УДК 004.056:004.77:519.17 **DOI:** 10.31673/2412-9070.2026.024909

**І. О. ШАХМАТОВ,** викладач,
ORCID: 0009-0004-9628-0365
Державний університет інформаційно-комунікаційних технологій, Київ

# ІНТЕГРОВАНИЙ КОНТУР ДОВІРИ У ВЕБЗАСТОСУНКАХ НА ОСНОВІ ГРАФОВОГО ОЦІНЮВАННЯ РИЗИКУ ТА НЕЗМІННОГО ЖУРНАЛЮВАННЯ РІШЕНЬ

***Порушення довіри та цілісності у вебзастосунках часто виникають на межі від вводу даних до виконання функціонального завдання, коли зловмисний або аномальний ввод (вебформа/API) переходить у дію в бізнес-логіці та транзакціях. У цій роботі розглядаються такі епізоди як єдиний потік подій і запропоновано інтегрований контур довіри, а саме контекстне оцінювання ризику на графі взаємодій та верифіковану фіксацію результатів у незмінному журналі. Ризик оцінюється графовою моделлю, яка враховує зв'язки подій у ковзному часовому вікні між актором (користувачем або сервісом), сесією, пристроєм і агрегованими мережевими сигналами. Для узгодження різнорідних джерел ми використовуємо уніфіковане подання подій і спільний підхід до ознак, для вебформ це змістові, структурні та часові характеристики, для транзакцій це пояснюваний інтегральний показник ризику, що агрегує вплив суми, часу, геопросторових відхилень і типу платіжного інструмента. Точність прийняття рішень забезпечується нормалізацією подій, хешуванням і цифровим підписом сервісу-реєстратора перед записом у незмінний журнал мінімально достатніх атрибутів, а саме хешів подій, ідентифікаторів, контрольних сум версій моделі, параметрів політики реагування та підсумкових оцінок. Запропонований підхід дозволяє виявляти підміни історії та приймати обґрунтовані рішення в межах зафіксованих даних, що важливо для аудиту й розслідування інцидентів у вебсередовищі.***

**Ключові слова:** веббезпека, графові нейронні мережі, виявлення спаму, незмінний журнал, аудит, версіонування моделей, програмне забезпечення, архітектура.

## *Вступ*

Вебзастосунки залишаються вразливими до бот-активності, спаму, шахрайських транзакцій і підміни даних. Особливо небезпечними є сценарії, у яких підозрілі дії у вебформах пов'язані з подальшими транзакціями, тому ізольований контроль окремих подій часто є недостатнім [10–12]. Сучасні підходи до побудови контуру довіри у вебзастосунках поєднують графове подання подій, інтелектуальне оцінювання ризику та незмінне журналювання рішень. Графи походження дають змогу відстежувати зв'язки між подіями, процесами й артефактами, а графові та темпоральні моделі, зокрема GNN, дозволяють враховувати не лише ознаки окремої події, а й структуру та динаміку взаємодій [3–7]. Важливим доповненням є незмінне журналювання, у якому фіксуються події, рішення, часові мітки, версії моделей і контрольні відбитки, з використанням permissioned-blockchain, append-only журналів або хеш-ланцюжків [1], [12]–[14], [17], [18], [20]. Література підтверджує доцільність такого поєднання, але вказує на потребу балансу між доказовістю, продуктивністю, приватністю та повнотою фіксації подій [21]. Проблематика полягає в недостатній ефективності ізольованого виявлення загроз у вебформах і транзакціях та відсутності надійної перевірки прийнятих рішень. Метою дослідження є підвищення довіри й цілісності у вебзастосунках шляхом інтегрованого оцінювання ризику подій вебформ і транзакцій та забезпечення перевірки прийнятих рішень. Постановка завдання полягає у побудові моделі, яка поєднує контекстний аналіз подій у спільному часовому та графовому середовищі з незмінним журналюванням результатів, що дає змогу не лише підвищити

якість виявлення загроз, а й забезпечити аудитну відтворюваність рішень. Практична цінність підходу полягає у зменшенні хибних спрацювань, підвищенні якості детекції та формуванні доказової історії рішень для аудиту й розслідування інцидентів.

### *Основна частина*

*Модель та архітектура системи довіри у вебзастосунках*

Модель виконує дві основні функції: контекстно оцінює ризик події та формує перевірюваний аудитний слід рішення. Розглядається потік подій вебзастосунку, де критичними є події надсилання форми та транзакції (рис.1).

Подія у момент часу $t$ задається як (1):

$$e_t = \langle id, type, ts, actor, session, resource, ctx, payload \rangle, \quad (1)$$

де $type \in$ SUBMIT,TX. Для забезпечення перевірюваності використовується канонізація події та її криптографічна фіксація (2):

$$h_e = \text{SHA256}\big(canon(e_t)\big), \; sig_e = \text{Sign}(sk_{svc}, h_e). \quad (2)$$

Після цього для кожної події формується вектор ознак $x_t$. Для транзакцій додатково використовується інтегральний показник ризику (3):

$$R_{TX} = R_P(P) + R_C(C) + R_T(T) + R_K(K), \quad (3)$$

де окремі компоненти визначаються, зокрема, за сумою операції та часовим фактором (4, 5):

$$R_P(P) = a \log(P + 1), \quad (4)$$

$$R_T(T) = b \exp\left(-\frac{(T-\mu)^2}{2\sigma^2}\right). \quad (5)$$

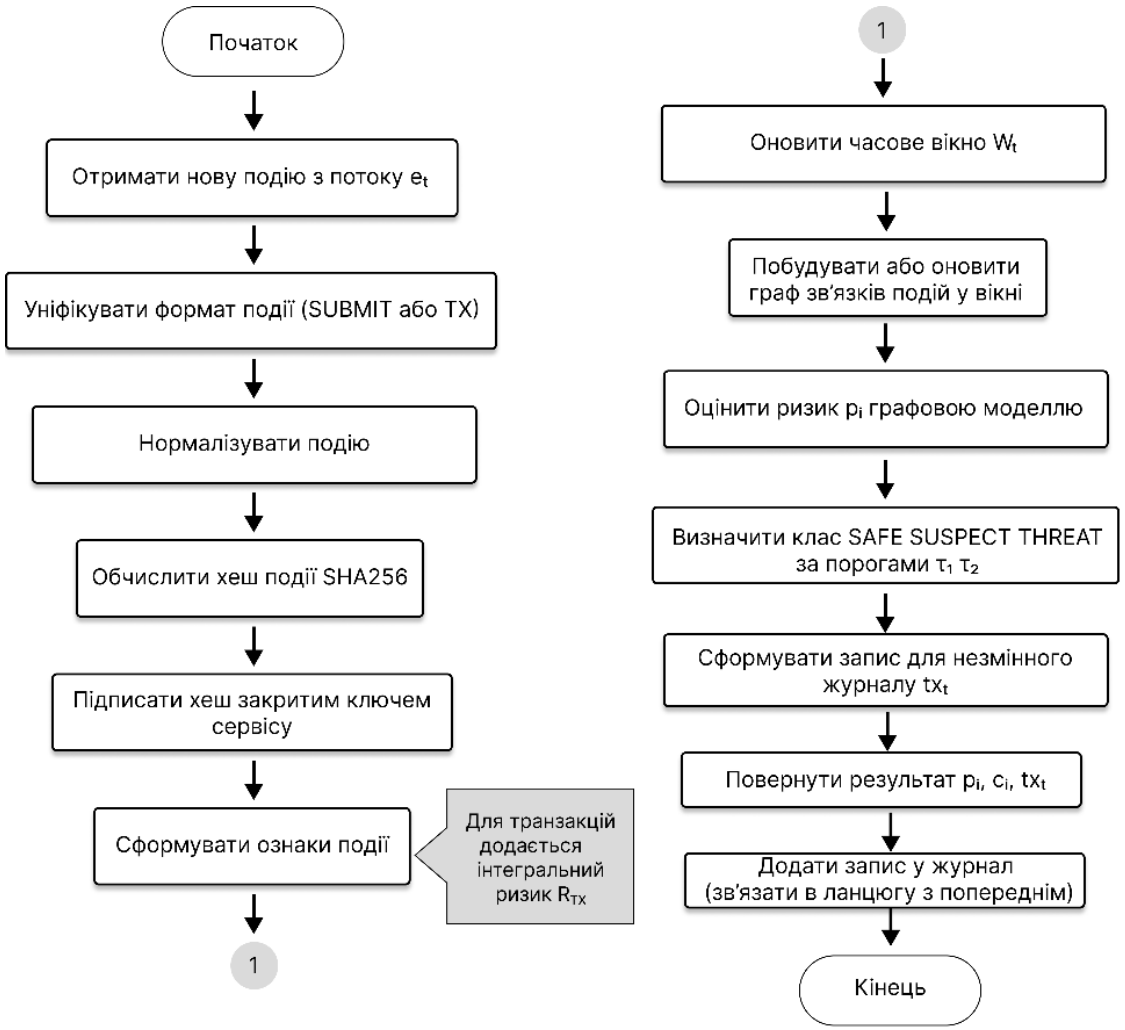


**Рис. 1. Узагальнена схема роботи моделі TALM**

Контекст події враховується в межах ковзного часового вікна (6):

$$W_t = [ts - \Delta_W, ts], \tag{6}$$

у якому будується граф зв'язків між подіями та технічними сутностями. На основі цього графа модель обчислює ймовірність ризику $p_i$ та відносить подію до одного з класів:

$$c_i = \begin{cases} SAFE, & p_i < \tau_1, \\ SUSPECT, & \tau_1 \le p_i < \tau_2, \\ THREAT, & p_i \ge \tau_2, \end{cases} \tag{7}$$

де $\tau_1, \tau_2$ - пороги політики реагування. Для кожної події формується аудитний запис:

$$tx_t = \langle h_e, sig_e, model_{id}, model_{hash}, \tau_{hash}, window_{id}, p_i, c_i, ts, sig_{svc} \rangle. \tag{8}$$

Такий запис фіксує відбиток події, версію моделі, параметри політики та результат оцінювання, що забезпечує виявлення підміни й відтворюваність рішення під час аудиту. Вхід моделі - потік подій $e_t$, а вихід - пара $(p_i, c_i)$ та аудитний запис $tx_t$, який використовується для реагування і перевірки рішень у часі.

Архітектура підсистеми довіри включає дві синхронні лінії: оперативне прийняття рішення для поточної події та формування доказового сліду для подальшої перевірки. Потік починається у вебзастосунку, де виникають події двох типів: надсилання вебформ і транзакційні операції. Далі вони надходять у модуль збору та нормалізації, де приводяться до єдиного формату. Після цього виконується криптографічна фіксація події: формується канонічний опис, обчислюється хеш і створюється підпис сервісного компонента, що забезпечує виявлюваність підміни даних. На наступному етапі подія перетворюється на ознаковий опис. Для вебформ використовуються змістові, структурні, технічні та часові характеристики, для транзакцій – атрибути операції, техніко-часовий контекст і інтегральний ризик $R_{TX}$. Аналіз виконується не ізольовано, а в межах часового вікна, де будується граф зв'язків між подіями, сесіями, пристроями та мережевими слідами. Графовий модуль формує ризикову оцінку, а політика реагування перетворює її на один із трьох станів: безпечний, контрольований або загрозливий.

Паралельно з рішенням формується компактний аудитний запис, який містить хеш і підпис події, ідентифікатор моделі, параметри політики та ідентифікатор контекстного вікна. Цей запис фіксується у незмінному журналі, що забезпечує перевірку цілісності, валідності підписів і узгодженості версій моделі та політики. Підтверджені інциденти й результати модерації повертаються у контур навчання, де формується оновлена версія моделі. Експериментальна перевірка виконувалась як порівняльне дослідження, у якому оцінювались якість виявлення загроз і властивості доказовості рішень. Дані оброблялися у потоковому режимі з часовим поділом на train, validation і test, що зменшувало ризик витоку інформації. Для порівняння використовувалися три конфігурації: система правил Rule WAF, графова модель GNN та повний контур TALM GNN Ledger.

Результати показали, що графовий підхід істотно покращує якість виявлення порівняно з системою правил, а повний контур TALM забезпечує найкращі значення precision, recall, F1 і найменшу частку хибних спрацювань як для вебформ, так і для транзакцій (табл. 1).

Таблиця 1

**Порівняння якості виявлення загроз для подій вебформ і транзакцій**

| Domain | System | Precision | Recall | F1 | FPR, % |
|---|---|---|---|---|---|
| SUBMIT | Rule WAF | 0.83 | 0.74 | 0.78 | 2.6 |
| SUBMIT | GNN | 0.91 | 0.89 | 0.90 | 1.2 |
| SUBMIT | TALM | 0.93 | 0.91 | 0.92 | 0.9 |
| TX | Rule WAF | 0.79 | 0.70 | 0.74 | 1.9 |
| TX | GNN (без $R_{TX}$) | 0.86 | 0.84 | 0.85 | 1.2 |
| TX | TALM | 0.91 | 0.89 | 0.90 | 0.8 |

Для транзакцій додавання інтегрального ризику $R_{TX}$ дало додаткове покращення метрик (рис. 2).

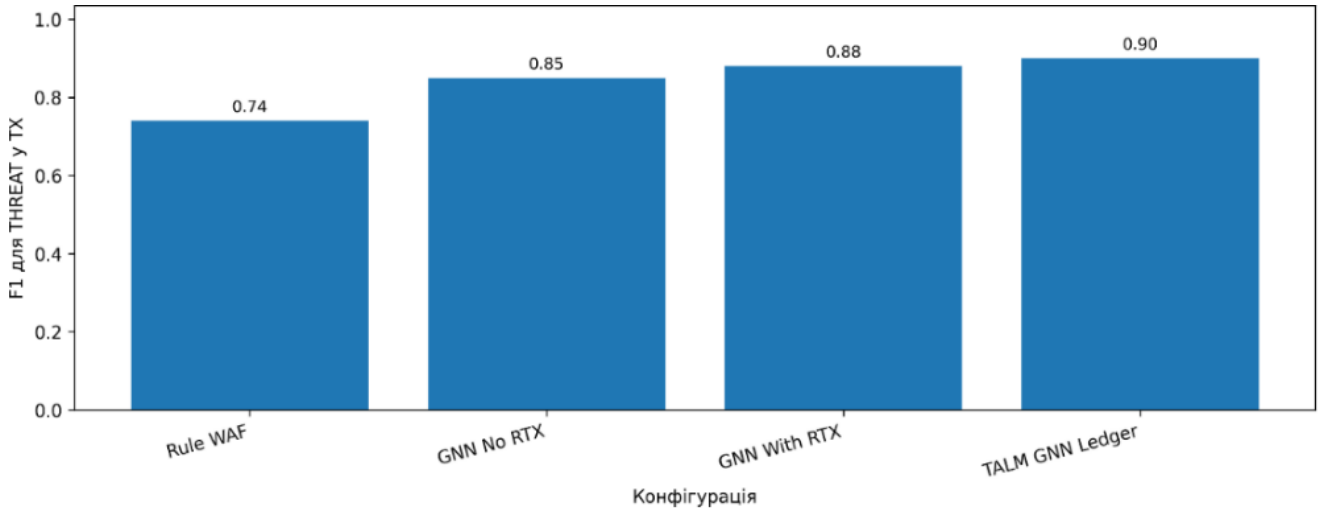


**Рис. 2. Вплив додавання інтегрального транзакційного ризику $R_{TX}$ на F1**

Часові характеристики показали помірне зростання затримки через криптографічну фіксацію і журналювання. Перехід від GNN до TALM збільшив p95 затримку з 52 до 64 мс, але ці витрати залишилися прийнятними для потокових сценаріїв (табл. 2).

Таблиця 2

**Часові характеристики та продуктивність різних конфігурацій системи**

| Система | p50, мс | p95, мс | Швидкість обробки, подій/с | Запис у журнал, мс | Перевірка в аудиті, мс |
|---|---|---|---|---|---|
| Rule WAF | 12 | 28 | 2400 | 0 | 0 |
| GNN | 18 | 52 | 1600 | 0 | 0 |
| TALM | 23 | 64 | 1350 | 8 | 6 |

Поведінка системи за різних сценаріїв навантаження наведена на (рис. 3). Навіть у комбінованому режимі p95 затримка для повного контуру становила 98 мс, що свідчить про збереження передбачуваності роботи системи. Основний приріст якості забезпечує контекстний графовий аналіз, додавання $R_{TX}$ підсилює транзакційний канал, а незмінне журналювання додає помірний оверхед, але водночас забезпечує відтворюваність і аудитну перевірюваність рішень.

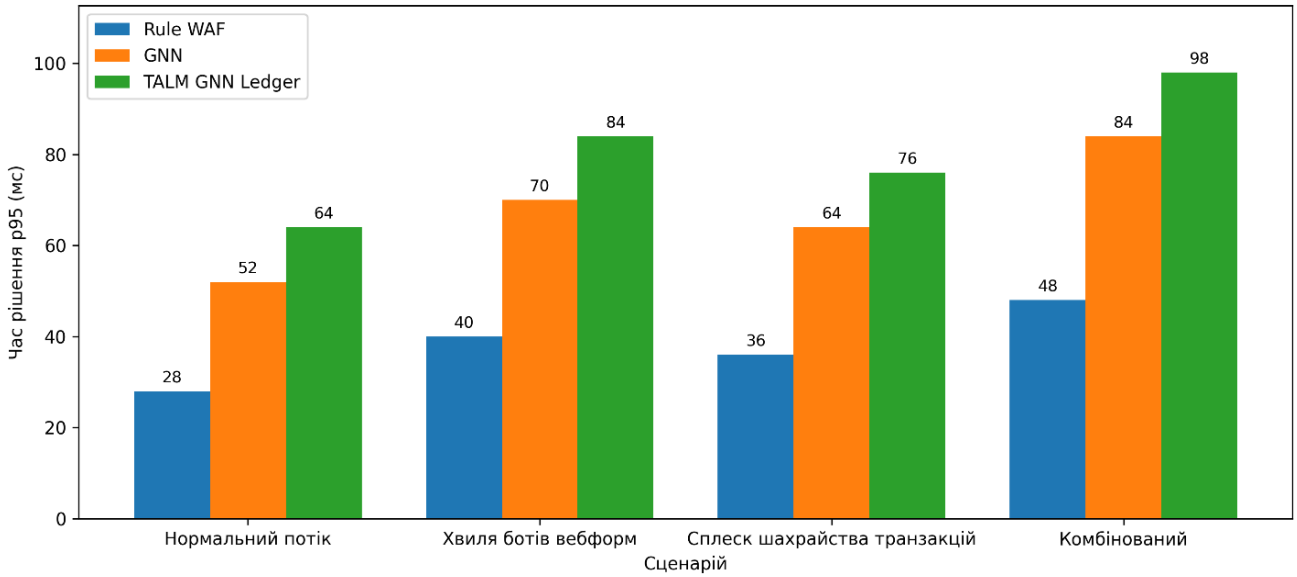


**Рис. 3. p95 затримка прийняття рішення (мс) у різних сценаріях навантаження**

### *Висновки*

Запропонований інтегрований контур довіри поєднує контекстне графове оцінювання ризику для подій вебформ і транзакцій із журналюванням рішень, придатним для перевірки під час аудиту. Експерименти показали, що перехід від правил до графової моделі підвищує якість виявлення та зменшує частку хибних спрацювань; у повному контурі це доповнюється відтворюваністю рішень завдяки прив'язці до версій моделі й параметрів політики. Накладні витрати на доказовість залишаються керованими: p95 затримка для повного контуру зростає з 52 до 64 мс, що зберігає придатність підходу для потокових сценаріїв і сплесків активності. Подальший розвиток доцільно спрямувати на розширення типів подій, підвищення стійкості до адаптивних атак та оптимізацію графового модуля для збереження низьких затримок у високих навантаженнях.

**Декларація про штучний інтелект**

Автори декларують, що штучний інтелект не використовувався для генерування наукового змісту, результатів дослідження, інтерпретації даних або формулювання висновків. Усі наукові положення статті є результатом самостійної роботи авторів.

**Конфлікт інтересів**

Автори заявляють про відсутність конфлікту інтересів та підтверджують, що під час підготовки цієї роботи не існувало жодних комерційних, фінансових чи інших взаємовідносин, які могли б бути розцінені як такі, що здатні вплинути на результати дослідження або їх інтерпретацію. Робота виконана відповідно до принципів академічної доброчесності, етичних норм проведення наукових досліджень та вимог редакційної політики щодо запобігання конфлікту інтересів.

### *Список використаної літератури*

*1. Chong C., Peng, Z., Hartel, P.H. (2003). Secure Audit Logging with Tamper-Resistant Hardware. In: Gritzalis, D., De Capitani di Vimercati, S., Samarati, P., Katsikas, S. (eds) Security and Privacy in the Age of Uncertainty. SEC 2003. IFIP — The International Federation for Information Processing, vol 122. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-35691-4_7*

*2. Pal R. Secure Cloud Storage with Attribute-Based Encryption and Audit Logs / International Journal of Advanced Research in Computer Science and Engineering (IJARCSE). – 2025. – Vol. 1, No. 3. – P. 27–34. – DOI: 10.63345/v1.i3.304.*

*3. M. A. Inam et al., "SoK: History is a Vast Early Warning System: Auditing the Provenance of System Intrusions," 2023 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 2023, pp. 2620-2638, doi: 10.1109/SP46215.2023.10179405.*

*4. Y. Cui, X. Han, J. Chen, X. Zhang, J. Yang and X. Zhang, "FraudGNN-RL: A Graph Neural Network With Reinforcement Learning for Adaptive Financial Fraud Detection," in IEEE Open Journal of the Computer Society, vol. 6, pp. 426-437, 2025, doi: 10.1109/OJCS.2025.3543450.*

*5. Z. Shao, X. Wang, E. Ji, S. Chen and J. Wang, "GNN-EADD: Graph Neural Network-Based E-Commerce Anomaly Detection via Dual-Stage Learning," in IEEE Access, vol. 13, pp. 8963-8976, 2025, doi: 10.1109/ACCESS.2025.3526239.*

*6. Cherif A., Ammar H., Kalkatwi M., Alshehri S., Imine A. Encoder–decoder graph neural network for credit card fraud detection / Journal of King Saud University – Computer and Information Sciences. – 2024. – Vol. 36, Iss. 3. – Art. 102003. – DOI: 10.1016/j.jksuci.2024.102003.*

*7. Li P., Yu H., & Luo X. (2025). Context-aware Graph Neural Network for Graph-based Fraud Detection with Extremely Limited Labels. Proceedings of the AAAI Conference on Artificial Intelligence, 39(11), 12112-12120. https://doi.org/10.1609/aaai.v39i11.33319.*

*8. Kulothungan V. Using Blockchain Ledgers to Record AI Decisions in IoT / IoT. – 2025. – Vol. 6, No. 3. – Art. 37. – DOI: 10.3390/iot6030037.*

*9. Campbell R. Zero Trust for AI Systems: A Reference Architecture and Assurance Framework. Preprints 2026, 2026020085. https://doi.org/10.20944/preprints202602.0085.v1*

*10. Замрій І. В. і Шахматов, І. О. (2024) «Підвищення безпеки вебзастосунків з допомогою інноваційних патернів інтеграції штучного інтелекту», Сучасний стан наукових досліджень та технологій в промисловості, (1(27), с. 67–80. DOI: 10.30837/ITSSI.2024.27.067.*

*11. Шахматов І. О., Замрій І. В. Integrated security systems for protecting payment synchronization from MITM attacks / Problems in programming. – 2025. – № 2. – С. 28–39. – DOI: 10.15407/pp2025.02.028.*

*12. Mishra P. Securing e-governance against shadow attacks with blockchain technology. Sci Rep 15, 42306 (2025). https://doi.org/10.1038/s41598-025-26326-0*

*13. Dowling B., Günther F., Herath U., Stebila D. (2016). Secure Logging Schemes and Certificate Transparency. In: Askoxylakis, I., Ioannidis, S., Katsikas, S., Meadows, C. (eds) Computer Security – ESORICS 2016. ESORICS 2016. Lecture Notes in Computer Science(), vol 9879. Springer, Cham. https://doi.org/10.1007/978-3-319-45741-3_8*

*14. Guardiola-Múzquiz G., Soriano-Salvador E. SealFSv2: combining storage-based and ratcheting for tamper-evident logging. Int. J. Inf. Secur. 22, 447–466 (2023). https://doi.org/10.1007/s10207-022-00643-1*

*15. Arunraju Chinnaraju Kannan Avalurpet Loganathan (2026)Trustworthy Agentic Supply Chains: A Governance Framework for Digital Twin Orchestrated AI Decisioning Under Compliance, Auditability, and Data Sovereignty ConstraintsInternational Journal of Latest Technology in Engineering Management & Applied Science. 10.51583/IJLTEMAS.2026.150100018. 15:1.(245-318). Online publication date: 23-Jan-2026*

*16. Ojewale V., Suresh H., Venkatasubramanian S. Audit Trails for Accountability in Large Language Models / arXiv. – 2026. – arXiv:2601.20727 [cs.CY]. – DOI: 10.48550/arXiv.2601.20727.*

*17. Kamimura, Tokachi, Hybrid Post-Quantum Signatures for Tamper-Evident Audit Trails: Formal Security Analysis and Design Trade-offs (December 08, 2025). Available at SSRN: https://ssrn.com/abstract=5883842 or http:/dx.doi.org/10.2139/ssrn.5883842*

*18. Gade U. R. Designing a Ledger-Centric, Event-Driven Architecture for Consistent and Scalable Systems / Sarcouncil Journal of Engineering and Computer Sciences. – 2025. – Vol. 4, Iss. 9. – DOI: 10.5281/zenodo.17140937.*

*19. Kao L. Post-Quantum-Resilient Audit Evidence for Long-Lived Regulated Systems: Security Models, Migration Patterns, and Case Study // arXiv. – 2025. – arXiv:2512.00110 [cs.CR]. – DOI: 10.48550/arXiv.2512.00110.*

*20. S. Gummadi "Recursive Transaction Hash Chains for Immutable Audit Trails in Mortgage Platforms", IJETCSIT, vol. 6, no. 4, pp. 49–54, Oct. 2025, doi: 10.63282/3050-9246.IJETCSIT-V6I4P107.*

*21. Afiqah Azahari. Reliability of Application-Generated Data for Security Evidence. Computer Science [cs]. Sorbonne Université, 2025.*

*I. Shakhmatov*

**INTEGRATED TRUST CONTOUR IN WEB APPLICATIONS BASED ON GRAPH RISK ASSESSMENT AND IMMUTABLE DECISION LOGGING**

*Violations of trust and integrity in web applications often arise at the boundary between data input and the execution of functional tasks, when malicious or anomalous input from a web form or API is transformed into actions within business logic and transactional processes. This paper treats such episodes as a unified event stream and proposes an integrated trust contour that combines contextual risk assessment on an interaction graph with verifiable recording of outcomes in an immutable log. Risk is assessed by a graph-based model that captures event relationships within a sliding time window among the actor (user or service), session, device, and aggregated network signals. To reconcile heterogeneous sources, we employ a unified event representation and a common feature engineering approach: for web forms, this includes semantic, structural, and temporal characteristics; for*

*transactions, it involves an explainable composite risk score aggregating the effects of amount, time, geospatial deviations, and payment instrument type. Verifiability of decisions is ensured through event normalization, hashing, and digital signing by the logging service prior to recording a minimally sufficient set of attributes in an immutable journal, namely event hashes, identifiers, model version checksums, response policy parameters, and final risk scores. This approach enables the detection of history tampering and supports the reconstruction of decision rationale within the scope of the recorded data, which is particularly valuable for auditing and incident investigation in web environments.*

**Keywords:** web security, graph neural networks, spam detection, immutable log, audit, model versioning, software, architecture.

Надійшла до редакції: 25.02.2026
Прийнята до друку: 21.04.2026
Опубліковано: 27.04.2026