МЕТОД РАННЬОГО ВИЯВЛЕННЯ МІЖЛАНЦЮГОВИХ SANDWICH-АТАК У МОСТОВИХ ПРОТОКОЛАХ ДЕЦЕНТРАЛІЗОВАНИХ СИСТЕМ

Антіпін В.С., Олійников Р.В.
Харківський національний університет радіоелектроніки, Харків, Україна

Розвиток міжланцюгових протоколів у децентралізованих системах ускладнив структуру MEV-загроз. Якщо в межах одного блокчейна перевага нападника зазвичай формується через спостереження за публічним мемпулом і зміну порядку включення транзакцій, то в міжланцюговому середовищі вона може виникати ще на етапі появи події у вихідному ланцюгу, коли операція в цільовій мережі ще не стала локально видимою [2, 3, 5]. Саме тому міжланцюговий MEV сьогодні розглядається як окремий напрям, для якого визначальними є затримки між доменами, різна фіналізація та прозорість міжланцюгових повідомлень.

Одним із найнебезпечніших проявів цього механізму є міжланцюгові sandwich-атаки в мостових протоколах на основі пулів ліквідності [5]. Їхня особливість полягає в тому, що подія у вихідному ланцюгу може розкривати параметри майбутнього обміну раніше, ніж відповідна операція стане видимою в мережі призначення. Унаслідок цього нападник отримує структурну інформаційну перевагу й може підготувати front-run-транзакцію ще до появи транзакції жертви в локальному середовищі. Дослідження на даних протоколу Symbiosis за період з 10 серпня до 10 жовтня 2025 року показало, що сукупний прибуток таких атак перевищив 5,27 млн дол. США, або 1,28 % загального обсягу міжланцюгових переказів у вибірці [5].

Звичайні засоби виявлення MEV, орієнтовані на локальний мемпул, перестановку транзакцій у блоці або аномальну поведінку окремого пулу ліквідності, у такому випадку виявляються недостатніми. Ключова ознака атаки виникає не в одному блокчейні, а в причинному зв'язку між подією у вихідному ланцюгу та виконанням у цільовому. Це підтверджують і дослідження міжланцюгового арбітражу, які показують, що вирішальну роль у таких сценаріях відіграють саме часові затримки та інфраструктурні переваги [1].

Запропонований підхід спирається не на аналіз окремої транзакції, а на виявлення причинно пов'язаного міжланцюгового ланцюжка подій. На відміну від підходів, що фіксують лише локальні ознаки front-running у межах одного блокчейна, така схема враховує зв'язок між подією ініціації, мостовим повідомленням і виконанням у цільовому ланцюгу. Основою виявлення виступають не ізольовані транзакції, а цілісний шаблон поведінки, для якого обчислюється показник ранньої поінформованості Δ_{info} та інтегральна оцінка підозрілості. Саме ця зв'язка – міжланцюговий причинно-часовий ланцюжок, показник ранньої поінформованості та комбінований вектор ознак – і становить головну відмінність запропонованого методу від локальних детекторів, що бачать тільки цільовий ланцюг.

Міжланцюгову sandwich-атаку зручно подавати у вигляді ланцюжка

$$\mathrm{p} = \langle e_s, \mathrm{m_r}, \mathrm{t_f}, \mathrm{t_v}, \mathrm{t_b} \rangle,$$

де e_s – подія ініціації міжланцюгового обміну у вихідному ланцюгу;

m_r – повідомлення мосту або релеєра;

t_f – front-run-транзакція нападника в цільовому ланцюгу;

t_v – транзакція користувача;

t_b – завершальна транзакція нападника (*back-run*), що фіксує результат після зміни ціни.

Атака має місце тоді, коли послідовність $t_f \to t_v \to t_b$ у цільовому ланцюгу з'являється не лише як реакція на локальний мемпул, а як наслідок події e_s у вихідному ланцюгу.

Інакше кажучи, нападник використовує міждоменний інформаційний витік для отримання переваги ще до того, як жертва стає помітною у звичайному локальному середовищі. Така постановка відповідає механіці, описаній у свіжому дослідженні саме цього класу атак.

Економічний зміст атаки можна подати через функцію

$$\Pi(p) = \Delta V - C_{gas} - C_{bridge} - C_{time} - R_{fail},$$

де ΔV – вартість, отримана внаслідок зміни ціни в пулі ліквідності;

C_{gas} – витрати на транзакції нападника;

C_{bridge} – супутні витрати міжланцюгової інфраструктури;

C_{time} – втрати, пов'язані із затримками між доменами;

R_{fail} – ризик неповного або повного зриву сценарію.

Для одно-ланцюгової sandwich-атаки головним є місце у блоці та ціна газу, тоді як у міжланцюговому сценарії критичною стає також затримка між появою події у вихідному ланцюгу й фактичним виконанням у цільовому. Саме часовий розрив між цими фазами створює простір для випереджального впливу на ціну в пулі ліквідності.

Для виявлення таких шаблонів використовується вектор ознак

$$x(p) = \langle \Delta_{sm}, \Delta_{mv}, \Delta_{fv}, \Delta_{vb}, \delta_{price}, \sigma_{slip}, \rho_{pool}, \kappa_{addr}, \eta_{route} \rangle,$$

де Δ_{sm} – проміжок між подією ініціації та появою мостового повідомлення;

Δ_{mv} – інтервал між мостовим повідомленням і входом операції користувача до вікна виконання в цільовому ланцюгу;

Δ_{fv} – випередження front-run-транзакції нападника відносно транзакції користувача;

Δ_{vb} – інтервал між операцією користувача та завершальною транзакцією нападника;

δ_{price} – відхилення ціни після front-run;

σ_{slip} – фактичне прослизання користувача;

ρ_{pool} – частка ліквідності пулу, на яку вплинув шаблон атаки;

κ_{addr} – повторюваність тих самих адрес або кластерів адрес у суміжних випадках;

η_{route} – стабільність одного й того самого міжланцюгового маршруту в серії підозрілих епізодів.

Такий набір поєднує часові, економічні, маршрутні й поведінкові сигнали, що особливо важливо для міжланцюгового середовища, де жодна окрема ознака сама по собі не є достатньою.

Метод виявлення будується як трирівневий позаланцюговий конвеєр і може застосовуватися в системах спостереження за мостовими маршрутами в режимі, наближеному до реального часу.

На першому етапі формується журнал подій, у який потрапляють події ініціації міжланцюгового обміну, повідомлення мостів і релеєрів, а також події обміну в пулах ліквідності цільового ланцюга.

На другому етапі у ковзному часовому вікні для заздалегідь визначених маршрутів і пар активів будуються кандидатні причинно пов'язані ланцюжки між двома мережами. Це зменшує обчислювальну складність, оскільки аналіз виконується не по всіх можливих подіях глобально, а в межах конкретних маршрутів, активів і часових інтервалів.

На третьому етапі для кожного ланцюжка обчислюється інтегральна оцінка підозрілості

$$S(p) = \sum w_i z_i(x_i),$$

де z_i – нормалізовані значення ознак;

w_i – ваги цих ознак.

Якщо $S(p) \geq \theta$ – ланцюжок позначається як потенційна міжланцюгова sandwich-атака.

Ваги w_i пропонується задавати комбіновано: експертно – на основі механіки атаки та емпірично – за історичними інцидентами і фоновою легітимною активністю. Це зменшує залежність від ручного налаштування та не потребує великої розміченої вибірки.

Показник ранньої поінформованості пропонується подавати у вигляді

$$\Delta_{info} = t(t_f) - t(e_s),$$

де $t(t_f)$ – час появи front-run-транзакції нападника;

$t(e_s)$ – час події ініціації у вихідному ланцюгу.

Якщо це значення систематично менше за типовий час реакції локального ринку, можна припустити, що нападник використовує не лише локальний мемпул, а й інформацію з іншого домену. Це дозволяє відокремити міжланцюгову sandwich-атаку від звичайної конкуренції торгових ботів у межах одного блокчейна. Метод є найбільш придатним для маршрутів, у яких мостові події або повідомлення дозволяють відновити параметри майбутнього обміну до етапу виконання.

Порогове значення для інтегральної оцінки підозрілості не слід вважати фіксованим. Його варто калібрувати на валідаційній вибірці з урахуванням хибнопозитивних і хибнонегативних спрацювань. Тому поріг має залежати від профілю ризику конкретного маршруту, активу або пулу ліквідності.

З метою оцінювання ефективності, для порівняння потрібен базовий детектор, який аналізує лише локальні події цільового ланцюга без урахування подій ініціації у вихідному. Такий підхід фіксує локальні шаблони front-running і цінового впливу, але не виявляє міждоменний інформаційний витік. Порівняння з ним дозволяє перевірити, чи справді врахування причинно-часового зв'язку між двома ланцюгами зменшує кількість пропущених атак. Ефективність методу можна оцінювати за метриками precision, recall, F1-score, а також за середнім часом випереджального виявлення, часткою випадків із підтвердженим прибутком нападника та повторюваністю маршрутів і адресних шаблонів.

Регулярна поява ланцюжків $e_s \to t_f \to t_v \to t_b$, мала варіативність інтервалу Δ_{fv}, повторюваність адресних кластерів і суттєве прослизання в користувачів вказують на те, що міст або маршрут розкриває надмірний обсяг інформації до етапу виконання. Це дозволяє використовувати метод не лише для виявлення атак, а і для оцінювання інформаційної вразливості конкретних мостових протоколів.

Обмеженням підходу є залежність від повноти журналу подій, точності часової синхронізації між мережами та доступності даних про маршрути виконання. У сценаріях із непрозорими каналами передавання або багатокроковими маршрутами побудова причинно пов'язаного ланцюжка ускладнюється, що підвищує вимоги до калібрування порогового значення та збільшує ризик хибнопозитивних спрацювань.

Протидія таким атакам має бути спрямована насамперед на зменшення інформативності подій у вихідному ланцюгу. Для цього можуть застосовуватися відкладене розкриття параметрів обміну, схеми commit–reveal, приватні канали передавання повідомлень або зміна структури мостових подій так, щоб із них неможливо було завчасно відновити намір користувача. Отже, міжланцюгові sandwich-атаки доцільно виявляти не за ізольованими локальними подіями, а за причинно пов'язаними ланцюжками між двома мережами. Такий підхід дає змогу фіксувати сценарії, які залишаються непомітними для локальних детекторів, і краще відповідає природі міжланцюгового MEV.

Список літератури

1. Cross-Chain arbitrage: the next frontier of MEV in decentralized finance / B. Öz et al. URL: https://arxiv.org/abs/2501.17335.
2. Flash boys 2.0: frontrunning, transaction reordering, and consensus instability in decentralized exchanges / P. Daian et al. URL: https://arxiv.org/abs/1904.05234.
3. Mancino D., Sevim H. O. SoK: the evolution of maximal extractable value, from miners to cross-chain. URL: https://arxiv.org/abs/2603.07716.
4. Qin K., Zhou L., Gervais A. Quantifying blockchain extractable value: how dark is the forest?. URL: https://arxiv.org/abs/2101.05511.
5. The walls have ears: unveiling cross-chain sandwich attacks in defi / C. Li et al. URL: https://arxiv.org/abs/2511.15245.