УДК 004.056:004.85

А. В. Притула
Л. М. Куперштейн

БАЙЄСІВСЬКІ МЕТОДИ НАВЧАННЯ З ПІДКРІПЛЕННЯМ ДЛЯ ОЦІНЮВАННЯ НЕВИЗНАЧЕНОСТІ В ЗАДАЧАХ ТЕСТУВАННЯ БЕЗПЕКИ ВЕБЗАСТОСУНКІВ

Вінницький національний технічний університет.

Анотація

Розглянуто застосування байєсівських методів навчання з підкріпленням для оцінювання невизначеності у задачах тестування безпеки вебзастосунків. Показано, що інтеграція гаусових процесів з алгоритмами Temporal Difference та Actor–Critic дозволяє формувати не лише точкові оцінки функцій цінності станів MAL-графа атаки, а й дисперсійні характеристики, що виступають кількісними індикаторами ризику компрометації. Запропоновано інтерпретацію апостеріорної невизначеності як критерію пріоритетизації вузлів вебзастосунку для поглибленого аналізу безпеки.

Ключові слова: байєсівське навчання з підкріпленням, гаусівський процес, тестування безпеки, вебзастосунок, MAL-граф, невизначеність, час до компрометації.

Abstract

The application of Bayesian reinforcement learning methods for uncertainty estimation in web application security testing tasks is considered. It is shown that the integration of Gaussian processes with Temporal Difference and Actor–Critic algorithms enables the formation of not only point estimates of value functions for MAL attack graph states, but also variance characteristics that serve as quantitative indicators of compromise risk. An interpretation of posterior uncertainty as a criterion for prioritizing web application nodes for in-depth security analysis is proposed.

Keywords: Bayesian reinforcement learning, Gaussian process, security testing, web application, MAL graph, uncertainty, time to compromise.

Вступ

Сучасні підходи до тестування безпеки вебзастосунків дедалі частіше залучають методи машинного навчання та навчання з підкріпленням (Reinforcement Learning, RL) для моделювання складних сценаріїв кібератак [1-3]. Класичні механізми захисту, зокрема системи виявлення вторгнень та системи запобігання вторгненням, базуються на жорстко визначених правилах або сигнатурах, що робить їх малоефективними у випадку появи нових чи модифікованих атак [4, 5]. Водночас методи RL дозволяють формалізувати протиборство атакувальника й захисника у термінах марковських процесів прийняття рішень (Markov Decision Process, MDP), де простір станів визначається конфігурацією компрометації системи, а дії відповідають допустимим переходам у графі атак [6].

Проте класичні алгоритми RL оперують точковими оцінками функцій цінності та не враховують невизначеність, притаманну динамічним середовищам кіберзахисту. У реальних сценаріях розподіл часових характеристик компрометації (Time to Compromise, TTC) може залишатися невідомим або змінюватися внаслідок застосування контрзаходів [7]. Це робить актуальним завдання залучення байєсівських методів, здатних «вбудувати» невизначеність безпосередньо у процес прийняття рішень агентом.

Метою роботи є аналіз та адаптація байєсівських методів навчання з підкріпленням для кількісного оцінювання невизначеності в задачах тестування безпеки вебзастосунків на основі MAL-графів атак.

Результати дослідження

У контексті тестування безпеки вебзастосунків стани MDP інтерпретуються як поточні конфігурації компрометації системи, визначені виконаними та доступними кроками атаки в MAL-графі, тоді як дії відповідають допустимим переходам між ними. Часові параметри компрометації TTC інтегруються у функцію переходів, що формалізує ймовірнісно-часовий характер розвитку атаки [8].

MAL (Meta Attack Language) - це предметно-орієнтована мова моделювання загроз, що дозволяє формалізувати сценарії кібератак у вигляді спрямованих графів [7]. Вузли MAL-графа відповідають окремим крокам атаки або захисним заходам, а ребра - причинно-наслідковим або часовим

залежностям між ними. Кожному ребру може бути поставлено у відповідність часовий параметр TTC, що характеризує очікуваний час виконання відповідного кроку компрометації. Така структура дозволяє подати процес атаки як послідовність переходів у марковському процесі, де кожен стан описує поточну конфігурацію компрометації вебзастосунку, а простір дій визначається доступними переходами в графі.

Байєсівське навчання з підкріпленням виходить із припущення, що агент не може мати повної інформації про функції переходів або винагород. На відміну від класичних методів RL, байєсівський підхід формує апостеріорний розподіл параметрів політики, що дає змогу враховувати всю множину можливих моделей середовища. Для задач тестування безпеки вебзастосунків це є принципово важливим, адже атакувальник і захисник взаємодіють у просторі станів MAL-графа, в якому розподіл TTC може залишатися невідомим [9].

Ключовим інструментом запропонованого підходу є метод Gaussian Process Temporal Difference (GP-TD), що поєднує рівняння Беллмана з гаусівськими процесами. Функція цінності апроксимується як гаусівський процес, що дозволяє отримати як оцінку математичного сподівання, так і дисперсію прогнозу. На відміну від класичного Actor–Critic, сигнал TD-помилки у GP-TD має імовірнісний характер і описується не лише математичним сподіванням, а й дисперсією, що забезпечує більш обережне оновлення політики в ділянках MAL-графа з обмеженою або суперечливою інформацією [10].

Апостеріорні оцінки математичного сподівання та дисперсії функції цінності для стану s формуються відповідно до виразів:

$$\mu_*(s) = k_*^T (K + \sigma^2 I)^{-1} r\,, \qquad (1)$$

$$\sigma_*^2(s) = k(s,s) - k_*^T (K + \sigma^2 I)^{-1} k_*\,, \quad (2)$$

де $\mu_*(s)$ - апостеріорна оцінка математичного сподівання функції цінності стану s, K - ядерна коваріаційна матриця над відвіданими станами, k_* - вектор коваріацій між новим станом і навчальною траєкторією, r - вектор спостережуваних винагород, s - стан MAL-графа, σ^2 - дисперсія шуму спостережень, I - одинична матриця, а верхній індекс T позначає операцію транспонування. Вираз (2) виступає кількісним індикатором ризику: висока дисперсія сигналізує про ділянки MAL-графа, що потребують поглибленого аналізу.

Дисперсія апостеріорного прогнозу набуває значення кількісної характеристики ризику: підвищена варіативність оцінок TTC для окремих траєкторій у MAL-графі вказує на недостатню інформативність даних або на потенційно критичні сценарії компрометації. Водночас ядрова структура гаусівського процесу забезпечує перенесення знань між структурно подібними конфігураціями атак, що дозволяє отримувати обґрунтовані оцінки навіть за відсутності повної статистики для рідкісних векторів впливу.

Розвитком GP-TD є байєсівський підхід Actor–Critic (BAC), у якому критик реалізується у вигляді гаусівського процесу над Q-функцією, а оновлення політики ґрунтується на апостеріорній оцінці градієнта очікуваної винагороди з урахуванням невизначеності. Апріорна модель Q-функції задається як гаусівський процес з ядром, що може бути визначене, наприклад, як радіально-базисне ядро, яке безпосередньо відображає подібність за часовими характеристиками TTC. Невизначеність явно виражається через дисперсію апостеріорної оцінки. Якщо ця дисперсія висока, система може відкласти остаточне рішення або активувати лише часткові контрзаходи [11].

Таблиця 1 - Порівняльна характеристика байєсівських методів RL

Метод	**Простір дій**	**Оцінка невизначеності**	**Масштабованість**
Bayesian Q-learning	Дискретний	Дисперсія $Q(s,a)$	Низька
Bayesian Policy Gradient	Безперервний	Апостеріор градієнта	Середня
GP-TD	Будь-який	$\sigma^2(s)$	Середня
Bayesian Actor–Critic	Будь-який	$\sigma^2(s,a)$	Висока

Порівняльний аналіз байєсівських методів RL показує, що кожен із них має власну нішу застосування. Bayesian Q-learning є оптимальним для дискретних і структурованих політик, але погано масштабується до складних топологій. Bayesian Policy Gradient дозволяє працювати з багатовимірними безперервними діями. GP-TD забезпечує точне оцінювання ризиків завдяки дисперсії прогнозів. BAC поєднує адаптивність із надійністю, інтегруючи невизначеність безпосередньо у градієнтні оновлення

політики, що робить його найбільш придатним для практичного керування діями захисту вебзастосунків [12].

Висновки

Показано, що байєсівські методи навчання з підкріпленням, зокрема GP-TD та Bayesian Actor–Critic, забезпечують формальний інструментарій для роботи з невизначеністю в задачах моделювання протиборства «атака–захист» на MAL-графах вебзастосунків. Інтеграція гаусівських процесів із TD-методами дозволяє оцінювати не лише очікувану ефективність атакувальних або захисних дій, а й ризики, пов'язані з малоймовірними, проте критичними сценаріями компрометації. Запропонований підхід створює методологічні передумови для побудови адаптивних стратегій тестування безпеки, чутливих до варіативності динаміки атаки та обмежень середовища.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Толкачова А., Піскозуб А. Методи для тестування безпеки веб-застосунків // Кібербезпека: освіта, наука, техніка. - 2024. - №2 (26). 115–122. DOI: https://doi.org/10.28925/2663-4023.2024.26.668 (дата звернення: 14.03.2026).
2. Притула А. В., Куперштейн Л. М. Застосування методів навчання з підкріпленням у тестуванні на проникнення: ефективність, виклики та перспективи // Студентська конференція з Інформаційної, Функційної і Кібербезпеки (СКІФіК). - Харків, 2024. - С. 78–80. (дата звернення: 14.03.2026).
3. Aydos M., Aldan Ç., Coşkun E., Soydan A. Security testing of web applications: A systematic mapping of the literature // Journal of King Saud University - Computer and Information Sciences. - 2022. - Vol. 34, No. 9. - P. 6775–6792. DOI: https://doi.org/10.1016/j.jksuci.2021.09.018 (дата звернення: 14.03.2026).
4. Díaz-Verdejo J., Muñoz-Calle J., Estepa Alonso A. et al. On the Detection Capabilities of Signature-Based Intrusion Detection Systems in the Context of Web Attacks // Applied Sciences. - 2022. - Vol. 12, No. 2. - P. 852. DOI: https://doi.org/10.3390/app12020852 (дата звернення: 19.03.2026).
5. Притула А. В., Куперштейн Л. М., Маліновський В. І. Аналіз технологій тестування на проникнення web-додатків // Наукові праці Вінницького національного технічного університету. - 2024. - №2. DOI: https://doi.org/10.31649/2307-5376-2024-2-45-53. (дата звернення: 16.03.2026).
6. Song B., Sun L., Qin Z. Design of Web Security Penetration Test System Based on Attack and Defense Game // Scientific Programming. - 2022. - Vol. 2022. - P. 1–11. DOI: https://doi.org/10.1155/2022/8645969 (дата звернення: 15.03.2026).
7. Katsikeas S., Rencelj Ling E., Johnsson P., Ekstedt M. Empirical evaluation of a threat modeling language as a cybersecurity assessment tool // Computers & Security. - 2024. - Vol. 140. DOI: https://doi.org/10.1016/j.cose.2024.103743 (дата звернення: 15.03.2026).
8. Do C. T. et al. Game Theory for Cyber Security and Privacy // ACM Computing Surveys. - 2017. - Vol. 50, No. 2. - P. 1–37. DOI: https://doi.org/10.1145/3057268 (дата звернення: 15.03.2026).
9. Huang L., Zhu Q. A dynamic games approach to proactive defense strategies against Advanced Persistent Threats in cyber-physical systems // Computers & Security. - 2020. - Vol. 89. - P. 1-24. DOI: https://doi.org/10.1016/j.cose.2019.101660 (дата звернення: 19.03.2026).
10. Liu S., Cao J., Wang Y., Chen W., Liu Y. Self-play reinforcement learning with comprehensive critic in computer games // Neurocomputing. - 2021. - Vol. 449. - P. 207–213. DOI: https://doi.org/10.1016/j.neucom.2021.04.006 (дата звернення: 14.03.2026).
11. Mariselvam J., Alotaibi Y., Rajendran S. Proximal policy optimization (PPO) algorithm's role in autonomous decision-making for autism // Results in Engineering. - 2025. - Vol. 28. DOI: https://doi.org/10.1016/j.rineng.2025.107683 (дата звернення: 19.03.2026).
12. Li N., Zhang M., Li J. et al. A Game-Theoretical Self-Adaptation Framework for Securing Software-Intensive Systems // ACM Trans. Auton. Adapt. Syst. - 2024. - Vol. 19, No. 2. - P. 1–49. DOI: https://doi.org/10.48550/arXiv.2112.07588 (дата звернення: 19.03.2026).

Притула Андрій Вікторович – студент групи 125-23а, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: andrik.pritula@gmail.com.

Куперштейн Леонід Михайлович – к.т.н., доцент кафедри захисту інформації, Вінницький національний технічний університет, м. Вінниця, email: kupershtein.lm@gmail.com.

Prytula Andrii V. – Student of Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, e-mail: andrik.pritula@gmail.com.

Kupershtein Leonid M. – PhD, Associated Professor of Information Protection Chair, Vinnytsia National Technical University, Vinnytsia, email: kupershtein.lm@gmail.com.