

УДК 004.942

## ОБЗОР СОВРЕМЕННЫХ МЕТОДОВ АНАЛИЗА ПОВЕДЕНИЯ ТУРИСТОВ

Михайлов С.А. (Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»)

Научный руководитель – д.т.н., профессор Смирнов А.В.

(Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук)

**Аннотация.** Анализ поведения туристов позволяет повысить эффективность приложений, нацеленных на поддержку туризма за счет учета закономерностей в поведении туристов. Данная работа представляет собой обзор современных методов анализа поведения туристов, включая способы извлечения поведенческой информации туриста и различные модели.

**Введение.** Туристическая индустрия активно развивается на протяжении последних нескольких лет, в том числе и на рынке туристических услуг и приложений. Данная сфера привлекает к себе большое количество исследователей и является объектом научного интереса. Деятельность туристов формирует большой массив данных, который включает в себя различные поведенческие особенности, которые можно использовать для улучшения работы разнообразных туристических сервисов. Анализ поведения туристов позволит улучшить работу рекомендующих систем за счет выделения групп схожих по поведению пользователей; выявлять популярные точки интереса; прогнозировать дальнейшее поведение туриста и т.д.

**Основная часть.** Основным способом получения поведенческой информации о туристах – это автоматический сбор данных из доступных источников и последующая обработка. В качестве таких источников обычно используются: пользовательская информация, информация с устройств пользователя и информация о действиях туристов. Пользовательская информация включает в себя пользовательские фотографии с мета-тегами, а также генерируемую текстовую информацию в социальных сетях. Информация с устройств представляет собой как геопространственные данные, извлекаемые при помощи GPS, так и остальные «сырые» данные извлекаемые при помощи датчиков устройств. Информация о действиях туристов собирается на основе поисковой истории запросов туристов, посещений определенных страниц, списка покупок и т.д. В зависимости от типа полученных данных применяются различные способы фильтрации данных (токенизация и стеммизирование текстовых данных, устранение избыточности в GPS-данных). Для анализа поведения используются следующие методы и модели: Марковские цепи, кластеризация, нейронные сети и другие методы, применяемые в машинном обучении. Марковские цепи используются для обработки последовательной информации и последующего анализа – прогнозирование вероятности посещения следующих туристических мест на основе уже посещенных. Кластерный анализ используется для обнаружения популярных туристических мест и маршрутов. Одним из самых популярных способов кластеризации является метод DBSCAN. Нейронные сети с памятью также могут использоваться для моделирования поведения туристов – к примеру, нейронные сети с долгой краткосрочной памятью (LSTMNN) справляются с моделированием изменяемого с течением времени туристического потока точнее, чем интегрированная модель авторегрессии – скользящего среднего (ARIMA).

**Выводы.** Данная работа представляет обзор современных методов анализа поведения туристов. На ее основе возможно создание новой модели поведения, с учетом текущих научных разработок для последующего применения для дальнейших исследований в области туризма.