

Волгоградский Государственный Технический Университет
Факультет электроники и вычислительной техники
Кафедра САПР и ПК

**Метод кластеризации предпочтений жителей
города по перемещению.**

Исполнитель:
Чечеткин И. А.
Руководитель:
Щербаков М. В.

Волгоград 2016

Актуальность. В настоящее время формирование маршрутов в городской среде осуществляется на основе положений, заложенных в городской план развития. Обычно, эта информация достаточно устаревшая и не учитывает предпочтения жителей. На основе данных о предпочтениях жителей требуется разработать эффективный метод кластеризации предпочтений жителей города по перемещению.

Объект исследования – предпочтения жителей города, выраженные в географических координатах.

Предмет исследования – методы кластеризации предпочтений жителей.

Цель работы – разработка метода кластеризации предпочтений жителей для минимизации дискомфорта перемещения в городе.

Теоретические задачи:

- разработка алгоритма кластеризации;
- разработка метода учета географических особенностей местности;
- разработка критериев для оценки качества кластеризации.

Практические задачи:

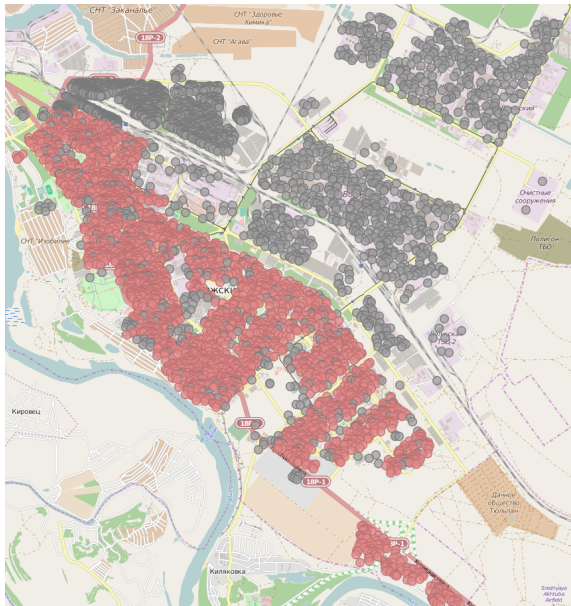
- генерация исходных данных;
- реализация разработанных алгоритмов и методов;
- построение полученных результатов на карте;
- оценка качества кластеризации.

Понятийный аппарат

- **Предпочтение** – пара узлов с определенными координатами и идентификатором пользователя.
- **Node (узел)** – точка с указанными координатами и тегами.
- **Tag (тег)** – пары «ключ – значение».
- **Дискомфорт** – совокупный параметр, определяющий время перемещения из начального узла в конечный.
- **Центроид** – центр тяжести фигуры (геометрический центр).
- **Метрика** – функция, определяющая расстояние в метрическом пространстве.
- **Framework (фреймворк)** – программное обеспечение, облегчающее разработку и объединение разных компонентов большого программного проекта.
- **OpenStreetMap (OSM)** – некоммерческий веб-картографический проект по созданию сообществом подробной свободной и бесплатной географической карты мира.
- **Project OSRM** – фреймворк для вычисления кратчайших путей в графе дорог. Разработан для использования с картографическим сервисом OSM.

- 1 Воронцов, К. В. Машинное обучение. — Режим доступа:
<http://www.machinelearning.ru/>
- 2 Mean Shift Clustering. — Available at:
http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/TUZEL1/MeanShift.pdf
- 3 Comaniciu, D. Mean Shift: A Robust Approach Toward Feature Space Analysis. / D. Comaniciu, P. Meer. — Available at:
<https://courses.csail.mit.edu/6.869/handouts/PAMIMeanshift.pdf>
- 4 Allard, D. Clustering geostatical data. / D. Allard, G. Guillot. — Available at:
http://people.compute.dtu.dk/gigu/article_capetown.pdf
- 5 Scikit learn: clustering. — Available at:
<http://scikit-learn.org/stable/modules/clustering.html>

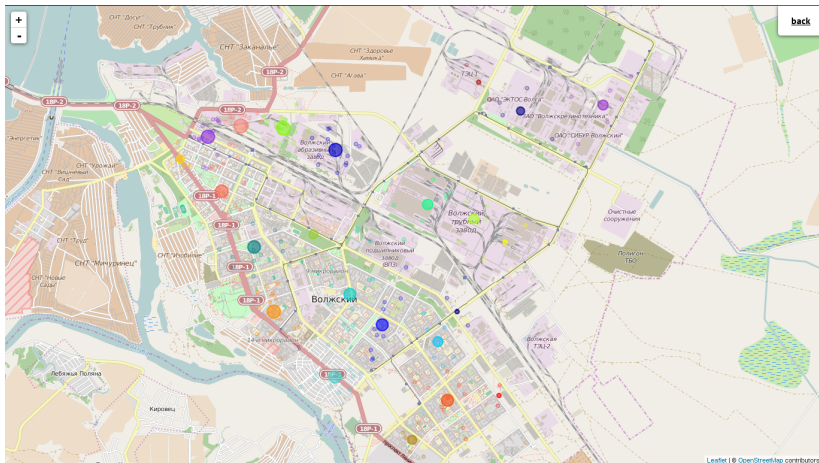
Исходная выборка



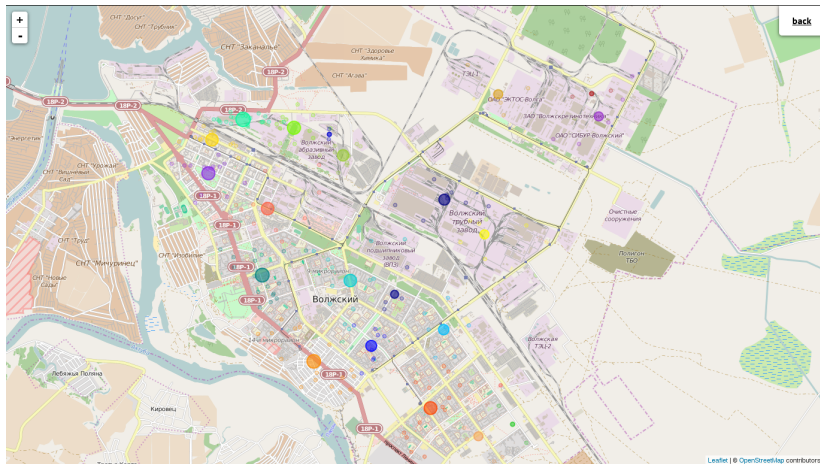
Псевдокод алгоритма Mean Shift:

- 1 Генерирование начального распределения центроидов
- 2 **ПОВТОРЯТЬ**
- 3 Определение соседних точек к центроидам
- 4 Определение среднего веса соседних точек к центроидам
- 5 Расчет нового положения центроидов
- 6 Сдвиг центроидов
- 7 **ПОКА** разница между рассчитанным положением и текущим не будет равна 0
- 8 **ВЫВОД** центроиды

До оптимизации

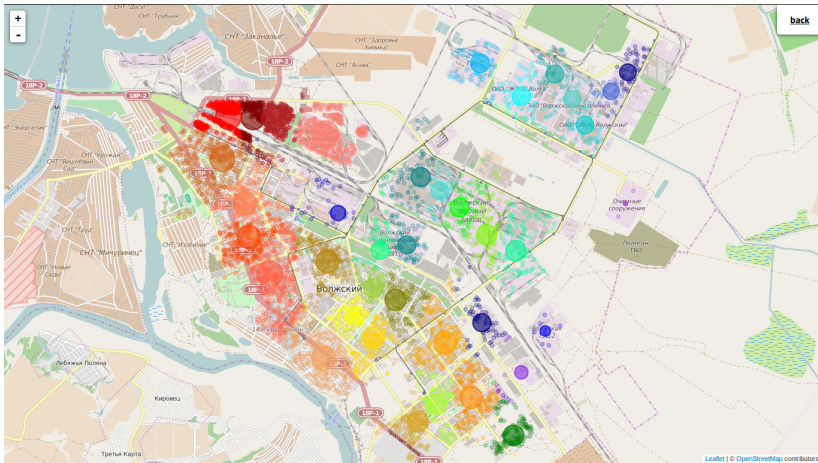


После оптимизации



Реализованный алгоритм:

<https://github.com/vstu-cad-stuff/clustering>



Ссылка на приложение:

<http://vstu-cad-stuff.github.io/clustering/>