SENTINEL-Chain: фреймворк для безопасной публикации медицинских данных с интеграцией блокчейна и защитой конфиденциальности

Источник: Frontiers in Digital Health

Оригинал: https://www.frontiersin.org/articles/10.3389/fdgth.2026.1807540

аналитика данных | блокчейн | кибербезопасность | конфиденциальность данных | электронные медкарты

Введение

Электронные медицинские карты (EHR — Electronic Health Records) играют центральную роль в аналитике здравоохранения, однако их высокая степень детализации при обмене данными повышает риск повторной идентификации пациентов. Традиционные методы обеспечения конфиденциальности, включая **k-анонимность (k-anonymity)**, **l-разнообразие (l-diversity)** и **дифференциальную приватность (differential privacy)**, зачастую защищают конфиденциальность в ущерб аналитической полезности, ослабляя клинически значимые корреляции.

Методы

Мы предлагаем **SENTINEL-Chain** — интегрированную с блокчейном платформу для обеспечения конфиденциальности при безопасной публикации **EHR**. Уровень конфиденциальности объединяет шесть механизмов: **адаптивное возмущение с учетом корреляций (ACAP — Adaptive Correlation-Aware Perturbation)**, **иерархическую многоуровневую обобщающую трансформацию (HMGG — Hierarchical Multi-Granularity Generalization)**, **семантически-ориентированную**

анатомизацию (SAA — Semantic-Aware Anatomization), **вероятностное подавление с границами полезности (PSUB — Probabilistic Suppression with Utility Bounds)**, **гео-временную неразличимость (GTI — Geo-Temporal Indistinguishability)** и **ансамблевую композицию конфиденциальности (EPC — Ensemble Privacy Composition)**. Блокчейн-уровень добавляет верификацию на основе **дерева хешей Меркла (Merkle Hash Tree)**, валидацию на базе протокола **PBFT (Practical Byzantine Fault Tolerance)**, проверку соответствия с помощью **доказательств с нулевым разглашением (zero-knowledge proof)** и контроль доступа на основе **смарт-контрактов**. Для оценки использовался синтетический набор данных (10 000 записей) и два реальных клинических эталонных набора (Wisconsin Breast Cancer, N = 569; Diabetes, N = 442).

Результаты

SENTINEL-Chain достигает показателя конфиденциальности 79,9% и полезности 98,2%, обеспечивая комбинированный показатель 178,1%, что превышает все 16 базовых моделей на 4%–95%. Точность корреляции достигает 99,9% для сумм страховых выплат, 99,6% для длительности пребывания в стационаре, 99,7% для возраста и 99,1% для индексов тяжести состояния. Платформа демонстрирует 100% устойчивость к атакам методом связывания записей (record linkage attacks), при этом преимущество злоумышленника при атаке методом вывода принадлежности (membership inference attack) остается ниже базового уровня случайного угадывания. Блокчейн-уровень обрабатывает 9 988 транзакций в 101 блоке с полной проверкой целостности. Формальная композиция **дифференциальной приватности Реньи (Renyi DP)** дает значение $\varepsilon = 7{,}08$ ($\delta = 10^{-5}$), а пропускная способность достигает примерно 3 600 записей в секунду при объеме до одного миллиона записей.

Обсуждение

SENTINEL-Chain устраняет пять выявленных пробелов в публикации медицинских данных: разрушение корреляций, разрыв между механизмами конфиденциальности и блокчейном, хрупкость использования единственного метода, верификацию без раскрытия данных и ограниченную оценку устойчивости к атакам. Оценка стоимости газа для смарт-контрактов в сети **Ethereum** указывает на затраты в размере 61 895 единиц газа на регистрацию одной записи; развертывание в **Layer-2** (втором уровне масштабирования) позволило бы снизить эти затраты в 10–100 раз.