

АРХИТЕКТУРА СИСТЕМЫ ИНТЕГРАЦИИ ГЕТЕРОГЕННЫХ ДАННЫХ ДЛЯ АНАЛИЗА ДТП

Волков А.А. (Университет ИТМО)

Научный руководитель – к.т.н, Тесля Н.Н.
(Университет ИТМО)

В статье рассмотрен вопрос интеграции гетерогенных данных для анализа ДТП в рамках систем умного города. Выявлены недостатки существующих подходов в использовании открытых источников данных в проектах "Карта ДТП", "ХоумХаб". Авторы предлагают новую архитектуру, основанную на онтологиях и функциональных возможностях баз данных, чтобы повысить точность анализа ДТП на примере системе умного города Km4City. Эта архитектура способствует более качественному анализу данных, что, в свою очередь, улучшит предложения по улучшению транспортной инфраструктуры и предотвращению аварий.

Введение. Анализ дорожно-транспортных происшествий (ДТП) играет важную роль в повышении безопасности и эффективности городской транспортной инфраструктуры. Системы умного города, интегрирующие гетерогенные данные из различных источников, включая государственные базы данных и информацию в реальном времени, предоставляют уникальные возможности для глубокого анализа и понимания причин аварий. Эти системы становятся ценным инструментом для городских властей и жителей, способствуя разработке целевых мер по предотвращению ДТП.

Основная часть. В России проект "Карта ДТП" показывает, как интеграция данных ГИБДД с картографическими сервисами может улучшить качество анализа мест происшествий. Хотя качество и точность исходных данных в виде карточек ДТП ограничивают потенциал анализа, но использование проверки характеристики, связанной с координатами в стороннем сервисе, позволяет повысить качество данных. Аналогично, проект "ХоумХаб" использует OpenStreetMap для оценки качества жилья, но сталкивается с проблемами неточности данных, подчеркивая необходимость улучшения качества исходной информации. Разработанная в рамках исследовательского проекта университета Флоренции система умного города Km4City выполняет сбор, хранение и анализ данных из разнообразных источников с использованием онтологий и базы знаний Virtuoso, что позволяет улучшить качество данных, обеспечить их эффективную интеграцию и анализ. Учитывая опыт и недостатки существующих систем, предлагается разработка архитектуры системы интеграции гетерогенных данных, основанной на использовании онтологий, для улучшения качества анализа ДТП. Это достигается перекрестной проверкой данных, не ограничиваясь одной характеристикой и расширению контекста описания ДТП за счет подключения большого количества источников. Основные направления улучшения включают повышение качества геоданных, агрегацию и семантическую интеграцию информации из различных источников.

Выводы. Исследование существующих проектов выявило ключевые проблемы в области анализа ДТП, связанные с качеством и интеграцией данных. Разработка новой архитектуры системы, учитывающей эти проблемы и предлагающей решения для их преодоления, открывает путь к созданию более безопасных и эффективных городских транспортных систем. Внедрение предложенной системы способно значительно улучшить анализ ДТП, обеспечивая основу для принятия обоснованных решений по улучшению транспортной инфраструктуры.