

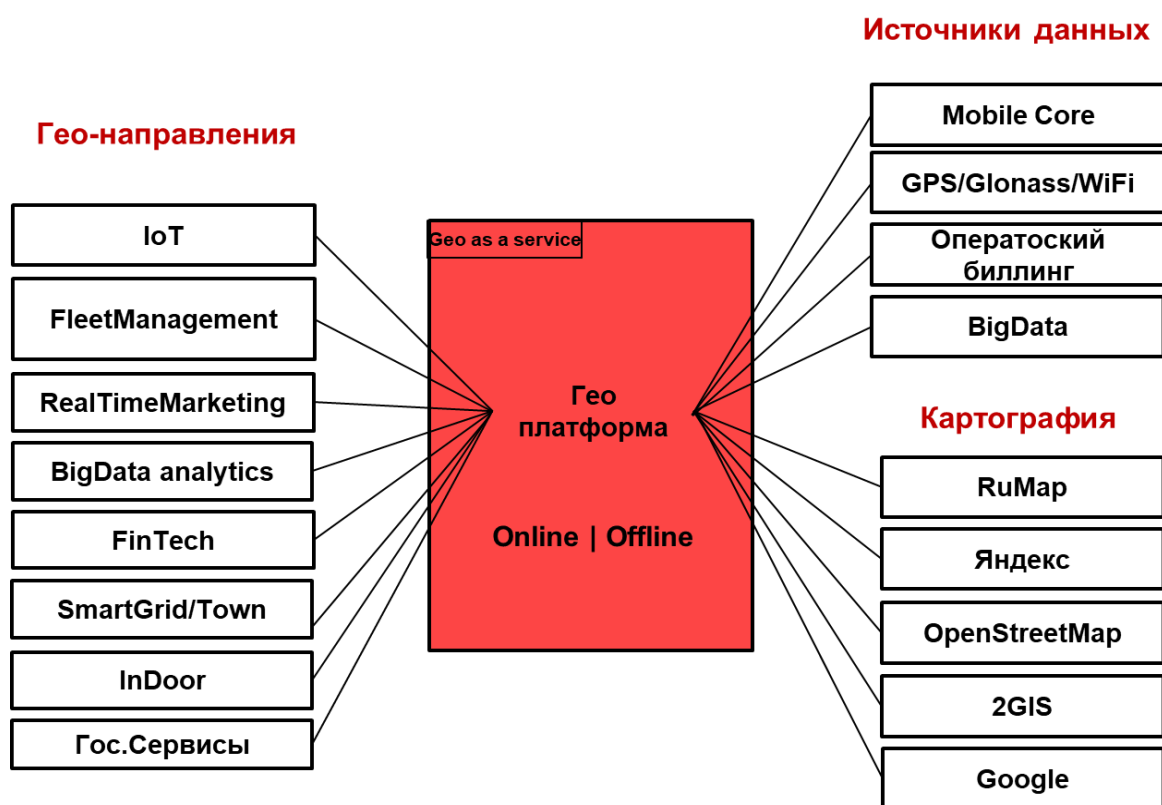
● Быть лучше каждый день

Задача

Гео-сервисы МТС

**Создать механизм по уточнению гео позиции
с учетом накопленных данных**

Направления и структура гео-сервисов МТС



Главное для нас – это точность

МТС Поиск – это:

- ✓ Удобный интерфейс с учетом лучших практик
- ✓ Геопозиция по GPS/Глонас/WiFi/LBS
- ✓ История перемещений
- ✓ Геозоны и уведомления по времени
- ✓ Маршрут до контакта
- ✓ Расширенный контроль телефона ребенка
- ✓ Функционал комьюнити, чат, задачи
- ✓ Кнопка SOS

● Быть лучше каждый день

Цель:

Улучшить точность определения местоположения объекта в online режиме/по запросу, при трекинге с настраиваемым интервалом

Оптимизировать и улучшить отображение истории перемещений, с привязкой к дороге, пешеходной части

Необходимо разработать алгоритм уточнения местоположения объекта учитывая:

- возможные методы, их совместное применение
- все возможные источники данных
- online и offline обработка данных

Дано:

Набор сырых данных системы:

В наборе содержатся координаты для всех пользователей системы за 1 день.

Координаты состоят из точек и окружностей, описывающих вероятное расположение пользователя, а так же временем, когда эта точка была получена.

Координаты в таблице двух видов ($\text{pos_method} = 8192$ и $\text{pos_method} < 8192$)

Координаты имеют в своём составе т.н. «вылеты»(точки, отличающиеся резко возросшей, по сравнению со средней, скоростью пользователя, K и OT точки)

Координаты имеют в своём составе кластера близко расположенных координат, где расстояние между соседними точками $<$ радиуса точки.

Координаты имеют в своём составе повторяющиеся координаты со следующим паттерном: $k_0-k_1-k_2-k_1-k_2-k_1-k_2-...n(k_1-k_2)-k_3$

Координаты, расположенные в границах г.Москва, имеют в своём составе точки, с радиусом ровно 100 метров, расположенные менее, чем в 100м. от координат вестибюлей станций метрополитена(согласно data.mos.ru)

Формат – csv

● Быть лучше каждый день

Необходимо:

- Реализовать алгоритм «сглаживания» трека, удалив из результирующего трека «вылеты» и повторения, описанные выше, а так же объединив кластеры точек по признаку ($d(n_0 - n_1) < r$, где $d(n_0 - n_1)$ – дистанция между точками n_0 и n_1 , а r – это радиус точки n_0).
- Вылетом считается точка, скорость движения «к» и «от» которой $> 200 \text{ км/ч}$
- Вылетом считается точка, если $d(n_0 - n_2) < d(n_0 - n_1)$, где d – дистанция между точками.
- Вылетом считается точка, если координата повторяется более 2 раз по формуле: $k_0 - k_1 - k_2 - k_1 - k_2 - \dots - n(k_1 - k_2) - k_3$ (в этом случае, необходимо убрать повторения k_1 и k_2)
- В результирующем треке, указать тип передвижения пользователя, согласно следующим правилам:
 - ❖ Скорость движения между точками 0-5 км/ч – пешком
 - ❖ Скорость 5-200 км/ч – автомобиль
 - ❖ Начальная и конечная точка сегмента расположена менее, чем в 100м от метро – метро
- Для каждой результирующей точки (при условии, что она является результатом объединения кластера из находящихся близко точек), указать период нахождения в ней

В результате:

- ✓ Демонстрация на макете и тестовых данных работу алгоритма, online и offline сценарии
- ✓ Опционально, интересны предложения для улучшения качества сервиса, дополнительного/нового функционала в направлении гео-сервисы