DOI: https://doi.org/10.30837/IST-2024.P1.174

Застосування алгоритмів машинного навчання для верифікації та валідації інформаційних систем через виявлення аномалій

Олександр Неліпа[1*] та Надія Калита[1]

[1] *Харківський національний університет радіоелектроніки, Проспект Науки, 14, Харків, 61166, Україна*

Анотація
У статті досліджується інтеграція алгоритмів машинного навчання в процеси верифікації та валідації інформаційних систем. У роботі розглядається взаємозв'язок аномалій із верифікацією та валідацією як ідентифікаторів дефектів та акцентується увага на обмеженнях традиційних методів виявлення аномалій.

Ключові слова
Функціональний дефект, виявлення аномалій, верифікація, валідація, машинне навчання.

1. Вступ

Інформаційні системи є основою критично важливих операцій у різних галузях, зокрема у фінансових послугах, охороні здоров'я, виробництві та державному секторі. Забезпечення надійності, безпеки та операційної стабільності цих систем є надзвичайно важливим, оскільки аномалії в таких середовищах можуть призвести до серйозних наслідків, таких як витоки даних [1], фінансові збитки [2] та переривання надання послуг [3]. Виявлення аномалій є аналітичним процесом, спрямованим на ідентифікацію точок даних, об'єктів або подій, що значно відхиляються від типових моделей або норм [4, 5]. Ця методологія відіграє критичну роль в автоматизованих системах моніторингу, особливо у виявленні потенційно шкідливих відхилень, що дозволяє забезпечити цілісність і безпеку даних [6].

Традиційні підходи до верифікації [7] та валідації [8] зосереджуються на тестуванні та моніторингу систем за допомогою попередньо визначених методів виявлення аномалій на основі правил або статичного аналізу. Хоча ці підходи є фундаментальними, вони стають дедалі менш ефективними для управління великими обсягами динамічно змінюваних даних.

Підхід, що пропонується, ґрунтується на використані алгоритмів машинного навчання для подолання обмежень, властивих методам верифікації та валідації на основі правил і статичного аналізу. Завдяки використанню таких методів, системи можуть не лише виявляти аномалії, що відхиляються від відомих шаблонів, але й адаптуватися та навчатися на основі еволюції даних. Методи, засновані на машинному навчанні, забезпечують гнучкий, адаптивний та проактивний підхід до виявлення аномалій [9], підвищуючи ефективність процесів для складних інформаційних систем.

Information Systems and Technologies (IST-2024), November 26-28, 2024, Kharkiv, Ukraine
* Corresponding author.
oleksandr.nelipa1@nure.ua (O. Nelipa); nadiia.kalyta@nure.ua (N. Kalyta)
0000-0003-1387-1414 (O. Nelipa); 0000-0001-6181-732X (N. Kalyta)

2. Виявлення аномалій в процесах верифікації та валідації

У контексті процесів верифікації та валідації інформаційних систем «дефект» означає будь-яку непередбачену поведінку, функціональну помилку чи відхилення від очікуваних результатів у системі [10]. У ході таких процесів дефекти ідентифікуються та класифікуються для розуміння їхнього впливу, джерела походження та необхідних дій.

Аномалії – це несподівана поведінка або відхилення від типових патернів системи, що вказують на наявність дефектів [11]. Не всі аномалії є дефектами, але багато з них свідчать про приховані проблеми, які можуть вплинути на надійність, безпеку або продуктивність системи. Аномалії зазвичай поділяються на дві основні категорії:

Статистичні аномалії – це відхилення від усталених патернів чи порогових значень, наприклад, незвично високий час відгуку, пікове використання ресурсів або несподівані патерни мережевого трафіку [11]. Такі аномалії можуть свідчити про проблеми з продуктивністю, помилки в конфігурації або загрози безпеці.

Структурні аномалії – це відхилення від нормальних структур даних, наприклад відсутність полів у записі бази даних або невідповідності у схемах даних [11]. Структурні аномалії часто виявляють проблеми з цілісністю даних або функціональні помилки в системі.

Певні аномалії класифікуються як дефекти, якщо вони впливають на функціональність, продуктивність або надійність системи. Найбільш поширені типи аномалій, які можуть бути визнані дефектами, включають:

Сплески продуктивності – несподівані стрибки часу відгуку, використання пам’яті або навантаження на процесор, які можуть сигналізувати про вузькі місця продуктивності або витоки ресурсів.

Відхилення даних – значення, що виходять далеко за межі очікуваних діапазонів, наприклад, аномально високі суми транзакцій, незвичайні часи входу в систему або неочікувані зміни активності користувачів. Такі аномалії можуть свідчити про пошкодження даних або порушення безпеки.

Несанкціоновані патерни доступу – численні невдалі спроби входу з невідомих IP-адрес або доступ поза звичайними годинами. Такі аномалії часто вказують на дефекти безпеки або спроби несанкціонованого доступу.

Відхилення конфігурації – зміни в конфігураціях систем, такі як модифікації налаштувань безпеки, правил доступу до ресурсів або політик мережі. Конфігураційні аномалії можуть призвести до погіршення продуктивності чи вразливостей безпеки.

3. Обмеження існуючих підходів в процесах верифікації та валідації

Попри те, що традиційні методи верифікації та валідації є ключовими для забезпечення якості та безпеки інформаційних систем, із зростанням масштабу та складності систем вони стикаються зі значними обмеженнями.

Методи виявлення аномалій на основі правил покладаються на заздалегідь визначені умови та статичні порогові значення для ідентифікації аномалій у системі [12]. Хоча вони ефективні для виявлення добре відомих патернів аномальної поведінки, ці підходи мають низку недоліків, такі як негнучкість до змін, висока частка хибнопозитивних результатів, та проблеми з масштабованістю.

Інструменти статичного аналізу оцінюють код і структуру системи для виявлення потенційних вразливостей або відповідності встановленим стандартам безпеки та якості. Подібним чином, попередньо визначені підходи використовують тестові сценарії, розроблені для оцінки очікуваної поведінки системи [13, 14]. Однак ці методи також мають свої обмеження, такі як відсутність динамічного контексту, обмежений обсяг тестування та високі витрати.

Крім того, класичні методи часто залежать від людського досвіду для інтерпретації результатів і адаптації правил. Ця залежність створює вузькі місця у швидко змінюваних середовищах [15], де для обробки аномалій потрібне майже миттєве виявлення та реагування.

4. Виявлення аномалій на основі машинного навчання як удосконалене рішення

Машинне навчання відкриває революційні можливості для процесів верифікації та валідації, дозволяючи системам самостійно навчатися на історичних даних, розпізнавати складні патерни та адаптуватися до нових потенційних аномалій без потреби у заздалегідь визначених і прописаних правилах. Переваги такого підходу включають:

Адаптивність – алгоритми постійно навчаються на вхідних даних [9], що дозволяє їм виявляти як відомі, так і невідомі типи аномалій.

Масштабованість – алгоритми здатні обробляти великі набори даних і складні, багатовимірні структури даних, що робить їх більш придатними для сучасних, орієнтованих на дані систем.

Вибір конкретного алгоритму залежить від природи аномалій, типу вхідних даних та доступності маркованих даних [16, 17]. Найбільш підходящі типи алгоритмів для виявлення аномалій у процесах верифікації та валідації:

Навчання з учителем – доцільне, коли в історичних даних є марковані приклади нормальних і аномальних подій. Наприклад, у фінансових системах відомі патерни шахрайських транзакцій можуть бути використані як марковані дані для навчання моделей з учителем, що дозволить їм виявляти подібні патерни у майбутніх даних.

Навчання без учителя – ідеальне для виявлення нових аномалій без потреби у маркованих даних. Наприклад, у середовищах із динамічними патернами даних, таких як моніторинг мережевого трафіку, де моделі без учителя можуть виявляти раніше невідомі загрози.

Гібридні моделі – ці підходи використовують обмежений набір маркованих даних для покращення точності моделі у сценаріях, коли аномалії є рідкісними, а маркування – дорогим [17].

5. Запропонований підхід впровадження машинного навчання в процеси верифікації та валідації

Для інтеграції машинного навчання у процеси верифікації та валідації з метою виявлення аномалій пропонується структурована методологія, що складається з чотирьох основних етапів.

Збір та попередня обробка даних – ефективне виявлення аномалій залежить від наявності комплексних і якісних даних. Дані мають збиратися з різних джерел, включаючи системні журнали, активність користувачів, записи транзакцій і зовнішню інформацію про загрози. Попередня обробка є важливим кроком для видалення шуму, стандартизації ознак і нормалізації значень, що робить дані придатними для використання у моделях машинного навчання [16, 18].

Вибір відповідної моделі є критично важливим і залежить від доступності маркованих даних та типу аномалій, які потрібно виявити [9]. Для відомих типів аномалій використовуються моделі з учителем, які навчаються на маркованих даних, тоді як у середовищах, де аномалії виникають непередбачувано, застосовуються моделі без учителя.

Після навчання моделі розгортаються у реальних умовах, де вони постійно моніторять вхідні дані. Для підтримання точності використовується механізм зворотного зв’язку,

який дозволяє регулювати чутливість виявлення для мінімізації хибнопозитивних результатів [16], а також періодично перенавчати моделі на основі нещодавніх даних для врахування нових патернів або можливих змін у поведінці системи.

Оцінювання продуктивності є необхідним для забезпечення ефективності систем на основі машинного навчання у виявленні аномалій. Регулярне застосування метрик оцінки моделі, таких як точність, повнота та F1-score [19], допомагає оцінити здатність моделі ідентифікувати аномалії у системі. Для адаптивного навчання важливим є постійний моніторинг даних, оскільки зміни в основних патернах даних можуть вимагати перенавчання моделі для підтримання достатнього рівня точності [17].

Для верифікації та валідації на основі великих мовних моделей для мультимодального виявлення аномалій застосовуються BERT, GPT, LLaMA2, які виконувати широкий спектр завдань – від прогнозування до виявлення аномалій – із залученням різних джерел даних, ефективно інтегруючи мультимодальні потоки для підвищення точності виявлення аномалій.

6. Висновки

Процеси верифікації та валідації є ключовими для забезпечення надійності, точності та безпеки інформаційних систем. Традиційні методи, засновані на правилах, мають обмеження, такі як негнучкість, висока частка хибнопозитивних результатів та проблеми з масштабованістю. Інтеграція машинного навчання дозволяє подолати ці виклики завдяки автоматизованому виявленню аномалій. Моделі здатні навчатися на нових даних, адаптуватися до змін та виявляти складні, непередбачені аномалії.

Запропонований підхід передбачає класифікацію дефектів та обробку багатовимірних потоків даних за допомогою великих мовних моделей.

Таким чином, інтеграція машинного навчання в процеси верифікації та валідації усуває недоліки підходів на основі правил, забезпечуючи гнучкість, масштабованість і точність у виявленні дефектів, що є важливим кроком у розвитку сучасних інформаційних систем.

Література

[1] X. Jin, C. Katsis, F. Sang, J. Sun, A. Kundu, and R. Kompella, "Edge security: Challenges and issues," arXiv preprint, arXiv:2206.07164, 2022.

[2] M. Ahmed, A. N. Mahmood, and M. R. Islam, "A survey of anomaly detection techniques in financial domain," Future Generation Computer Systems, vol. 55, pp. 278–288, 2016.

[3] Z. Sun, C. E. Rubio-Medrano, Z. Zhao, T. Bao, A. Doupé, and G.-J. Ahn, "Understanding and predicting private interactions in underground forums," in Proceedings of the Ninth ACM Conference on Data and Application Security and Privacy, 2019, pp. 303–314.

[4] L. Zhang, D. Lai, and A. V. Miranskyy, "The impact of position errors on crowd simulation," Simulation Modelling Practice and Theory, vol. 90, pp. 45–63, 2019.

[5] W. Dang, B. Zhou, W. Zhang, and S. Hu, "Time Series Anomaly Detection Based on Language Model," in Proceedings of the Eleventh ACM International Conference on Future Energy Systems, E-Energy '20, Association for Computing Machinery, New York, NY, USA, June 2020, pp. 544–547.

[6] W. Dang, B. Zhou, L. Wei, W. Zhang, Z. Yang, and S. Hu, "TS-Bert: Time Series Anomaly Detection via Pre-training Model Bert," in M. Paszynski, D. Kranzlmüller, V. V. Krzhizhanovskaya, J. J. Dongarra, and P. M. A. Sloot (Eds.), Computational Science – ICCS 2021, Lecture Notes in Computer Science, Springer International Publishing, Cham, 2021, pp. 209–223.

[7] ISTQB Standard Glossary of Terms Used in Software Testing: Verification. Available: https://glossary.istqb.org/en_US/term/verification

[8] ISTQB Standard Glossary of Terms Used in Software Testing: Validation. Available: https://glossary.istqb.org/en_US/term/validation
[9] S. K. Devineni, S. Kathiriya, and A. Shende, "Machine learning-powered anomaly detection: Enhancing data security and integrity," Journal of Artificial Intelligence & Cloud Computing, vol. 184, no. 2, pp. 2–9, 2023. doi:10.47363/JAICC/2023
[10] S. Apel, A. von Rhein, P. Wendler, A. Größlinger, and D. Beyer, "Strategies for product-line verification: Case studies and experiments," in Proceedings of the 2013 35th International Conference on Software Engineering (ICSE), San Francisco, CA, USA, 2013, pp. 482–491. doi:10.1109/ICSE.2013.6606594.
[11] Z. Li, X.-Y. Jing, and X. Zhu, "Progress on approaches to software defect prediction," IET Software, vol. 12, 2018. doi:10.1049/iet-sen.2017.0148.
[12] G. Costa, A. Forestiero, and R. Ortale, "Rule-Based Detection of Anomalous Patterns in Device Behavior for Explainable IoT Security," IEEE Transactions on Services Computing, vol. PP, pp. 1-13, 2023. doi:10.1109/TSC.2023.3327822.
[13] J. Zheng, L. Williams, N. Nagappan, W. Snipes, J. P. Hudepohl, and M. A. Vouk, "On the value of static analysis for fault detection in software," IEEE Transactions on Software Engineering, vol. 32, no. 4, pp. 240-253, Apr. 2006. doi:10.1109/TSE.2006.38.
[14] N. Ayewah, W. Pugh, D. Hovemeyer, J. D. Morgenthaler, and J. Penix, "Using Static Analysis to Find Bugs," IEEE Software, vol. 25, no. 5, pp. 22-29, Sept.-Oct. 2008. doi:10.1109/MS.2008.130.
[15] W. L. Oberkampf and T. G. Trucano, "Verification and validation benchmarks," Nuclear Engineering and Design, vol. 238, no. 3, pp. 716-743, 2008. doi:10.1016/j.nucengdes.2007.02.032.
[16] R. Ahamed, A. Habeeb, F. Nasaruddin, A. Gani, I. A. Targio Hashem, E. Ahmed, and M. Imran, "Real-time big data processing for anomaly detection: A Survey," International Journal of Information Management, vol. 45, pp. 289-307, 2019. doi:10.1016/j.ijinfomgt.2018.08.006.
[17] E. Quatrini, F. Costantino, G. Di Gravio, and R. Patriarca, "Machine learning for anomaly detection and process phase classification to improve safety and maintenance activities," Journal of Manufacturing Systems, vol. 56, pp. 117-132, 2020. doi:10.1016/j.jmsy.2020.05.013.
[18] M. Hosseinzadeh, E. Azhir, O. Ahmed, M. Ghafour, S. Ahmed, A. Rahmani, and B. Vo, "Data cleansing mechanisms and approaches for big data analytics: a systematic study," Journal of Ambient Intelligence and Humanized Computing, vol. 14, pp. 1-13, 2021. doi:10.1007/s12652-021-03590-2.
[19] B. J. Erickson and F. Kitamura, "Magician's Corner: 9. Performance Metrics for Machine Learning Models," Radiology: Artificial Intelligence, vol. 3, no. 3, p. e200126, 2021. doi:10.1148/ryai.2021200126.
[20] J. Su, C. Jiang, X. Jin, Y. Qiao, T. Xiao, H. Ma, R. Wei, Z. Jing, J. Xu, and J. Lin, "Large Language Models for Forecasting and Anomaly Detection: A Systematic Literature Review," arXiv, abs/2402.10350, 2024.