

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет радиофизики и компьютерных технологий
Кафедра интеллектуальных систем

Аннотация к магистерской диссертации

**«Методы обработки транзакционных данных для
оптимизации бизнес-процессов»**

специальность 1-98 80 03 «Аппаратное и программно-техническое
обеспечение информационной безопасности»

Дрозд Павел Сергеевич

Научный руководитель: д.т.н., профессор кафедры интеллектуальных
систем, Чернявский Александр Федорович

Минск, 2020

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

ПЕРСОНАЛЬНЫЕ ДАННЫЕ, КЛАСТЕРНЫЙ АНАЛИЗ, НЕЙРОННЫЕ СЕТИ, ЛОКАЛЬНО-ЧУВСТВИТЕЛЬНОЕ ХЕШИРОВАНИЕ, BIG DATA.

Цель работы - предложить новые эффективные алгоритмы К-анонимизации транзакционных данных на основе кластерного анализа и реализовать их в системе автоматической анонимизации.

Задачи - изучить проблему безопасной передачи персональных данных, рассмотреть существующие методы обеспечения К-анонимности, предложить новые алгоритмы К-анонимизации и L-разнообразия, разработать SaaS-систему анонимизации транзакционных данных.

Объект исследования - транзакционные наборы данных, содержащие персональную информацию клиентов.

В работе предлагаются новые методы обеспечения К-анонимности и L-разнообразия транзакционных данных. Методы основываются на алгоритмах кластерного анализа для предварительного разбиения исходных данных на домены и последующей генерации классов эквивалентности. Первый метод использует самоорганизующиеся карты Кохонена, второй – локально-чувствительное хеширование.

Реализована SaaS-система автоматической анонимизации транзакционных данных, предназначенная для использования розничными компаниями в странах с жёсткими законодательствами по регулированию и локализации данных.

Магистерская диссертация содержит 87 страниц, 33 рисунка, 12 таблиц, 20 использованных источников, 3 приложения.

АГУЛЬНАЯ ХАРАКТАРЫСТЫКА ПРАЦЫ

ПЕРСАНАЛЬНЫЯ ЗВЕСТКІ, КЛАСТАРНЫ АНАЛІЗ, НЕЙРОНАВЫЯ СЕТКІ, ЛАКАЛЬНА-АДЧУВАЛЬНАЕ ХЭШАВАННЕ, BIG DATA.

Мэта працы - прапанаваць новыя эфектыўныя алгарытмы К-ананімізацыі транзакцыйных дадзеных на аснове кластарнага аналізу і рэалізаваць іх у сістэме аўтаматычнай ананімізацыі.

Задачы - вывучыць праблему бяспечнай перадачы персанальных звестак, разгледзець існуючыя метады забеспячэння К-ананімнасці, прапанаваць новыя алгарытмы К-ананімізацыі і L-разнастайнасці, распрацаваць SaaS-сістэму ананімізацыі транзакцыйных дадзеных.

Аб'ект даследавання - транзакцыйныя наборы дадзеных, якія змяшчаюць персанальныя звесткі кліентаў.

У працы прапануюцца новыя метады забеспячэння К-ананімнасці і L-разнастайнасці транзакцыйных дадзеных. Метады грунтуюцца на алгарытмах кластарнага аналізу для папярэдняга разбіцця зыходных дадзеных на дамены і наступнай генерацыі класаў эквівалентнасці. Першы метады выкарыстоўвае карты Коханэну, другі - лакальна-адчувальнае хэшаванне.

Створана SaaS-сістэма аўтаматычнай ананімізацыі транзакцыйных дадзеных, прызначаная для выкарыстання рознічнымі кампаніямі ў краінах з жорсткімі заканадаўствамі па рэгуляванні і лакалізацыі дадзеных.

Магістарская дысертацыя змяшчае 87 старонак, 33 малюнка, 12 табліц, 20 выкарыстаных крыніц, 3 дадаткі.

GENERAL DESCRIPTION OF WORK

PERSONALLY, IDENTIFIABLE INFORMATION, CLUSTER ANALYSIS, NEURAL NETWORKS, LOCALITY-SENSITIVE HASHING, BIG DATA.

Objective - propose new effective K-anonymization algorithms of transactional data using cluster analysis and implement them in automatic anonymization system.

Tasks - research the problem of secure PII data transmission, review existing K-anonymization algorithms, propose new K-anonymization and L-diversity algorithms, implement SaaS-system for automatic transactional data anonymization.

The object of research - transactional datasets with personally identifiable information.

This research proposes new techniques for transactional data K-anonymization. The techniques are based on cluster analysis algorithms, which create similarity domains from origin dataset, and then equivalence classes are generated. The first one utilizes Kohonen's Self-Organizing Maps, the second one – Locality-Sensitive Hashing.

SaaS-system for automatic transactional data anonymization was implemented. It is intended for use by retail companies in countries with strict data residency regulations.

Thesis: 87 pages, 33 figures, 12 tables, 20 sources, 3 applications