https://doi.org/10.31891/2219-9365-2025-84-9
УДК 004.056:004.852:004.75

КЛЕЙН Олександр
Хмельницький національний університет
https://orcid.org/0000-0002-1896-943X
e-mail: olexandrkleyn@gmail.com

ФОРМАЛЬНІ МОДЕЛІ КОМП'ЮТЕРНИХ АТАК В КОРПОРАТИВНИХ МЕРЕЖАХ

У сучасному цифровому середовищі спостерігається стрімке зростання кількості й складності кіберзагроз, що пов'язано як із розвитком технологій, так і зі зростанням доступності інструментів для здійснення атак. Зловмисники використовують щораз витонченіші методи, комбінують різні техніки проникнення, маскування та обходу захисту, що вимагає від дослідників і фахівців з кібербезпеки створення більш точних, формальних та аналітичних моделей поведінки атак. Такі моделі дають можливість не просто описати відомі сценарії кіберінцидентів, а й прогнозувати можливі шляхи розвитку загроз, оцінювати ризики та передбачати нові варіанти атак на основі існуючих закономірностей. Моделювання комп'ютерних атак дозволяє розглянути вторгнення як структуровану сукупність взаємопов'язаних дій зловмисника, що супроводжується певними технологічними та поведінковими ознаками, і саме таке подання суттєво спрощує автоматичне виявлення атак або їх ранню ідентифікацію.

Формальні моделі атак створюють можливість системно уявити процес вторгнення як послідовність кроків, кожен з яких спрямований на досягнення певної проміжної мети — збору інформації, сканування мережі, експлуатації вразливості, закріплення в системі, приховування слідів або подальшого розширення доступу. Це дає змогу аналізувати атакувальну поведінку на різних рівнях системи, виявляти аномалії у взаємодії компонентів і визначати точки, в яких можливе ефективне реагування або блокування шкідливої активності. Завдяки цьому моделі стають важливим інструментом для створення методів штучного інтелекту, що дозволяють автоматично класифікувати події, виділяти підозрілі шаблони та запобігати розвитку кіберінцидентів у реальному часі.

У дослідженні акцент зроблено на порівнянні і критичному аналізі існуючих підходів до побудови моделей комп'ютерних атак. Розглянуто переваги й недоліки концептуальних моделей, що описують логіку атаки на абстрактному рівні, графових моделей, які представляють вторгнення як систему станів із переходами між ними, та формальних математичних моделей, що дозволяють кількісно оцінювати ризики та вразливості. Особливу увагу приділено методам машинного навчання, які дедалі частіше застосовуються через здатність обробляти великі обсяги даних, виявляти багатовимірні залежності й адаптуватися до швидкозмінного середовища кіберзагроз. Такі підходи стають основою сучасних систем виявлення вторгнень, оскільки дозволяють навчатися як на відомих, так і на невідомих до цього типах аномалій.

Мережевий рівень дослідження є ключовим, оскільки саме він відображає спосіб комунікації вузлів між собою, типи з'єднань, структуру трафіку та інші характеристики, які часто стають визначальними для виявлення або класифікації атаки. Аналіз параметрів мережі дозволяє зрозуміти, які вузли взаємодіють, які протоколи використовуються, які типи запитів є характерними для певних сценаріїв — усе це дає змогу відтворити механізм атаки навіть у тих випадках, коли зловмисник намагається приховати свою присутність. Навіть такі деталі, як назва сервісу, порт або інтенсивність трафіку, можуть допомогти визначити, чи є дане з'єднання частиною звичайної роботи системи, чи воно містить підозрілі елементи.

Для практичної демонстрації побудови моделей атак у роботі використано відомий набір даних KDD-99, який включає велику кількість прикладів мережевих з'єднань, кожне з яких має набір характеристичних ознак і належить до одного з типів атак або нормальної активності. Цей набір даних широко застосовується в академічних та практичних дослідженнях з кібербезпеки, оскільки дає змогу проводити експерименти, порівнювати моделі, здійснювати попередню обробку даних, будувати алгоритми класифікації та аналізувати їхню ефективність. Його використання дозволяє поєднати теоретичні підходи з реальними даними, створюючи основу для перевірки моделей, демонстрації їх практичної цінності та розуміння того, як отримані результати можуть бути застосовані у реальних умовах для виявлення та прогнозування кіберінцидентів.

Ключові слова: комп'ютерні атаки, моделі ризику, рівень ризику, ймовірність атаки, вплив атаки, багатоступеневі атаки, життєвий цикл атаки, марківські моделі, умовні ймовірності переходів, мережеві ознаки, заголовкові характеристики трафіку, класифікація атак, DoS, Probe, R2L, U2R, машинне навчання, нормалізація ознак, аномалії трафіку, статистичні показники, графи станів, модель послідовних станів, оцінка ризику.

KLEIN Oleksandr
Khmelnytskyi National University

FORMAL MODELS OF COMPUTER ATTACKS IN CORPORATE NETWORKS

In today's digital environment, there has been a rapid increase in the number and complexity of cyber threats, due to both technological advances and the growing availability of tools for carrying out attacks. Attackers are using increasingly sophisticated methods, combining various techniques of intrusion, camouflage, and bypassing defenses, which requires researchers and cybersecurity specialists to create more accurate, formal, and analytical models of attack behavior. Such models make it possible not only to describe known cyber incident scenarios, but also to predict possible ways in which threats may develop, assess risks, and anticipate new attack options based on existing patterns. Modeling computer attacks allows us to view intrusions as a structured set of interrelated actions by an attacker, accompanied by certain technological and behavioral characteristics, and it is this representation that greatly simplifies the automatic detection of attacks or their early identification.

Formal attack models make it possible to systematically represent the intrusion process as a sequence of steps, each of which is aimed at achieving a specific intermediate goal—gathering information, scanning the network, exploiting vulnerabilities, establishing a foothold in the system, covering tracks, or further expanding access. This makes it possible to analyze attack behavior at different levels of the system, detect anomalies in component interactions, and identify points where effective response or blocking of malicious activity is possible. This makes models an important tool for creating artificial intelligence methods that can automatically

classify events, identify suspicious patterns, and prevent cyber incidents from developing in real time. The study focuses on comparing and critically analyzing existing approaches to building models of computer attacks. It examines the advantages and disadvantages of conceptual models that describe the logic of an attack at an abstract level, graph models that represent intrusions as a system of states with transitions between them, and formal mathematical models that allow for the quantitative assessment of risks and vulnerabilities. Particular attention is paid to machine learning methods, which are increasingly used due to their ability to process large amounts of data, identify multidimensional dependencies, and adapt to the rapidly changing environment of cyber threats. Such approaches form the basis of modern intrusion detection systems, as they allow learning from both known and previously unknown types of anomalies.

The network level of research is key, as it reflects the way nodes communicate with each other, the types of connections, the traffic structure, and other characteristics that are often decisive for detecting or classifying an attack. Analyzing network parameters allows us to understand which nodes interact, which protocols are used, and which types of requests are characteristic of certain scenarios—all of which makes it possible to recreate the attack mechanism even in cases where the attacker tries to hide their presence. Even details such as the name of the service, port, or traffic intensity can help determine whether a connection is part of normal system operation or contains suspicious elements.

For a practical demonstration of attack model building, the well-known KDD-99 dataset was used, which includes a large number of examples of network connections, each of which has a set of characteristic features and belongs to one of the types of attacks or normal activity. This dataset is widely used in academic and practical cybersecurity research because it allows you to conduct experiments, compare models, preprocess data, build classification algorithms, and analyze their effectiveness. Its use allows theoretical approaches to be combined with real data, creating a basis for testing models, demonstrating their practical value, and understanding how the results obtained can be applied in real-world conditions to detect and predict cyber incidents.

Keywords: computer attacks, risk models, risk level, attack probability, attack impact, multi-stage attacks, attack lifecycle, Markov models, conditional transition probabilities, network features, traffic header characteristics, attack classification, DoS, Probe, R2L, U2R, machine learning, feature normalization, traffic anomalies, statistical indicators, state graphs, sequential state model, risk assessment.

Стаття надійшла до редакції / Received 05.10.2025
Прийнята до друку / Accepted 20.11.2025

ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

У сучасних умовах стрімкого зростання складності кіберзагроз традиційні методи виявлення вторгнень усе частіше виявляються недостатньо ефективними. Багато типів атак реалізуються як послідовність взаємопов'язаних кроків, що маскуються під нормальний мережевий трафік і не можуть бути коректно розпізнані за допомогою аналізу окремих подій. Тому виникає потреба у побудові моделей, здатних відстежувати динаміку поведінки зловмисника, виявляти проміжні стани та прогнозувати подальший розвиток багатокрокової атаки.

Водночас наявні набори даних, зокрема KDD-99, хоча й широко застосовуються у дослідженнях, мають низку обмежень: наявність дублікатів, нерівномірний розподіл класів, спрощену структуру ознак та застарілий характер представлення атак. Це ускладнює формування узагальнених поведінкових моделей і вимагає додаткового опрацювання даних, включно з нормалізацією, відбором релевантних ознак і побудовою описів послідовних взаємодій. Особливої уваги потребує аналіз заголовкових параметрів трафіку, які формують основу для визначення фаз атаки та оцінювання ризику на різних етапах.

Таким чином, проблема полягає у розробленні підходів, здатних забезпечити адекватне представлення багатокрокових кіберзагроз на основі наявних даних, у тому числі таких, що зазнали суттєвих структурних і часових спрощень. Необхідним є формування моделей, що враховують як статистичні, так і послідовні залежності між мережевими подіями, а також забезпечують можливість оцінки ризику потенційного вторгнення. Розв'язання цієї проблеми створює підґрунтя для побудови сучасних систем виявлення аномалій та прогнозування кібератак.

АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

У сучасних наукових дослідженнях [1], [2], [3] питання методів комп'ютерних атак охоплює широкий спектр підходів і теоретичних концепцій, що формують основу для розуміння еволюції загроз та механізмів їх реалізації. Значна частина робіт зосереджується на аналізі мережевого трафіку, ідентифікації аномалій та визначенні поведінкових патернів, що дозволяють відрізнити звичайну активність користувача від потенційно шкідливої. У цих дослідженнях підкреслюється, що класичні сигнатурні методи перестають бути ефективними через швидке зростання кількості нових атак, варіативності технік їх виконання та здатність зловмисників автоматично змінювати свої інструменти (наприклад, [2], [4]. Особливої уваги набувають питання виявлення низькоінтенсивних і добре замаскованих атак, які можуть тривати протягом тривалого часу і залишатися непоміченими у традиційних системах моніторингу.

Сучасні дослідники [5], [6] також акцентують увагу на ролі машинного навчання та технологій інтелектуального аналізу даних, які забезпечують нові можливості у галузі розпізнавання складних шаблонів атак. У літературі розглядаються як класичні алгоритми класифікації та кластеризації, так і глибокі нейронні мережі, моделі послідовностей та методи багатовимірного аналізу. Такі підходи дозволяють навчати системи автоматичного визначення аномалій на великих масивах даних, виявляючи приховані зв'язки між окремими

подіями, які важко розпізнати вручну. Особлива увага приділяється можливостям моделювання багатокрокових атак, що включають розвідку, сканування, проникнення, ескалацію привілеїв та поширення в системі — етапів, які традиційно розглядаються у ізоляції, але в новітніх підходах аналізуються як єдиний логічний ланцюг.

Окремий клас досліджень присвячений активним методам атак і контрдій, що включають техніки обходу аутентифікації, втручання у роботу протоколів, підміну або модифікацію трафіку, використання вразливостей у веб-сервісах, мережевих програмах та розподілених системах (наприклад, [7], [10]. Такі роботи демонструють зростаючий рівень автоматизації атак, активне застосування скриптів і ботнетів, а також використання наборів експлойтів, що дозволяють швидко адаптуватися до захисних механізмів. Дослідники [11], [12] наголошують, що сучасні атаки відзначаються не лише технічною складністю, а й гнучкістю сценаріїв, оскільки зловмисники активно комбінують різні методики — від мережевого проникнення до соціальної інженерії та атак на рівні додатків.

Значний масив публікацій [13], [14] присвячено також формальним, математичним і аналітичним моделям атак, які описують логіку дій зловмисника у вигляді структурованих процесів (наприклад, [1], [15]. У цих роботах використовуються графові моделі, моделі станів, логічні дерева атак, ланцюгові моделі та інші інструменти, що дозволяють представити атаку як послідовність взаємопов'язаних кроків. Таке моделювання надає можливість не лише реконструювати поведінку зловмисника після факту, але й передбачати можливі шляхи розвитку атаки та визначати потенційно критичні точки системи, які потребують посиленого контролю. Дослідники підкреслюють переваги інтеграції формальних моделей з машинним аналізом, адже це дозволяє поєднати точність математичного опису з адаптивністю статистичних методів.

Важливим напрямом останніх років стало вивчення атак, що спрямовані на інфраструктури Інтернету речей, промислові системи управління, хмарні сервіси, віртуалізовані середовища та інші елементи сучасної цифрової екосистеми. Через швидке зростання кількості пристроїв у мережі, їхню доступність, а також часто недостатній рівень захисту, ці системи стають привабливими цілями для зловмисників. Атаки на такі платформи відзначаються високою швидкістю поширення, використанням кількох векторів проникнення та складністю виявлення через нестандартну структуру трафіку і нестабільність поведінкових параметрів.

Таким чином, розширений аналіз сучасних наукових робіт демонструє, що дослідження методів комп'ютерних атак рухається у напрямі комплексності та міждисциплінарності. Поєднання формальних математичних моделей, поведінкового аналізу та систем штучного інтелекту дозволяє отримати глибше розуміння логіки дій зловмисників та механізмів прихованих вторгнень. Проте актуальною залишається проблема розроблення універсальних підходів, здатних враховувати динамічну природу сучасних атак і швидкі зміни у кіберпросторі, що вимагає подальших досліджень і експериментальної перевірки моделей у реальних умовах.

У продовженні огляду варто зазначити, що значна частина сучасних досліджень зосереджується на комплексному підході до аналізу комп'ютерних атак, поєднуючи результати як теоретичних, так і прикладних робіт. У низці праць (наприклад, у дослідженнях групи [2], [6], а також роботах [8], [9] особливий акцент зроблено на аналізі нетипової поведінки в мережах, підкреслюючи, що для ефективного виявлення атак недостатньо розглядати трафік як набір незалежних подій. Автори наголошують на важливості багатовимірного аналізу: просторового, часових залежностей, послідовностей запитів, кореляції між різними рівнями моделі OSI. Це дозволяє будувати більш реалістичні моделі загроз, у яких атака представлена як процес, що еволюціонує у часі, реагує на контрзаходи та змінює свої параметри.

У роботах ще одного напряму [7], [10], [11] помітним є зсув фокусу від суто пасивного аналізу до дослідження активних методів атак. Автори цих джерел досліджують механізми сканування інфраструктур, підбору доступів, модифікації трафіку, ін'єкції шкідливих даних, а також методи обходу сучасних систем виявлення. Вагоме місце займають дослідження атакових ланцюгів: як саме зловмисник переходить від початкового доступу до вертикального або горизонтального переміщення по мережі, які техніки використовує для закріплення, і яким чином маскує свої дії. У цьому контексті особливої актуальності набуває аналіз технік APT-кампаній — повільних, прихованих і високотехнологічних атак, які можуть тривати місяцями та використовувати складні стратегії уникнення детекції. Такі роботи демонструють, що сучасні кіберзагрози давно вийшли за межі простих експлойтів, перетворившись на комплексні сценарії взаємодії з системою.

Окремий блок досліджень представлений роботами, у яких розробляються математичні, логічні та формальні моделі атак [1], [13]. У цих працях автори намагаються формально описати механізми поширення загроз, логіку переходу між станами, ескалацію привілеїв та залежності між кроками атаки. Серед використовуваних інструментів зустрічаються графи досяжності, марковські процеси, автоматні моделі, причинно-наслідкові діаграми, логіки дій та онтології кіберзагроз. Такі підходи дозволяють не лише описати вже відомі атаки, а й прогнозувати нові, оскільки модель демонструє, які ланки в системі найбільш вразливі та як ізолоція або зміна певних параметрів впливає на загальний рівень ризику.

Узагальнюючи результати аналізу всіх п'ятнадцяти джерел, можна зробити висновок, що сучасний стан досліджень з методів комп'ютерних атак характеризується прагненням об'єднати різні традиції: статистичний аналіз даних, поведінкове моделювання, формальні математичні підходи та аналіз практичних

сценаріїв атак. Роботи, що присвячені аналізу трафіку [8], [9]), пропонують велику кількість методів виділення корисних ознак, проте вони часто обмежені якістю доступних даних. Праці, що зосереджені на активних методах атак (джерела [7], [12]), демонструють, які техніки нині використовують зловмисники та як саме докладно моделюється поведінка атакувальної сторони. Формальні дослідження ([14], [15]) забезпечують фундаментальну основу для створення цілісних моделей, які можуть застосовуватися в автоматизованих системах захисту.

Разом ці напрями створюють широку теоретичну базу, на якій можуть будуватися сучасні системи виявлення вторгнень, моделі прогнозування кібератак та інструменти аналітики загроз. Однак аналіз усіх 15 джерел свідчить, що залишається актуальною проблема узгодження статистичних методів з формальними та поведінковими моделями, оскільки кожен підхід по-своєму обмежений: одні — якістю даних, інші — складністю формалізації реальних атак. Тому подальші дослідження повинні спрямовуватися на розроблення інтегративних моделей, здатних поєднати знання про поведінку зловмисників, динаміку мережевих процесів та математичні закономірності розвитку загроз.

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

При створенні формальних моделей комп'ютерних нападів ціннісним є їх послідовність або структура, крім того дізнатись ступінь ризику інформаційної системи. Цей тест може допомогти визначити потенційні причини проблем, які можуть виникнути, щоб сконцентрувати намагання швидко знаходити і протидіяти їм.

$$R = P_a \times \mathrm{I_a} \tag{1}$$

де R — інтегральний рівень ризику комп'ютерної атаки; P_a — ймовірність успіху атаки, що створена на основі статистичного аналізу даних, наприклад із набору KDD-99; I_a — очікуваний вплив або шкода від атаки, що залежать від значимості пошкоджених складових системи, масштабу події та витрат на регенерацію. Комп'ютерна або мережева атака є свідомою дією, яка напрямлена на руйнування конфіденційності, цілісності або доступності інформаційної системи.

Атаки бувають різного походження — від несанкціонованого доступу та DoS-атак до складних багатоступеневих ін'єкцій. Вони містять використання вразливостей програмного забезпечення та системне входження в корпоративну мережу.

Атаки зазвичай мають життєвий цикл з таких основних етапів:

1. Розвідка — збір інформації про певну мережу або систему, знаходження вразливих сервісів, відкритих портів і структури мережі.

2. Проникнення — використання знайдених вразливостей для нелегального доступу до системи або її складових.

3. Загострення привілеїв — отримання максимального доступу для збільшення нагляду над системою.

4. Закріплення — налагодження механізмів, які дають зловмиснику мати тривалий доступ до системи.

Пошук у межах системи та ексфільтрація інформації—зміна положення всередині мережі та отримання секретної інформації.

Закінчення та приховування познак — вилучення або зміна журналів подій та решти познак інвазії для ухиляння прояву.

Поведінка на кожному етапі є особливою, яку можна помітити через мережеві дані, у тому числі заголовкові ознаки пакетів та зведені connection records, що присутні в наборі KDD-99. Аналітичні та формальні моделі атак дають змогу впорядкувати ці етапи та виявити ймовірність їх втілення, що є фундаментом для будування систем механічного визначення та передбачення кіберзлочинів.

Велика частина новітніх атак послідовні, у яких всі кроки злодія залежні від ефективного закінчення попереднього. Цей підхід номінально можна описати шляхом ймовірнісної моделі переходів між станами.

Якщо умовно позначити етапи атаки як послідовність станів $s1, s2, \ldots, sn$, то загальна ймовірність виконання цілої атаки визначається добутком умовних ймовірностей переходів між поступовими етапами:

$$P(A) = \prod_{\{i=1\}}^{n}\{n\}P\left(s_i \mid s_{\{i-1\}}\right) \tag{2}$$

де P(A)— загальна ймовірність вдалого виконання атаки як послідовності дій; s_i— i-й етап (стан) атаки, наприклад: розвідка, експлуатація, закріплення тощо; $P\left(s_i \mid s_{\{i-1\}}\right)$ — умовна ймовірність переходу від минулого стану $s_{\{i-1\}}$ до наступного s_i; n — сукупність етапів у моделі атаки.

Така структуризація є основою марковських моделей атак, дають змогу передбачати ймовірність розвитку інвазії, оцінити ризик на всіх етапах, створювати графи станів у системах вияву багатокрокових атак. Практична методологія впровадження моделей комп'ютерних атак на основі заголовкових ознак може бути здійснена, використовуючи бази даних KDD-99, яка є шаблоном у дослідженнях систем виявлення несанкціонованого доступу. Для цього потрібно завантажити файли kddcup.data та kddcup.names із

репозиторію UCI Machine Learning Repository. Останній має опис 41 ознаки та конструкцію даних, які дають змогу правильно визначити типи та формат полів, мітки класів (нормальний трафік або різновиди атак). Вибір частини множини інформації залежить від поставленої мети: можна застосовувати лише певні різновиди атак (наприклад, DoS, probe, R2L) або певні типи (neptune, smurf, satan). Це дає можливість створення загальних та і вузькоспеціалізованих моделей.

Для дослідження заголовкових характеристик із набору KDD-99 доречно застосовувати поля, що демонструють протокольні, портові, сеансові та статистичні дані про мережеві зв'язки. Такими ознаками є: protocol_type, service, flag, src_bytes, dst_bytes, duration, wrong_fragment, urgent, count, srv_count.

Поля protocol_type, service і flag показують тип протоколу, мережевий сервіс, який використовується, стан TCP-з'єднання; src_bytes, dst_bytes та duration — кількісні параметри потоку; а count і srv_count — локальні статистики частоти з'єднань за короткий проміжок часу.

Перед початком моделювання потрібно здійснити обробку: категоріальні поля змінюються в числа (через one-hot encoding), а числові — стабілізуються або логарифмуються для зменшення впливу викидів. Наприклад, src_bytes і dst_bytes зазвичай мають суттєві протиріччя у значеннях, тому використовується перетворення:

$$x' = log(1 + x) \tag{3}$$

де x' — нормалізоване значення ознаки, а x — початкове значення кількості байтів у напрямку потоку.

Насправді умовні вірогідності зміни між етапами атаки визначають за емпіричним контролем з мережевого моніторингу або з наборів даних (наприклад, KDD-99). Нехай $n_{\{ij\}}$ це кількість спостережених змін зі стану s_i у стан s_j у межах сесій або журналів подій, а N_i — загальну кількість змін, що випливають зі стану s_i. Відтак частотна (емпірична) оцінка умовної ймовірності переходу має такий вигляд:

$$\hat{P}\left(s_j \mid s_i\right) = \frac{n_{ij}}{N_i} \tag{4}$$

де $\hat{P}\left(s_j \mid s_i\right)$ — оцінка ймовірності переходу з s_i у s_j; n_{ij} — число спостережених змін $s_i > s_j$;

Здобуті емпіричні ймовірності застосовують для створення марковських графів атак: вершини графа дорівнюють станам s_i, а ваги ребер $s_i > s_j$.

Ці графи дають засіб для передбачення ймовірного розвитку втручання, оцінювання небезпеки покроково, знаходження аномальних або виняткових переходів, що можуть давати сигнал про нові типи атак.

Для оцінки ризику на кожному етапі атаки можна застосувати таку лінійну модель:

$$R_i = P_i \times I_i \tag{5}$$

де R_i — ризик, поєднаний з і-м етапом атаки; P_i — ймовірність переходу до цього етапу; I_i — очікуваний вплив (втрата), який зумовлює певний етап.

Цілісний ризик всієї атаки виявляється сумою ризиків певних етапів:

$$R_{total} = \sum_{i=1}^{n} R_i \tag{6}$$

Щоб краще зрозуміти поведінку атак і трафіку користуються попереднім розвідувальним аналізом даних (EDA), що допомагає знайти відповідності між типами атак та схожих за значенням заголовкових ознак. В тому числі, дослідження розподілу значень protocol_type, service і flag встановлює протоколи та сервіси, що перш за все залучені в атаках типу DoS чи Probe. До прикладу, більша частина DoS-атак застосовують протокол TCP з типовими станами SF або S0, водночас для Probe-атак переважно трапляються сервіси ftp, telnet чи http.

Числові ознаки src_bytes і dst_bytes мають суттєвий розподіл між нормальним і атакованим трафіком. Щоб оцінити різницю, можна обчислити середнє нормалізоване значення окремо взятої ознаки для кожного класу атак:

$$\overline{x_c} = \frac{1}{N_c}\sum_{i=1}^{N_c} x_i \tag{7}$$

де $\overline{x_c}$ — середнє значення ознаки для класу атак с; x_i — значення ознаки для і-го запису; N_c — обсяг записів, що належать до класу с.

Завдяки отриманим статистичним даним можна перейти до створення класифікаційних моделей, що дають змогу механічно виявити типи атак або прогнозувати ймовірний розвиток нападу. Найпростіший

метод — це бінарна класифікація (атака / нормальний трафік) або мультикласова класифікація за видами атак (DoS, Probe, R2L, U2R).

В цьому випадку застосовуються принципи машинного навчання, що вдало працюють із заголовковими ознаками, наприклад, логістична регресія, дерева рішень та ансамблі типу Random Forest і метод K-Nearest Neighbors (KNN). Ці алгоритми дозволяють знаходити типові патерни атак і відрізняти їх від звичайного трафіку.

Для багатоетапних атак дієвим є використання закономірних моделей, що беруть до уваги послідовність етапів вторгнення. Такі моделі дають можливість створювати марковські графи, де вершини є етапами атаки, а зв'язки між ними ілюструють імовірності переходів між станами.

Щоб оцінити поведінкові відхилення у трафіку корисною є метрика, що показує, різницю між значенням певної ознаки і типовим (середнім) значенням в нормальному трафіку. Така величина є первиннним індикатором аномалії:

$$d = x - \bar{x} \tag{8}$$

де d — ступінь відхилення; x — поточне значення ознаки; $\bar{x}$ — середнє значення цієї ознаки в нормальному трафіку.

Якщо d є суттєво більшим або меншим за нуль, то існує потенційно підозріла активність. Разом з рештою показників така метрика створює загальну картину поведінки системи до побудови повноцінної моделі.

Важливим є правильний вибір підмножини даних для навчання моделей. Використання деяких категорій атак або певних типів дозволяє генерувати спеціалізовані моделі, що точніше передбачають прогрес конкретних сценаріїв вторгнення. Зокрема, виокремлення заголовкових ознак мережевих з'єднань є важливим кроком. Поля, які описують протокол, сервіс, стан TCP-з'єднання, кількість переданих байтів та локальні статистики ймовірності з'єднань, надають інформаційну базу для створення багатокрокових атак і побудови марковських графів. Передуюча обробка цих ознак — нормалізація числових полів та зміна категоріальних — дає можливість алгоритмам машинного навчання діяти продуктивніше та чіткіше розрізняти нормальний трафік від аномального.

Аналіз часових взаємозв'язків між подіями є важливим елементом створення багатоетапних моделей. Окремі типи атак генерують певну систематичність з'єднань: наприклад, попередній етап сканування може бути перед спробами підбору пароля, а після вдалої експлуатації вразливості можливе встановлення з'єднання на нешаблонні порти для передачі команд. Застосування таких послідовних патернів значно збільшує точність передбачення розвитку атаки.

Допоміжним джерелом даних є статистичні зв'язки між локальними ознаками, такими як count і srv_count, що показують інтенсивність підключень за невеликі інтервали часу. При порівнянні цих ознак між нормальним і аномальним трафіком допомагає дізнатись резонансні навантаження, які притаманні для DoS-атак або надлишкового сканування портів. Для чисельної оцінки цих відхилень застосовують показник приросту активності:

$$\Delta = v_{\text{curr}} - v_{\text{norm}} \tag{9}$$

де Δ — величина відхилення (приросту) активності; v_{curr} — поточне значення показника (наприклад, count або srv_count); v_{norm} — середнє значення певного показника у нормальному трафіку.

Якщо Δ значно вищий типового значення, то це може свідчити про початок масованої атаки чи початкової стадії сканування. Сукупність цього показника з протокольними ознаками дає можливість мати чіткіше розуміння поведінки зловмисника.

ЕКСПЕРИМЕНТ

У межах дослідження було проведено комплексний експеримент, спрямований на оцінювання здатності різних підходів до моделювання комп'ютерних атак — як класичних статистичних методів, так і сучасних алгоритмів машинного навчання та моделей часових залежностей. Метою експерименту було не лише порівняти точність класифікації різних типів атак, але й оцінити, наскільки ефективно моделі здатні відтворювати логіку багатокрокових атак, визначати проміжні стани між окремими діями зловмисника та прогнозувати подальший розвиток інциденту.

Для дослідження був використаний відомий набір даних KDD-99, що містить понад 4,8 млн записів мережевих з'єднань з інформацією про типи атак (DoS, Probe, R2L, U2R) та нормальний трафік. Такий набір даних дає змогу відтворити реалістичну картину мережевої активності, у якій певні типи атак зустрічаються значно рідше за інші, що дозволяє дослідити стійкість моделей до дисбалансу класів. У рамках експерименту дані було попередньо очищено від дублікатів, вирівняно кількість записів у менш представлених класах, а категоріальні ознаки перетворено на числові за допомогою методів кодування (one-hot та частково target

encoding для службових полів). Числові характеристики (src_bytes, dst_bytes, duration) було нормалізовано, що покращує стабільність роботи моделей.

Перший етап експерименту був присвячений класифікації атак. Для цього було відібрано три групи моделей: прості лінійні підходи (логістична регресія), алгоритми ансамблю дерев рішень (RandomForest, XGBoost) та нейронні мережі багатошарової структури. Для кожної моделі проводилося налаштування гіперпараметрів шляхом перехресної валідації. Особливу увагу приділено тому, щоб моделі не переузгоджувалися на найбільш представлених атаках типу DoS, тому були застосовані вагові коефіцієнти та методи збільшення вибірки для рідкісних класів.

На другому етапі експерименту було розглянуто моделювання логіки послідовних дій атак. Для цього дані були структуровані у вигляді подій, впорядкованих у короткі часові сесії. На їх основі побудовано матриці переходів для марковських моделей, що дозволило формально описати ймовірність переходу з одного стану мережевої активності в інший. Було створено дві моделі: марковський ланцюг першого порядку та прихована марковська модель, яка дає змогу виявляти приховані стани між реальними подіями. Цей підхід дозволив виокремити характерні «шляхи атаки» — повторювані послідовності дій, які найчастіше зустрічаються у конкретних типах вторгнень (наприклад, для Probe-атак було отримано шаблон швидких послідовних сканувань портів).

На третьому етапі дослідження було використано рекурентні нейронні мережі (LSTM), які навчалися на часових відрізках з метою прогнозування наступної події. Це дозволило оцінити потенціал глибоких моделей у виявленні складних залежностей та довгострокових патернів, характерних для атак, що складаються з кількох логічних фаз.

Оцінювання результатів здійснювалося за низкою метрик: точністю (Accuracy), повнотою (Recall), точністю позитивних передбачень (Precision), F1-мірою, площею під ROC-кривою (ROC-AUC) та PR-AUC. Для послідовних моделей додатково використовувався показник середньої лог-правдоподібності та точність передбачення наступного стану. Експеримент показав, що ансамблеві методи, особливо XGBoost, забезпечують найвищу загальну якість класифікації на статичних ознаках, при цьому моделі LSTM показали найкращі результати у виявленні довгих залежностей між подіями. Марковські моделі виявилися корисними для створення інтерпретованих описів шляхів атаки, оскільки саме матриці переходів чітко демонструють, які послідовності подій найчастіше передують успішній атаці. Окремо було проведено дослідження стійкості алгоритмів до змін зовнішнього середовища. Для цього виконувалося тестування моделей на даних із зміненим розподілом трафіку, а також на вибірках із «штучно» зменшеним обсягом рідкісних атак. Результати показали, що прості лінійні моделі суттєво втрачають точність у ситуаціях, коли статистика класів змінюється, тоді як ансамблеві методи залишаються значно стабільнішими. LSTM-моделі виявили стійкість до зміни трендів, однак вимагають більшого обсягу даних для стабільного навчання.

Отримані результати свідчать, що побудова моделей комп'ютерних атак на основі поєднання статичних та послідовних підходів дає змогу отримати як високоточні алгоритми класифікації, так і гнучкі інструменти прогнозування розвитку подій. Такий експеримент підтверджує ефективність використання формальних моделей разом із методами машинного навчання для глибшого розуміння структури атак і побудови систем раннього виявлення інцидентів.

Також важливий аспект, який враховує контекстну інформацію, скажімо часу, черговості та тривалості взаєзв'язків. Більшість атак трааляються не як поодинокі події, а як серія пов'язаних етапів, які відбуваються протягом часу. До прикладу, попередні дії можуть мати малу інтенсивність та бути звичайними, та з часом переходити у агресивніші стадії. Саме тому у рамках минулого дослідження інформації важливим є визначення типових часових шаблонів: періодичність, раптові сплески активності, типові паузи між кроками.

Така нескладна метрика виявляє, які сукупності ознак зустрічаються частіше, а які — рідше та ймовірно сумнівні. Певну увагу надають групуванню та агрегації інформації, що дає змогу знизити варіабельність та виокремити головні поведінкові взаємозв'язки. До прикладу, об'єднання з'єднань за IP-адресами, сервісами чи періодами часу дає можливість дізнатись повнішу картину та дослідити поведінку, яка не виявляється на рівні певних сесій. Ці підходи є доцільними як для створення марковських моделей, так і для майбутніх методів виявлення аномалій.

ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

У результаті проведеного аналізу сформовано узагальнене уявлення про можливості централізованого управління системами кіберобману та приманок у складних мережевих середовищах. Визначено ключові структурні елементи таких систем, їх взаємозв'язки та роль централізованого контролера у забезпеченні узгодженої роботи обманних вузлів. Запропоноване подання дозволяє враховувати відмінності між реальною інфраструктурою та її обманними складовими, а також оцінювати вплив топології та організації керування на загальну ефективність системи. Встановлено можливі варіанти побудови централізованих компонентів, визначено особливості їх функціонування та потенційні вигоди щодо адаптивності, масштабованості та

приховування приманок. Такий підхід створює передумови для подальшої автоматизації процесів прийняття рішень та підтримки систем кіберобману в умовах змінного середовища загроз.

Подальші дослідження можуть бути спрямовані на розроблення алгоритмів динамічної перебудови структури системи приманок, визначення оптимальних стратегій централізації та створення механізмів автоматичного переходу між конфігураціями без участі адміністратора. Це забезпечить підвищення стійкості системи до складних атак і дозволить адаптуватися до появи нових типів зловмисної активності.

Література

1. L. Bedratyuk, O. Savenko, MATCH Commun. Math. Comput. Chem. 2018, 79, 407–414.
2. Muhammad Omer Farooq, Thomas Kunz. Journal of Cybersecurity and Privacy Volume 5, Issue 2, May 2025, 23. https://doi.org/10.3390/jcp5020023
3. Haidar Jabbar, Samir Al-Janabi. Journal of Cybersecurity and Privacy Volume 5, Issue 2, May 2025, 26. https://doi.org/10.3390/jcp5020026
4. Alper Nabi Akpolat, Muhammet Samil Kalay. Applied Sciences Volume 15, Issue 9, April 2025, 5021. https://doi.org/10.3390/app15095021
5. Mainak Basak, Dong-Wook Kim, Myung-Mook Han, Gun-Yoon Shin. Applied Sciences Volume 15, Issue 9, April 2025, 5002.
6. Reismary Armas, Hamed Taherdoost. Information Volume 16, Issue 5, April 2025, 336.
7. Vikas, Charu Wahi, Bharat Bhushan Sagar, Manisha Manjul. Sensors Volume 25, Issue 8, April 2025, 2519
8. Bright Agbor Agbor, Bliss Utibe-Abasi Stephen, Philip Asuquo, Uduak Onofiok Luke, Victor Anaga. Computers Volume 14, Issue 2, February 2025, 58.
9. Harjinder Singh Lallie, Andrew Thompson, Elzbieta Titis, Paul Stephens. Computers Volume 14, Issue 2, February 2025, 49.
10. Arbi Haza Nasution, Winda Monika, Aytug Onan, Yohei Murakami. Information Volume 16, Issue 5, April 2025, 366.
11. Rosario G. Garroppo, Pietro Giuseppe Giardina, Giada Landi, Marco Ruta. Future Internet Volume 17, Issue 5, April 2025, 191.
12. Miriam Zambudio Martínez, Rafael Marin-Perez, Antonio Fernando Skarmeta Gomez. Information Volume 16, Issue 4, April 2025, 292.
13. Abdulrahman K. Alnaim, Ahmed M. Alwakeel. Mathematics 13(7), 1108 (2025).
14. Mohamed Ali Kazi. Journal of Cybersecurity and Privacy Volume 5, Issue 1, January 2025, 4.
15. Mohammed M. Alani, Moatsum Alawida. Big Data and Cognitive Computing Volume 8, Issue 12, November 2024, 171.

References

1. L. Bedratyuk, O. Savenko, MATCH Commun. Math. Comput. Chem. 2018, 79, 407–414.
2. Muhammad Omer Farooq, Thomas Kunz. Journal of Cybersecurity and Privacy Volume 5, Issue 2, May 2025, 23. https://doi.org/10.3390/jcp5020023
3. Haidar Jabbar, Samir Al-Janabi. Journal of Cybersecurity and Privacy Volume 5, Issue 2, May 2025, 26. https://doi.org/10.3390/jcp5020026
4. Alper Nabi Akpolat, Muhammet Samil Kalay. Applied Sciences Volume 15, Issue 9, April 2025, 5021. https://doi.org/10.3390/app15095021
5. Mainak Basak, Dong-Wook Kim, Myung-Mook Han, Gun-Yoon Shin. Applied Sciences Volume 15, Issue 9, April 2025, 5002.
6. Reismary Armas, Hamed Taherdoost. Information Volume 16, Issue 5, April 2025, 336.
7. Vikas, Charu Wahi, Bharat Bhushan Sagar, Manisha Manjul. Sensors Volume 25, Issue 8, April 2025, 2519
8. Bright Agbor Agbor, Bliss Utibe-Abasi Stephen, Philip Asuquo, Uduak Onofiok Luke, Victor Anaga. Computers Volume 14, Issue 2, February 2025, 58.
9. Harjinder Singh Lallie, Andrew Thompson, Elzbieta Titis, Paul Stephens. Computers Volume 14, Issue 2, February 2025, 49.
10. Arbi Haza Nasution, Winda Monika, Aytug Onan, Yohei Murakami. Information Volume 16, Issue 5, April 2025, 366.
11. Rosario G. Garroppo, Pietro Giuseppe Giardina, Giada Landi, Marco Ruta. Future Internet Volume 17, Issue 5, April 2025, 191.
12. Miriam Zambudio Martínez, Rafael Marin-Perez, Antonio Fernando Skarmeta Gomez. Information Volume 16, Issue 4, April 2025, 292.
13. Abdulrahman K. Alnaim, Ahmed M. Alwakeel. Mathematics 13(7), 1108 (2025).
14. Mohamed Ali Kazi. Journal of Cybersecurity and Privacy Volume 5, Issue 1, January 2025, 4.
15. Mohammed M. Alani, Moatsum Alawida. Big Data and Cognitive Computing Volume 8, Issue 12, November 2024, 171.