

О. Р. Мысливец
(УО «БГУИР», Минск)

РАЗРАБОТКА МОДУЛЯ СИСТЕМЫ СБОРА И ХРАНЕНИЯ ДАННЫХ СОЦИАЛЬНЫХ СООБЩЕСТВ В ИНТЕРНЕТЕ

В настоящее время существует большое количество социальных сетей различной тематики и направленности, где пользователи размещают информацию о себе. Информация подобного рода может быть полезна при анализе характерных признаков и поведения пользователя. Наиболее распространенными этапами при создании приложений, связанных с анализом данных социальных сообществ, являются сбор и накопление данных, моделирование распространения сети, развитие систем рекомендаций, а также прогнозирование связей и анализ объектов социальных сообществ.

Предлагаемой решение связано с созданием единой системы сбора, хранения и анализа данных социальных сообществ из различных источников, приведение разнородных данных к единому формату

для последующего анализа с помощью алгоритмов обработки больших массивов данных.

При разработке подобной системы встает вопрос получения данных из различных социальных сетей. Настройки конфиденциальности пользователей, ограничения на использование API-функций многих социальных сетей и необходимость эффективного хранения больших объемов информации делает этап сбора данных важным шагом при проектировании и реализации системы.

Разрабатываемый модуль сбора и хранения информации позволит производить сбор данных из наиболее распространенных социальных сетей (Вконтакте, Facebook, LinkedIn) с возможностью получения различной служебной информации о процессе сбора данных через Telegram-бот.

Исходя из вышеперечисленных особенностей, модуль сбора и хранения информации состоит из двух серверов: сервер с СУБД и сервер, обрабатывающий запросы Telegram-бота и предоставляющий советующую информацию на запрос. Спроектированный подобным образом модуль сбора и хранения информации позволит оперативно реагировать на непредвиденные ситуации в процессе сбора данных, что позволит, с учетом уже имеющихся у многих социальных сетей ограничений на количество запросов в секунду, избежать лишней траты времени.