Міжнародний конкурс студентських робіт зі штучного інтелекту 2026

II тур
Перший етап

Р Е Ц Е Н З І Я

на наукову роботу

назва Адаптивне виявлення мережевих атак у зашифрованому трафіку на основі каскадної моделі з накопиченням відхилень за джерелом

шифр ADAPTIVE SENTINEL

представлену на Міжнародний конкурс студентських робіт зі штучного інтелекту

№ з/п	Характеристики та критерії оцінки рукопису наукової роботи	Рейтингова оцінка. Максимальна кількість балів (за 120-бальною шкалою)	Бали
1	Актуальність проблеми	10	9,2
2	Новизна та оригінальність ідей	15	11,8
3	Використані методи дослідження	15	11,2
4	Теоретичні наукові результати	10	7,0
5	Практична направленість результатів	20	15,6
6	Рівень використання наукової літератури та інших джерел інформації	5	4,4
7	Ступінь самостійності роботи	10	6,4
8	Якість оформлення	5	4,8
9	Наукові публікації	10	3,8
10	Недоліки роботи (пояснення зниження максимальних балів у пунктах 1-9): Не вистачає практичних прикладів, які демонструють перевагу запропонованого підходу. Є елементи використанні ШІ (GPT/Claude). Основні зауваження: нерівні умови порівняння; неповні абляції; недостатні теоретичні гарантії; відсутність оцінок невизначеності та швидкодії.		
10.1	Наявність розробленого програмного забезпечення, у т.ч. веб-сервісу, мобільного застосунку	10	6,8
10.2	Рівень отриманих практичних результатів наукової роботи	10	7,4
Сума балів			88,4

Загальний висновок рекомендується для захисту на ПНПК
(рекомендується, не рекомендується для захисту на науково-практичній конференції)

Рецензент *члени Конкурсної комісії напряму «Обчислювальний інтелект»*

«06» листопада 2026 року

Виявлення атак без розшифрування корисного навантаження та накопичення слабких відхилень за джерелом мають чітку прикладну мотивацію. Постановка враховує потокову обробку, дрейф і забруднення навчальних даних.

Сильна сторона – поєднання двобічного накопичення за актором, каскадної маршрутизації, рекурентного варіаційного ансамблю та інцидентної агрегації. Водночас складові спираються на відомі CUSUM, HST, VAE та SPOT. Немає окремих абляцій нижньої статистики, поправки актора і групування за джерелом; не доведено, яка саме модифікація забезпечує приріст.

Позитивно оцінено часову постановку, відкладене оцінювання, п’ять базових методів, чотири набори та три початкові значення генератора. Проте каскад адаптується й отримує похідні ознаки, а базові методи заморожені та їх не отримують. Контроль одночасно вимикає донавчання і зміну порога, тому не відокремлює їх внески. Не наведено розкиду трьох запусків, довірчих інтервалів, повних матриць помилок і повного переліку налаштувань каскаду.

Модель докладно формалізовано формулами (2.1)–(2.24). Однак твердження про зменшення дисперсії ковзного середнього рівно у G разів потребує некорельованості оцінок; для часової залежності слід ураховувати коваріації. Абсолютна неможливість виявлення будь-яким порогом сформульована надто сильно. Мінімум реконструкційної похибки на моді не є загальною гарантією для довільного автоенкодера. Квантиль також задає ймовірність перевищення за відповідного розподілу; перевагу EVT слід пояснювати оцінюванням хвоста, а не принциповою відсутністю ймовірнісного змісту квантиля. Для залежних даних, забруднення та обмеження порога (2.22) не доведено збереження цільового FPR.

Описано роботу на метаданих, два режими виконання та об’єднання тривог у інциденти. Разом із тим немає перевірки на безпосередньо захопленому TLS/QUIC-трафіку з підтвердженням складу протоколів; використання лише метаданих саме по собі цього не доводить. Не наведено пропускної здатності, затримки виявлення, пам’яті залежно від кількості акторів і вимірювань під NAT. Ретроспективне підвищення категорії тривог слід відокремити від своєчасного онлайн-виявлення.

Використано джерела з потокового аналізу, кібербезпеки, послідовних критеріїв і варіаційного навчання. Водночас для експериментальних наборів бракує окремих джерел із точними версіями та адресами отримання; сучасні споріднені методи переважно обговорено без прямого експериментального зіставлення.

Рукопис демонструє зв’язний виклад і деталізовані інженерні рішення, що підтримує позитивну оцінку. Проте відсутні опис особистого внеску, історія розробки.

За табл. 3.3 каскад має найвище MCC на трьох наборах: 0,762; 0,901; 0,466. На четвертому поступається щільному автоенкодеру: 0,209 проти 0,300. Позитивно, що показано і невдалі класи: TPR сканування 0,0014, проникнення 0,0149. FPR 0,071 на NF-ToN-IoT-v3 означає 7,1 % нормальних потоків і потребує операційної оцінки навантаження. Без інтервалів невизначеності, окремих інцидентних метрик та польової перевірки висновок про загальну перевагу передчасний.

Питання та рекомендації для захисту

Яким є внесок двобічного CUSUM окремо від похідних ознак, EVT і донавчання? Потрібні контрольні порівняння за однакових ознак і режимів адаптації.

Як забезпечено хронологічний порядок для наборів без часових позначок? Чому відсутність часу унеможливлює CUSUM і вікно з W потоків, хоча ці формули можна визначити за порядком надходження за наявності ключа актора?

Як побудовано вікно рівня 2 та інциденти для 02-15-2018 без ідентифікатора джерела? Опис на с. 24 потребує явного правила заміни ключа.

Чи може перенесення стійкого відхилення в поправку актора за (2.14) нормалізувати тривалу атаку? Як це узгоджується із захистом від отруєння?

Якими є онлайн-затримка до першої тривоги, частота хибних інцидентів і метрики до ретроспективної перекласифікації?

Робота є змістовним студентським дослідженням із перспективним поєднанням методів і позитивними експериментальними результатами. Рекомендується для захисту на підсумковій науково-практичній конференції.