031242293622237
0.008492785863792604	0.0017734360683349098	0.003916599801790467	0	0.0017734360683349098	0.003916599801790467
0.054235840071181575	0.0066690839611641154	0.004710545357725302	0	0.0066690839611641154	0.004710545357725302

Рис. 2. Дані експерименту для моделі ГРІБ

static_characteristics	theoretical	practical
π_0	0.6066	0.6015
π_1	0.3934	0.3985
total_system_uptime	-	1.0000

Рис. 3. Результати теоретичних та імітаційних статистичних характеристик

Серія експериментів моделювання функціонування M/M/1/0 з підвищенням інтенсивності надходження вимоги за умови незайнятості системи в рамках одного експерименту формує вибірку стаціонарних характеристик системи для статистичного дослідження (рис. 4).

practical π_0	practical π_1
0.6022455267485098	0.39775447325149005
0.6406159078058036	0.3593840921941965
0.6442642936624592	0.355735706337541
0.6220262690359886	0.3779737309640115
0.5913479104059476	0.4086520895940524
0.5914650743122567	0.4085349256877433
0.571009460933337	0.428990539066663
0.6423990176429085	0.3576009823570915
0.6670377194389605	0.3329622805610394
0.5827047548230427	0.41729524517695715

Рис. 4. Приклад потоку даних для вибірки експериментальних даних

Загалом від визначення об'єму вибірки, тобто кількості точок, залежить забезпечення отримання результатів з заданою точністю [15]. Аналіз проведених експериментів показав, що для збільшення точності на порядок варто збільшити кількість експериментів у 100 разів.

Обговорення результатів та перспективи подальшого розвитку досліджень. Для статистичної обробки результатів функціонування M/M/1/0 з підвищенням інтенсивності надходження вимоги за умови незайнятості системи (ГРІБ в умовах навантаження) використовується одномірний статистичний аналіз.

В рамках статистичного аналізу використано бібліотеки Python statistics, яка працює зі списком даних та має відносно невелику кількість статистичних функцій, однак її часто використовують поряд з чистим Python, Numpy, яка вирізняється потужністю обробки, Pandas, яка опрацьовує дані, представлені об'єктом, що пришвидшує опрацювання та Matplotlib, Plotly для візуалізації даних.

Статистичний аналіз стаціонарних характеристик досліджуваної моделі включає використання центральних метрик для дослідження центрів концентрації даних та метрики оцінки варіативності даних, які аналізують розкид значень. На рис. 5 подано розрахунки основних метрик для 1D-послідовності значень стаціонарних характеристик (див. рис. 4) M/M/1/0 з підвищенням інтенсивності надходження вимоги за умови незайнятості системи.

measures of central tendency	experimental dataset π_0	experimental dataset π_1
Mean	0.6077101348848145	0.3922898651151855
Harmonic mean	0.6055925660892444	0.3890531374357461
Geometric mean	0.6066554534951076	0.3906703712616243
Median	0.610170863966655	0.3898291360333451
median_low	0.609342318756844	0.3890005908235341
median_high	0.6109994091764659	0.39065768124315614
Mode	0.6022455267485098	0.39775447325149005

measures of variability	experimental dataset π_0	experimental dataset π_1
Variance	0.001284260209640262	0.0012842602096402613
Standard deviation	0.03583657642186627	0.035836576421866266
Skewness	-0.1334056932198437	0.1334056932198432
Percentiles 5%	0.5525206305931332	0.3387669735083679
Percentiles 95%	0.661233026491632	0.4474793694068666
Ranges	0.17670516260187918	0.17670516260187924

Рис. 5. Розрахунки центральних метрик та метрики оцінки варіативності даних для вибірки даних роботи ГРІБ з навантаженням

Графічно вибірка для отримання висновків про спостереження представлена набором точок з даних імітаційного експерименту, границею набору теоретичних даних стаціонарних характеристик, які мають сталу величину на сітці кількості проведених експериментів по відношенню до отриманих значень (рис. 6). Множина даних демонструє викиди, які пов'язані зі зміною поведінки досліджуваної системи. На базі отриманих даних відповідними лініями позначені середні арифметичне, геометричне, гармонічне, медіана та мода.

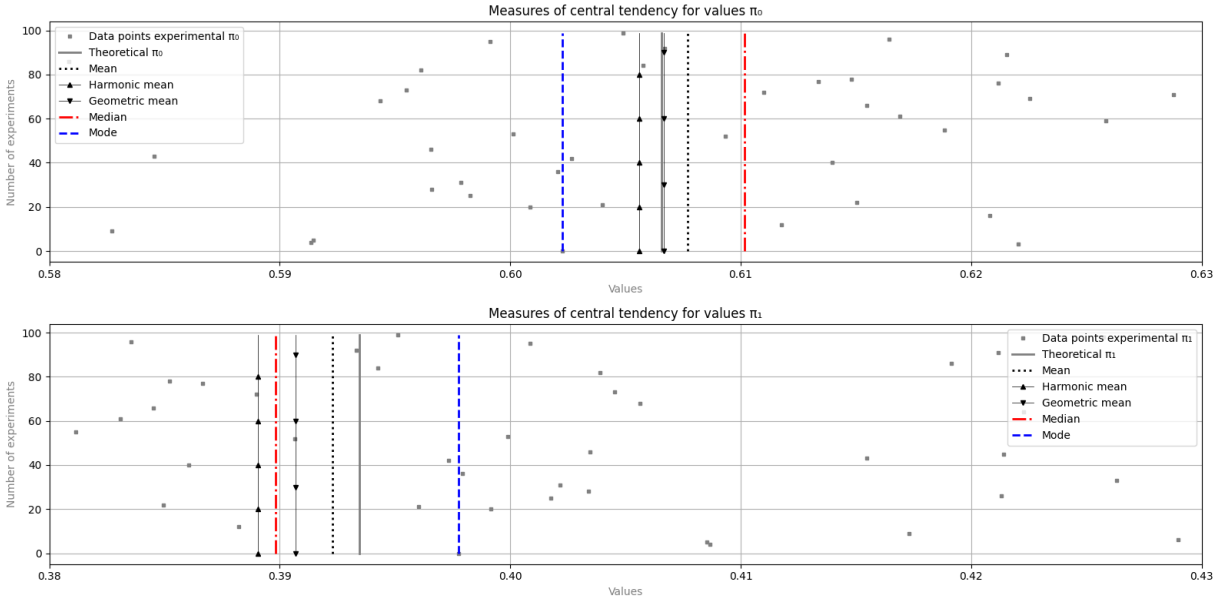

Рис. 6. Графічні дані розрахованих центральних метрик для вибірки даних проведених імітаційних експериментів

Середнє геометричне значення практично співпадає з теоретичним значенням для π_0, а для π_1 найближчим до теоретичного виявилось середнє арифметичне значення. Найбільше від середнього арифметичного відрізняється середнє гармонічне середнє. Воно розраховується як обернена величина від середнього значення обернених величин всіх елементів набору даних та може використовуватись як сукупний показник оцінки алгоритмів та систем, у випадку обох значень аналітично прогнозованих стаціонарних характеристик системи.

Основна відмінність між поведінкою середнього значення та медіани пов'язана з викидами чи екстремальними значеннями, зокрема медіана є менш чутливою до викидів. Вона, ділить ряд значень на дві частини, в кожній із яких розміщується однакова кількість одиниць сукупності. Для даних імітаційного експерименту значення медіани для π_0 перевищує значення середнього, для π_1 є меншим.

З метою визначення моди у випадках, коли кожне експериментальне значення буде зустрічатися один раз, зазвичай застосовують дискретизацію даних при побудові гістограм, після якої значення частоти виникнення визначається інтервалами, де інтервалу призначається значення його середньої точки. Мода, як правило, є тим значенням, де гістограма досягає свого максимуму [11]. У випадку наших вибірок результат подібної процедури продемонстрований на рис. 5.

На рис. 7 представлені кореляційні дані у вигляді діаграми розмаху для розсіяння числових даних.

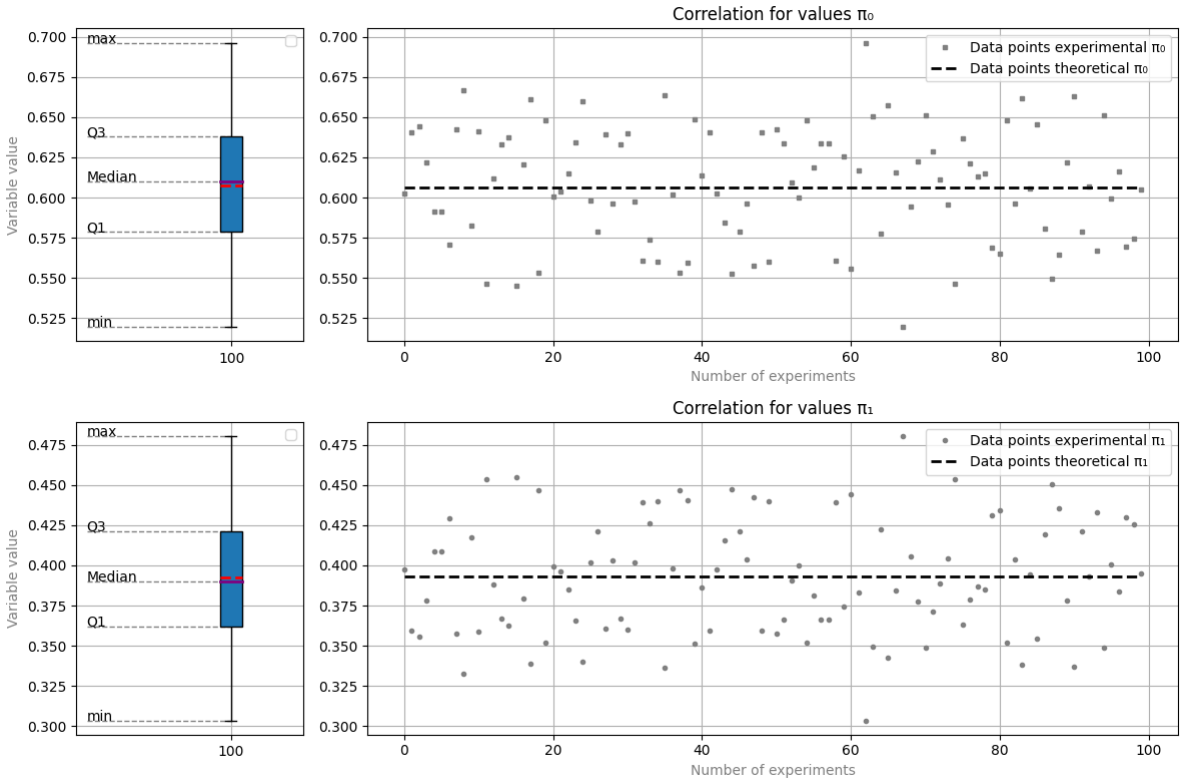

Рис. 7. Графічні дані розрахованих метрики оцінки варіативності даних для вибірки даних проведених експериментів

В правій стороні рис. 7 подано приклад набору точок, отриманих в імітаційному експерименті для π_0 та π_1 відповідно. Пунктиром позначено стаціонарне значення, до якого будуть наближатися статичні результати.

На основі даних розрахунків рис. 5 та набору точок, описаного вище, в лівій стороні рис. 7 продемонстровані вертикальні діаграми (для вибірок випадкових даних π_0 та π_1 для 100 імітаційних експериментів), які пропонують візуалізацію групи числових даних через їх квантилі, зокрема, додатково розрахованих при побудові, першого (Q_1 – зразок 25-го процентиля) та третього (Q_3 – зразок 75-го процентиля). Відмічені значення другого квантиля (0,5 процентиль), тобто медіани, а також, виділено пунктиром червоного кольору числове значення середнього. Границі розмаху варіацій показників у лівій частині рис.7 (діаграми розмаху – виділено синім кольором) представлені відповідно мінімальним та максимальним значенням відповідного показника. Відстані між різними частинами варіацій вказують на ступінь дисперсії (рис. 5), асиметрію (рис. 5) в даних та наявність викидів.

Аналізуючи розраховані метрики оцінки варіативності даних, продемонстрованих на рис. 5 варто також зазначити:

– значення дисперсії показує наскільки далеко розміщені точки даних від середнього значення, зокрема дані стаціонарних характеристик особливо не різняться і мають відхилення порядку 10^{-1};

– стандартне відхилення пов'язане з оцінкою дисперсії і аналогічно попередньому особливо не різняться для стаціонарних характеристик;

– коефіцієнт відхилення вказує на асиметрію вибірки даних, причому для даних π_0 він має від'ємне значення, що свідчить про наявність домінантного об'єму даних з лівого боку, а для π_1 він має додатне значення, тобто наявний довший хвіст даних з правого боку.

Проведене імітаційне моделювання засвідчило можливість врахування статистичних закономірностей кібератак на ефективність роботи ГРІБ. Це дозволяє запропонувати такий метод для розрахунку очікуваної ефективності функціонування ГРІБ в умовах зміни статистичних характеристик кібератак.

Спочатку знаходяться такі характеристики, використовуючи загальну базу даних для кібератак що вже відбулися.

1. Розраховуються статистичні характеристики для часових проміжків між початками окремих інцидентів під час кібератаки. В якості основної характеристики можна вибрати середнє значення часу очікування між окремими інцидентами, що дозволить розрахувати числове значення показника λ.

2. Розраховуються статистичні показники для наростання інтенсивності окремих інцидентів під час кібератаки, що дозволить отримати показник $\alpha\lambda$.

3. Розраховуються статистичні показники для часу опрацювання окремого інциденту під час кібератаки μ.

4. Формується класифікація кібератак. Така класифікація залежить від поставленої перед ГРІБ задачі. В цій класифікації обов'язково враховується наявність наростання інтенсивності окремих інцидентів під час кібератаки.

5. Формується база даних для показників $\alpha\lambda$, λ і μ та їх статистичних характеристик для кожного із класів кібератак.

База даних, сформована на етапі 5, може бути використана для оптимізації захисту від кібератак. Для нього необхідно здійснити такі кроки.

Використовуючи інформацію про клас кібератаки та статистичні показники $\alpha\lambda$ та λ для цього класу, здійснюється імітаційне моделювання. В результаті даного імітаційного моделювання знаходяться статистичні характеристики (див. рис. 5) для очікуваної ефективності діяльності ГРІБ (яка характеризується показником μ) в умовах даного класу кібератак. Наприклад, середні значення та середньоквадратичне відхилення. Це дозволяє здійснювати прогноз щодо вибору такої групи реагування, яка дозволить забезпечити потрібний рівень захисту від кібератак, що відносяться до заданого класу.

Нарешті, відслідковуючи в процесі розгортання поточної кібератаки статистичні характеристики інцидентів, можна вже в процесі проведення кіберзахисту замінювати одну ГРІБ на іншу, статистичні характеристики якої будуть забезпечувати потрібний рівень захисту.

Таким чином, діяльність по оптимізації організації захисту від кібератак із наростанням інтенсивності може бути коротко описана такою послідовністю: класифікація кібератак → аналіз статистичних характеристик інцидентів під час кібератаки → аналіз статистичних характеристик окремих ГРІБ → імітаційне моделювання для розрахунку статистичних характеристик ефективності ГРІБ для окремих класів кібератак → прогноз очікуваного рівня захисту (в залежності від характеристик кібератаки та ГРІБ) → попередній прогноз вибору ГРІБ для досягнення потрібного рівня захисту → моніторинг статистичних характеристик поточної кібератаки → оптимізація вибору ГРІБ для досягнення потрібного рівня захисту.

Висновки з даного дослідження і перспективи подальших розвідок у даному напрямі

Зростання кількості та впливу кібертатак з метою дестабілізації суспільства вимагає модифікації діяльності ГРІБ передовсім у напрямку врахування стохастичності фрагментів кібератаки та наростання її інтенсивності із часом.

В статті здійснено імітаційне моделювання діяльності ГРІБ з урахування статистичних характеристик як розгортання кіберінциденту у часі, так і діяльності самої групи. Результати проведеного моделювання свідчать, що медіанне значення є найбільш ефективним для моделювання середніх (стаціонарних) значень характеристик ефективності діяльності ГРІБ.

Описано метод для використання отриманих результатів та імітаційного моделювання взагалі для оптимізації діяльності ГРІБ.

Література

1. Звіт за результатами роботи системи виявлення вразливостей і реагування на кіберінциденти та кібератаки. Оперативний центр реагування на кіберінциденти Державного центру кіберзахисту та протидії кіберзагрозам. 2021. URL: https://cert.gov.ua/files/pdf/SOC_Annual_Report_2022.pdf (дата звернення: 12.01.2022).

2. Олексюк Л. В. Кращі практики управління кібербезпекою. Оглядовий звіт. 2019. 130 с. URL: file:///C:/Users/User/Downloads/ Report_on_Cybersecurity_04.pdf (дата звернення: 12.01.2022).

3. Про рішення Ради національної безпеки і оборони України від 14 травня 2021 року «Про Стратегію кібербезпеки України»: Указ Президента України; Стратегія від 26.08.2021 № 447/2021 // База даних «Законодавство України» / Верховна Рада України. URL:https://zakon.rada.gov.ua/go/447/2021 (дата звернення: 12.01.2022).

4. Carley K.M.: Social cybersecurity: an emerging science. Computational and Mathematical Organization Theory. 2020. V. 26(4). P. 365–381.

5. Eltantawy N., Wiest J. The Arab Spring. Social Media in the Egyptian Revolution: Reconsidering resource mobilization theory. *Int. J. Commun*. 2011. V. 5. P.1207–1224.

6. Wolfsfeld G, Segev E, Sheafer T. Social media and the Arab Spring: Politics comes first. Int. J. Press/Politics. 2013. V. 18. P. 115–137.

7. Toraman C., Şahinuç F., Yilmaz E. H. What Happened in Social Media during the 2020 BLM Movement? An Analysis of Deleted and Suspended Users in Twitter. 2021. arXiv:2110.00070v1 [cs.SI].

8. Ross B., Pilz L., Cabrera B.etal. Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks. European Journal of Information Systems. 2019. V. 28(4). P. 394–412.

9. Вентцель Е. С., Овчаров Л. А. Теория случайных процессов и её инженерные приложения. М. : Наука, 1991. 384с.

10. Голоскоков О. Є., Голоскокова А. О., Мошко Є. О. Основи теорії експоненціальних систем масового обслуговування. Харків : НТУ "ХПІ", 2017. 312 с.

11. Маталыцкий М., Хацкевич Г. Теория вероятности и математическая статистика. ЛитРес, 2021. 350 с.

12. Кельтон В., Лоу А. Имитационное моделирование. Классика CS. Киев : Издательская группа BHV, 2004. 847 с.

13. Копей В. Б. Мова програмування Python для інженерів і науковців. Івано-Франківськ : ІФНТУНГ, 2019. 272 с.

14. Ложковський А. Г. Теорія масового обслуговування в телекомунікаціях. Одеса : ОНАЗ ім. О.С. Попова, 2010. 112 с.

15. Кононюк А. Е. Обобщенная теория моделирования. Начала. К1. Ч.3. К.4: "Освіта України", 2012. 568 с.

References

1. Report on the results of the system of vulnerability detection and response to cyber incidents and cyber attacks. Operational Center for Cyber Incident Response of the State Center for Cyber Defense and Counteraction to Cyber Threats. 2021. Available at: https://cert.gov.ua/files/pdf/SOC_Annual_Report_2022.pdf (Accessed: 12 January 2022). (In Ukrainian).

2. Oleksyuk L.V. Best practices of cybersecurity management. Review report. 2019. 130 p. Available at: file:///C:/Users/User/Downloads/ Report_on_Cybersecurity_04.pdf. (Accessed: 12 January 2022). (In Ukrainian).

3. On the decision of the National Security and Defense Council of Ukraine of May 14, 2021 "On the Cyber Security Strategy of Ukraine": Decree of the President of Ukraine; Strategy from 26.08.2021 № 447/2021 // Database "Legislation of Ukraine". The Verkhovna Rada of Ukraine. Available at: https://zakon.rada.gov.ua/go/447/2021. (Accessed: 12 January 2022). (In Ukrainian).

4. Carley K.M.: Social cybersecurity: an emerging science. *Computational and Mathematical Organization Theory*, 2020, vol. 26(4), pp. 365–381.

5. Eltantawy N., Wiest J. The Arab Spring. Social Media in the Egyptian Revolution: Reconsidering resource mobilization theory. *Int. J. Commun*, 2011, vol. 5, pp.1207–1224.

6. Wolfsfeld G, Segev E, Sheafer T. Social media and the Arab Spring: Politics comes first. *Int. J. Press/Politics*, 2013, vol. 18, pp. 115–137.

7. Toraman C., Şahinuç F., Yilmaz E. H. What Happened in Social Media during the 2020 BLM Movement? An Analysis of Deleted and Suspended Users in Twitter. 2021. Available at: arXiv:2110.00070v1 [cs.SI]. (Accessed: 12 January 2022).

8. Ross B., Pilz L., Cabrera B.et al. Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks. *European Journal of Information Systems*, 2019. vol. 28(4), pp. 394–412.

9. Ventsel' E.S., Ovcharov L.A. Teoriya sluchaynykh protsessov i yeyo inzhenernyye prilozheniya [Theory of random processes and its engineering applications]. Moscow, Nauka, 1991. 384 p.

10. Holoskokov O.YE., Holoskokova A.O., Moshko YE.O. Osnovy teoriyi eksponentsial′nykh system masovoho obsluhovuvannya [Fundamentals of the theory of exponential queuing systems]. Kharkiv : NTU "KhPI",, 2017. 312 p.

11. Matalytskiy M., Khatskevich G. Teoriya veroyatnosti i matematicheskaya statistika [Probability theory and mathematical statistics]. LitRes, 2021. 350 p.

12. Kel'ton V., Lou A. Imitatsionnoye modelirovaniye. Klassika CS [Simulation modeling. CS classic]. Kiyev : Izdatel'skaya gruppa BHV, 2004. 847 p.

13. Kopey V.B. Mova prohramuvannya Python dlya inzheneriv i naukovtsiv [Python programming language for engineers and scientists]. Ivano-Frankivs'k : IFNTUNH, 2019. 272 p.

14. Lozhkovs'kyy A.H. Teoriya masovoho obsluhovuvannya v telekomunikatsiyakh [Theory of queuing in telecommunications]. Odesa : ONAZ im. O.S. Popova, 2010. 112 p.

15. Kononyuk A.Ye. Obobshchennaya teoriya modelirovaniya. Nachala. K1. CH.3. K.4 [Generalized modeling theory]. "Osvita Ukrayiny", 2012. 568 p.

Рецензія/Peer review : 05.11.2021 Надрукована/Printed :30.12.2021