

Министерство образования Республики Беларусь  
Учреждение образования  
«Белорусский государственный университет  
информатики и радиоэлектроники»

УДК 004.94:629.7.06(086.44)

Иванин  
Никита Сергеевич

Автоматизированное конструирование карт на основе GPS траекторий

**АВТОРЕФЕРАТ**

магистерской диссертации на соискание степени  
магистра

по специальности 1-40 80 04 – Информатика и технологии  
программирования

---

Научный руководитель  
Волорова Н.А.  
Кандидат технических наук,  
доцент

---

Минск 2021

## ВВЕДЕНИЕ

Для автоматизированного конструирования карт нужно решить две задачи: конструирования карт и сопоставления карт, которая позволит улучшить полученную карту. Для построения карт будут использоваться GPS траектории. Сначала дадим определение траектории.

Траектория  $Tr$  – это последовательность пространственных точек(местоположений транспортного средства)  $Tr = \{p_1, p_2, \dots, p_n\}$ , записанные некоторым GPS-устройством, установленном в транспортном средстве. Каждая точка  $p_i$  состоит из 2-х координат: долготы и широты и времени измерения координат, т.е.  $p_i = \{lat, lon, t\}$ . Точки в  $Tr$  упорядочены хронологически, т. е.  $\forall i < j, t_i < t_j$  и каждые две последовательные точки  $p_i p_{i+1}$  соединены и образуют сегмент траектории.

При решении задачи построения карт строится дорожная карта(дорожная сеть). По определению, дорожная сеть представляет собой ориентированный граф  $G = (V, E)$ , в котором вершина  $v = (x, y) \in V$  представляет перекресток или конец дороги, а дуга  $e = (v_1, v_2)$  – это направленный сегмент дороги, соединяющей вершины  $v_1$  и  $v_2$ .

Таким образом, построение карт – это процесс нахождения исходной дорожной сети на основе коллекции GPS траекторий. Входными данными задачи построения карт является набор траекторий, а выходными данными – дорожная сеть.

Пусть дана дорожная сеть  $G = (V, E)$ , и траектория  $Tr$ , задача сопоставления карт заключается в нахождении исходного маршрута  $Route_{tr}$  в сети  $G$ , представляющий собой последовательность дорожных сегментов(дуг дорожной сети), пройденных транспортным средством.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

### Актуальность темы исследования

В последние годы растущая популярность устройств позиционирования, особенно устройств, оснащенных GPS, позволила непрерывно отслеживать местоположение пользователей и транспортных средств. Данные о местоположениях и траекториях движения стали источником данных для различных приложений, основанных на геопозиционировании, таких как навигация, отслеживание автомобилей, управление трафиком и рекомендаций на основе местоположения. Все подобные приложения значительной степени зависят от наличия высококачественных цифровых карт. Однако, несмотря на повсеместное распространение, качество как

траекторий GPS, так и карт остается низким по следующим причинам:

- в результате нестабильности сигналов GPS траектории GPS обычно неточно отображают местоположение объектов;
- поскольку карты редко обновляются из-за значительных затрат на обновление, они, как правило, являются устаревшими.

Ранее дорожные карты составлялись и обновлялись с помощью сложных наземных обследований, что обеспечивало высокую точность и полноту карт. Однако, поскольку наземные съемки, как правило, трудоемки и отнимают много времени, в настоящее время карты обновляются реже, так что последние изменения в дорожных сетях не могут быть своевременно внесены в карты. До 15% дорог ежегодно каким-либо образом меняются по всему миру. Проблема актуальности карт стала решающим фактором, влияющим на надежность приложений, основанных на местоположении и вызывает проблемы, такие как неправильные результаты навигации и сбой отслеживания транспортных средств. Поэтому существует острая необходимость в автоматическом процессе обновления карт, который обеспечивает частое обновление карт с гарантированным качеством.

Другим толчком к развитию системы автоматического построения карт является развитие беспилотных автомобилей. В настоящее время многие производители, исследовательские организации, ИТ-компании занимаются разработкой полностью автономных беспилотных автомобилей. Таким автомобилям необходимы очень точные карты, содержащие все последние изменения дорожной структуры, поскольку неточности могут привести к различным аварийным ситуациям.

Таким образом, задачи построения и сопоставления карт являются на данный момент весьма актуальной темой для проведения исследований.

### **Степень разработанности проблемы**

Проблема построения карт была впервые изучена в начале 2000-х годов Эделькампом и Шредлом. В своем исследовании авторы использовали высококачественные дифференциальные траектории GPS(DGPS). В качестве метода построения карт был выбран простой метод кластеризации(k-средних).

С тех пор были предложены различные подходы, направленные на построение высококачественных карт с использованием GPS-траекторий, в том числе с низкой дискретизацией. В своих работах Ахмед и Биагиони проанализировали существующие алгоритмы построения карт с разных точек зрения. В частности, Ахмед произвел всесторонний обзор существующих алгоритмов, осуществил их категоризацию, оценил качество и эффективность с помощью метрик, а также сравнил на основе описанных метрик.

Метод, описанный Станоевичем, построен на основе алгоритма

кластеризации k-means с некоторыми оптимизациями. Алгоритм, описанный в работе Лю – это метод, в ходе своей работы осуществляющий кластеризацию дорожных сегментов(или, другими словами, ломаных состоящих из некоторого числа точек). В работа Дейвиса описан алгоритм построения карты на основе оценки плотности. Метод, предложенный Биагиони является методом, основанным на KDE. Основное внимание в этой работе направлено на решение проблем несоответствия траекторий. Вместо того, чтобы устанавливать глобальный порог при бинаризации плотности сетки, этот метод предлагает многослойную архитектуру, в которой сетки с разной плотностью попадают в разные слои. Ахмед и др. предложили метод, который использует расстояние Фреше для объединения новой траектории с существующими ребрами дорожной сети. Этот алгоритм может эффективно сопоставлять части траекторий на основании максимального расстояния между траекторией и дорожным сегментом. Алгоритм, предложенный в работе Карагиорду, является походом, относящимся к семейству методов, объединяющих точки перекрестка. На первом шаге вычисляется изменение направления для каждой точки траектории. Далее ищутся все точки, которые находятся вблизи поворотов. Затем такие точки группируются в соответствии с их близостью и заданным максимальным радиусом.

Исследование проблемы сопоставления карт началось в начале 1990-х годов. Одной из первых работ была работа Бернштейна. Лю представил алгоритм инкрементного сопоставления карт, который учитывает пересечения между траекторией и сегментом дороги для определения исходного маршрута. Вэй и др. предлагают использование подобия на основе расстояния Фреше. Алгоритм изначально находит маршруты, расстояние Фреше до траектории которых меньше порогового значения. Жу использует самую длинную общую подпоследовательность (LCSS) для расчета подобия. В частности, алгоритм сегментирует траекторию и находит кратчайший путь на карте для каждой пары конечных точек сегмента. Работа Ньюсона – это первая работа, в которой дается подробное объяснение того, как модель НММ может быть адаптирована в задачах сопоставления карт. В работе Хантера был предложен алгоритм на основе CRF, который может быть применен как в реальном времени, так и автономно. Йин и др. решают задачу сопоставления карт через построение графа на основе траекторий-кандидатов. После извлечения дорог-кандидатов для каждой выборки траекторий алгоритм строит граф, соединяя кандидатов с их кратчайшим путем на карте. Тагучи и др. предлагают алгоритм, который обеспечивает высокую точность за счет использования модели прогнозирования маршрута. Работа Шарата – самая последняя работа в категории сопоставления карт на уровне полос движения, которая обеспечивает сопоставление карт сопоставляя траекторию с полосой движения. Алгоритм сначала идентифицирует полосы движения на

каждой дороге, используя информацию о ширине дороги на карте и соответственно разбивая их на сетки. Затем алгоритм находит сетки полос-кандидатов вокруг наблюдаемого местоположения и оценивает результат сопоставления с сеткой в каждой временной точке.

### **Цель и задачи исследования**

Целью диссертации является разработка архитектуры приложения и методов, позволяющих решать задачу автоматизированного конструирования карт. Входными данными служат GPS траектории, записанные различными транспортными средствами.

Для выполнения поставленной цели в работе были сформулированы **следующие задачи**:

- разработать методы для предварительной обработки траекторий с целью удаления из них точек-аномалий;
- разработать модуль для предварительной обработки данных траекторий;
- разработать модуль для решения задачи построения карт;
- произвести оценку качества работы модуля для построения с помощью метрик;
- разработать модуль для решения задачи сопоставления карт;
- произвести оценку качества работы модуля для сопоставления карт с помощью метрик;
- разработать архитектуру для приложения, осуществляющего решения задач построения и сопоставления карт;
- произвести верификацию полученной архитектуры;
- добавить поддержку разворачивания системы в облаке.

**Объектом** исследования являются цифровые карты.

**Предметом** работы выступают методы автоматизированного построения карт.

**Область исследования.** Содержание диссертационной работы соответствует образовательному стандарту высшего образования второй ступени (магистратуры) специальности 1-40 80 04 «Информатика и технологии программирования».

### **Теоретическая и методологическая основа исследования**

В диссертационной работе были использованы методы математической статистики, теории вероятностей, методы оптимизации и методы вычислительной математики.

**Информационная база** диссертации состояла теоретических и методологических источников, посвященных проблемам построения и сопоставления карт. Среди исследователей проблемы построения карт можно отметить таких ученых, как Ахмед, Биагиони, Станоевич, Ли, Лю, Дейвис, Хэ, Каригиоргу, Эделькамп, Шредл. Задачу сопоставления карт

активно освещали такие авторы, как Кудус, Вей, Жу, Женг, Ньюсон, Гох, Хантер, Лу, Хуан, Тагучи, Шарат.

**Научная новизна** диссертационной работы заключается в создании классификации возможных аномалий(зашумленностей) в GPS траекториях и способах их удаления, описании метода для генерации искусственных траекторий, оценки алгоритмов построения и сопоставления карт для дорожной сети г.Минска.

### **Основные положения, выносимые на защиту**

1. Модель для построения карт.
2. Оценка качества работы модели для построения карт.
3. Модель для сопоставления карт, построенная с применением скрытых марковских моделей.
4. Оценка качества работы модели для сопоставления карт.
5. Модель для генерации искусственных траекторий.
6. Архитектура приложения для автоматизированного построения карт.

**Теоретическая значимость** диссертации заключается в создании классификации возможных аномалий(зашумленностей) в GPS траекториях и способах их удаления, описании метода для генерации искусственных траекторий, оценки алгоритмов построения и сопоставления карт для дорожной сети г.Минска.

**Практическая значимость** диссертации состоит в разработке архитектуры для эффективного решения задачи автоматизированного построения карт и ее реализации.

### **Апробация и внедрение результатов исследования**

Результаты исследования были представлены на V Международной научно-практической конференции BIG DATA and Advanced Analytics, проходившей 13–14 марта 2019 г. в Белорусском государственном университете информатики и радиоэлектроники, а также на конференции Информационные технологии и системы 2019 (ИТС 2019), проходившей 30 октября 2019 г. в Белорусском государственном университете информатики и радиоэлектроники.

### **Публикации**

Основные положения работы и результаты диссертации изложены в двух опубликованной работе общим объемом 5 п.л. (авторский объем 5 п.л.).

**Структура и объем работы.** Структура диссертационной работы обусловлена целью, задачами и логикой исследования. Работа состоит из введения, четырёх глав и заключения, библиографического списка и приложения. Общий объем диссертации – 69 страниц. Работа содержит 3 таблицы, 23 формулы, 29 рисунков. Библиографический список включает 66 наименований.

## **ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ**

Во **введении** рассмотрено современное состояние задачи автоматизированного построения карт, приведены предпосылки появления интереса к этой задаче, определены основные направления исследований, а также сделано обоснование актуальности темы диссертационной работы.

В **первой главе** анализируется качество собранных GPS траекторий, приводится классификация аномалий в траекториях, а также способы удаления найденных аномалий.

Во **второй главе** был представлен всесторонний обзор существующих алгоритмов построения карт. Также в этой главе, описывается модель для построения карт на основе GPS траекторий, которая проверяется экспериментально на различных тестовых данных.

В **третьей главе** были перечислены существующие алгоритмы сопоставления карт. Были обсуждены основные сильные стороны и слабые стороны существующим алгоритмов. Была описана модель для решения задачи сопоставления карт. Эта модель была проверена экспериментально на искусственно сгенерированных тестовых данных с помощью метрик полноты, точности, F-меры, доли правильных ответов.

В **четвертой главе** представлена описана архитектура реализованного приложения.

В **приложении** приведен исходный код полученного программного продукта.

## ЗАКЛЮЧЕНИЕ

В рамках магистерской диссертации было проведено исследование по вопросам качества GPS-траекторий и дорожных карт. В частности, были проведены исследования задачи автоматизированного построения карт. Также была предложена архитектура приложения, которое решает задачи сопоставления и построения карт, при этом решая проблему зашумленности. Основная цель приложения автоматизированного построения карт заключалась в решении следующих подзадач:

- сопоставления записанного GPS трека с реальной дорожной сетью. Путь транспортного средства по реальной дорожной сети находился с помощью применения скрытых марковских моделей;
- построение карт на основе коллекции GPS треков;
- добавление найденных изменений в оригинальную дорожную сеть.

Для используемых моделей были найдены значения параметров, позволяющие успешно работать с GPS траекториями, записанными велосипедистами и пешеходами. Именно эта область представляла

наибольший интерес, поскольку участки карты, не посещаемые автомобилями, имеют наименьшее покрытие картами. Разработанные модели показали удовлетворительные результаты на тестовых данных.

### **Список опубликованных работ**

1. Иванин, Н. С. Сравнение методов извлечения осевых траекторий в задаче обновления карт / Н. С. Иванин // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сборник материалов V Международной научно-практической конференции, Минск, 13–14 марта 2019 г. В 2 ч. Ч. 1 / Белорусский государственный университет информатики и радиоэлектроники; редкол. : В. А. Богуш [и др.]. – Минск, 2019. – С. 257 – 261.

2. Иванин, Н. С. Сравнение алгоритмов кластеризации в задаче обработки GPS данных / Иванин Н. С., Романов А. А. // Информационные технологии и системы 2019 (ИТС 2019) = Information Technologies and Systems 2019 (ITS 2019) : материалы международной научной конференции, Минск, 30 октября 2019 г. / Белорусский государственный университет информатики и радиоэлектроники; редкол. : Л. Ю. Шилин [и др.]. – Минск, 2019. – С. 212 – 213.