

**ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ОБРАБОТКИ
ПОЛЬЗОВАТЕЛЬСКИХ ДАННЫХ С
ИСПОЛЬЗОВАНИЕМ АНАЛИЗА
ЕСТЕСТВЕННОГО ЯЗЫКА**

Савенков Павел Анатольевич

Старший преподаватель

Тульский государственный университет, Тула, Россия

E-mail: pavel@savenkov.com

Научный руководитель — Ивутин Алексей Николаевич

Повышение доступности вычислительных мощностей позволило реализовать на практике интеллектуальные системы, способные формировать нетривиальные выводы о деятельности пользователей на основе собранных массивов данных. Данные системы осуществляют — анализ поведения пользователей и именуются как UBA (User Behavior Analytics). Применение анализа естественного языка, в реализации данных решений, позволяет идентифицировать глубокие зависимости, включая поиск изменений в поведении пользователей [1]. В данный момент актуальной является задача реализации подобных систем именно в мобильном исполнении, ввиду смещения акцента анализа в сторону применения мобильных устройств [2]. Между тем, обрабатываемые на мобильных устройствах данные, существенно отличаются от собираемых с персональных компьютеров, что требует выбора источников данных, их предварительной обработки и очистки от информационного шума. Специфика работы пользователей на мобильных устройствах накладывает дополнительные условия. Диапазоны агрегации наборов данных и их объем являются параметрами, влияющими на качество анализа.

Предлагаемое в данной работе программное решение обеспечивает сбор набираемых пользователями текстов на мобильных устройствах, их предварительную обработку, а применение методов анализа естественного языка, позволяет снизить объем анализируемой вручную информации за счет формирования краткого объема высокоинформативных результирующих значений, при большом количестве входных данных. Разрабатываемая мобильная система идентификации типов сценариев использования мобильных устройств имеет возможности сбора наборов текстовых данных и идентификации нетиповых сценариев на их основе. Архитектура разрабатываемой системы представлена на рис.1.

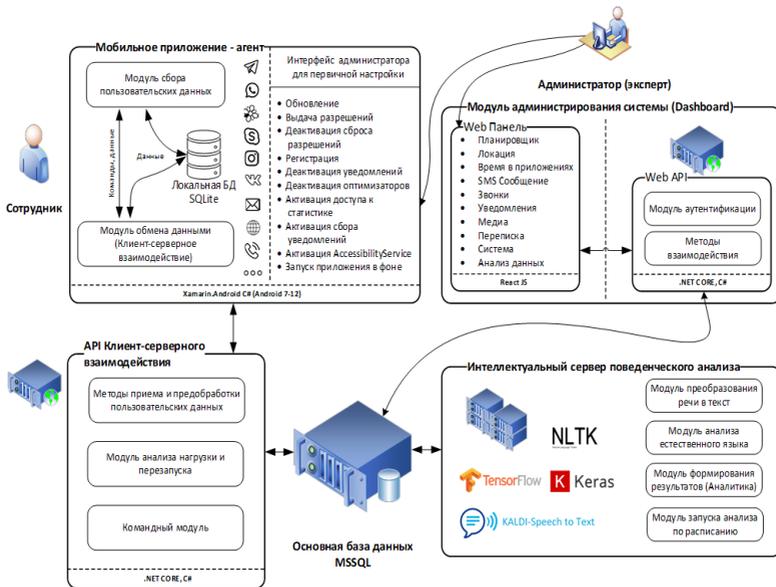


Рисунок 1 - Архитектура системы

Благодаря использованию программного комплекса, организации смогут осуществлять контроль за персональной / конфиденциальной информацией и активностью пользователей, идентифицировать нетиповую активность использования мобильных устройств, для своевременного реагирования и блокирования утечек, оценивать риски и минимизировать их.

Литература

1. Желябин Д. В. Применение методов машинного обучения для решения задачи NLP классификации текста на основе анализа семантики естественного языка // Вестник Алтайской академии экономики и права. – 2020. – №. 6-2. – С. 229-235.
2. Талалаева Г. В., Пермяков К. А. Социальная кибернетика как инструмент комплексной безопасности в эпоху глобальной цифровизации // ФУНДАМЕНТАЛЬНАЯ И ПРИКЛАДНАЯ НАУКА: СОСТОЯНИЕ И ТЕНДЕНЦИИ РАЗВИТИЯ. – 2022. – С. 156-231.