

Московский государственный университет имени М.В.Ломоносова
Факультет вычислительной математики и кибернетики

Магистерская программа «Программное обеспечение вычислительных сетей»

Магистерская диссертация

Анализ матриц корреспонденций для описания транспортных потоков

Работу выполнила:
Пучкина Ирина Андреевна

Научный руководитель:
к.ф.м.н. Намиот Дмитрий Евгеньевич

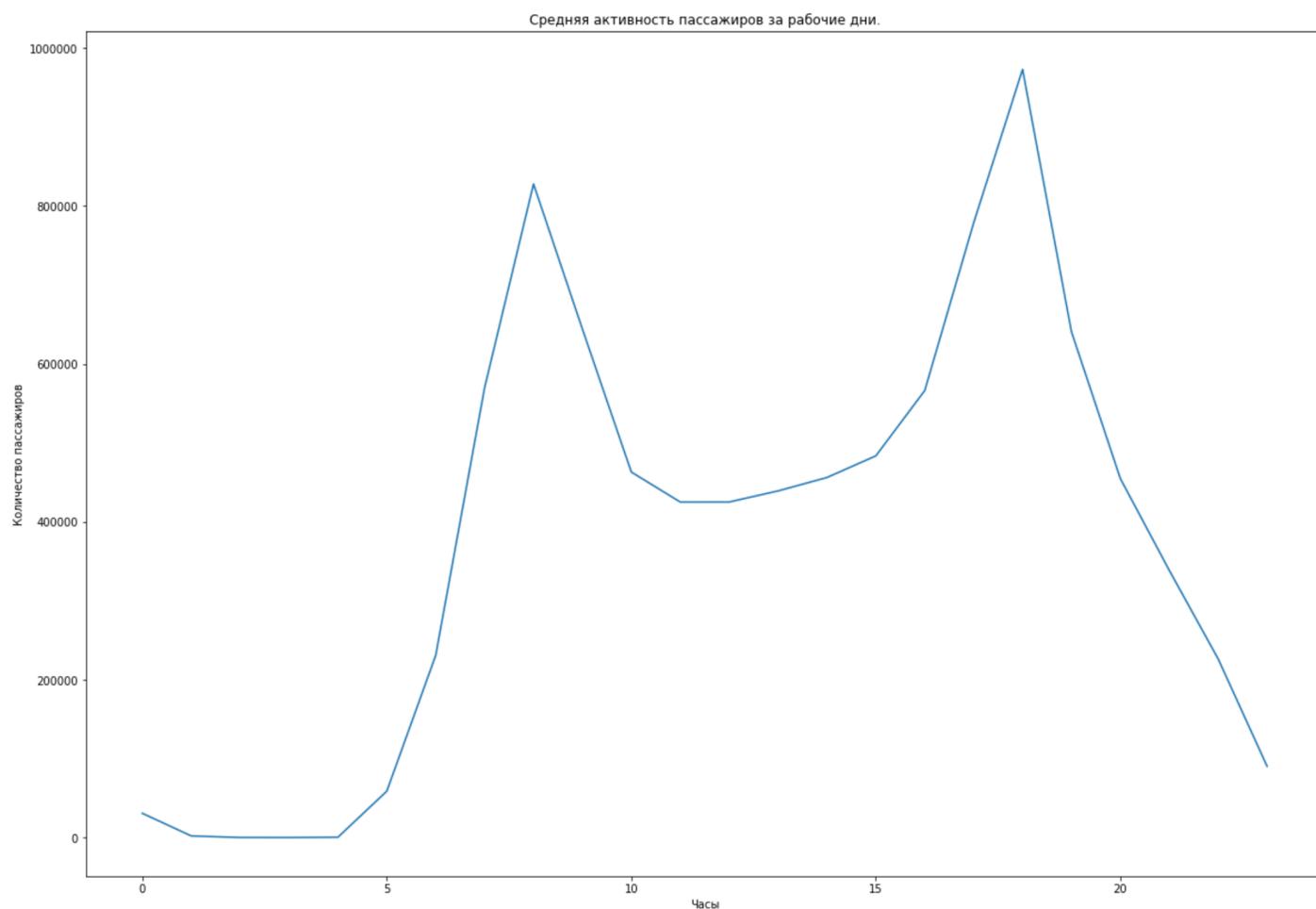


Москва, 2020

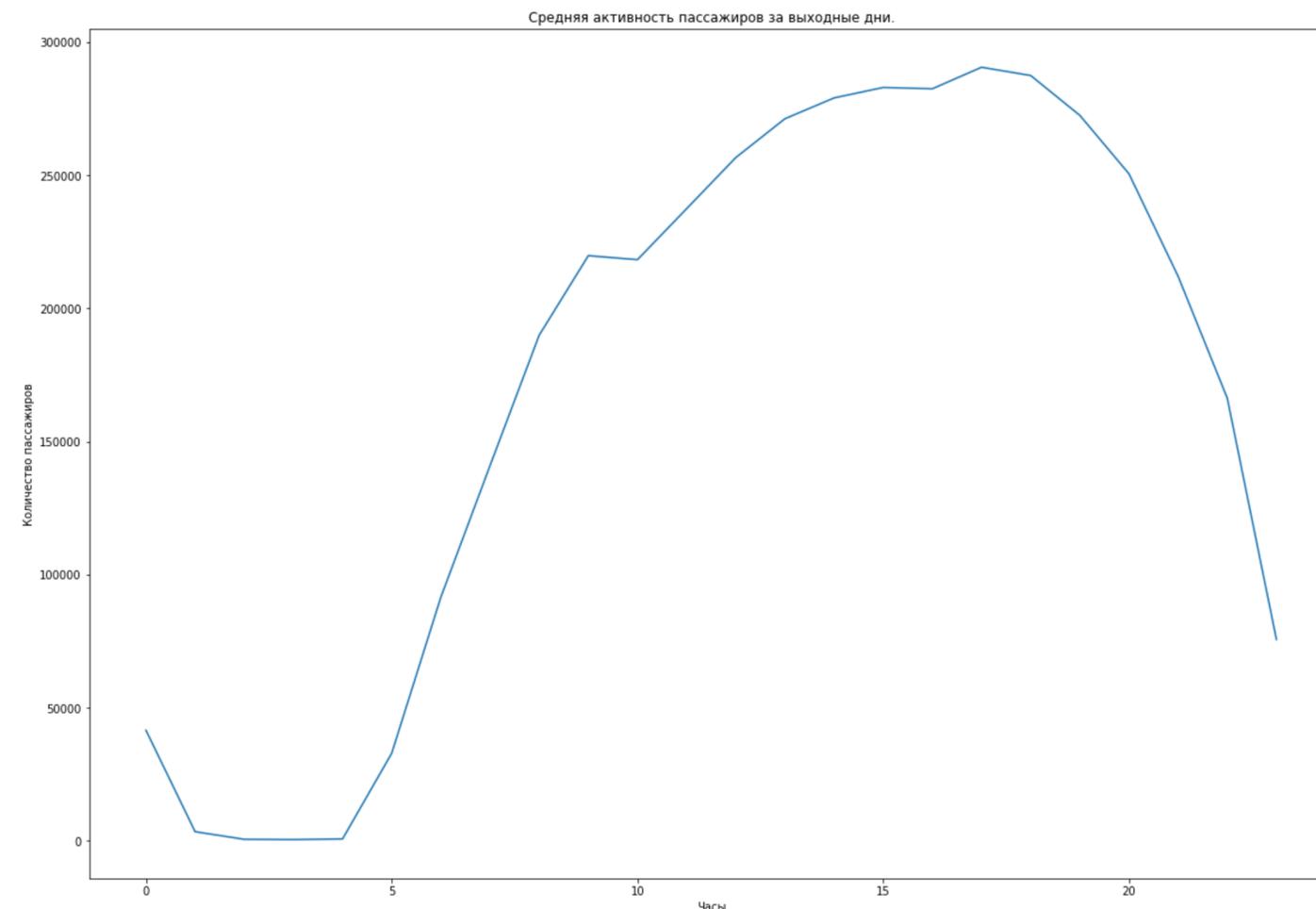
Цель работы

- Научиться определять шаблоны использования станций Московского метрополитена.
- Изучить закономерности соотношения входов и выходов по станциям.
- Предложить способ определения аномалий в данных по перемещениям пассажиров.

Исследование возможных шаблонов

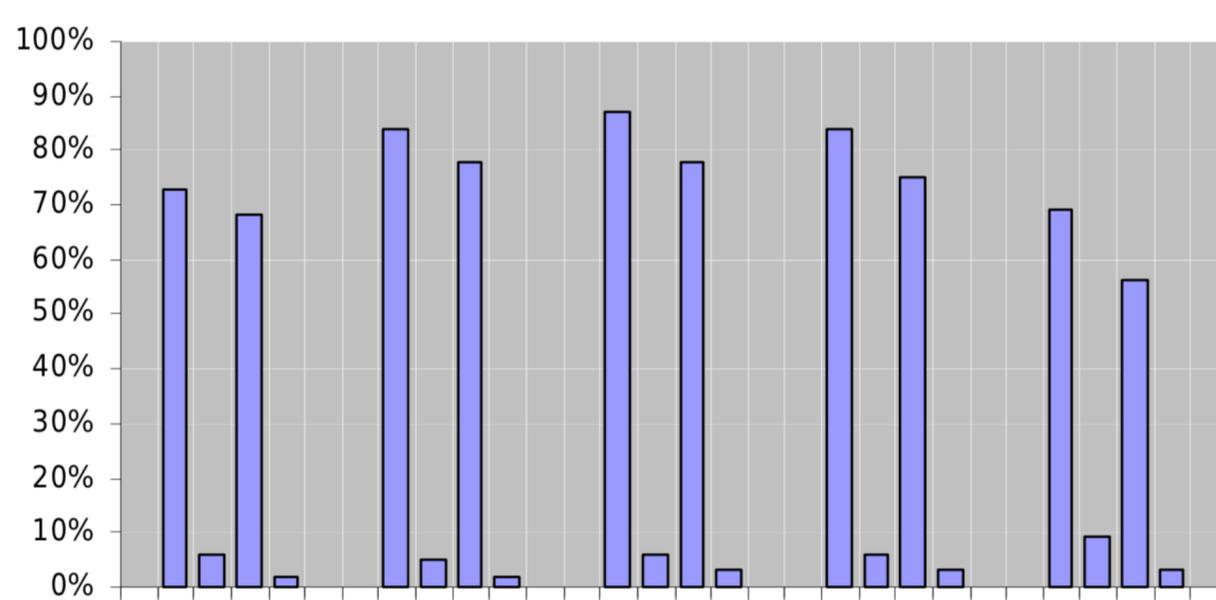


Объем пассажиропотока
в рабочие дни в
распределении по часам

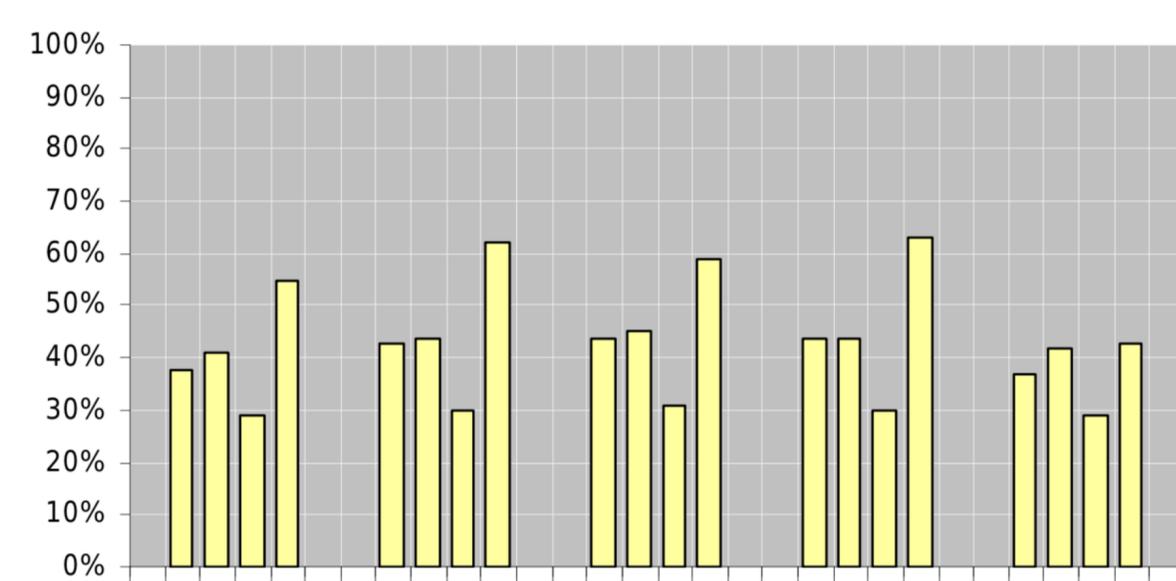


Объем пассажиропотока
в выходные дни в
распределении по часам

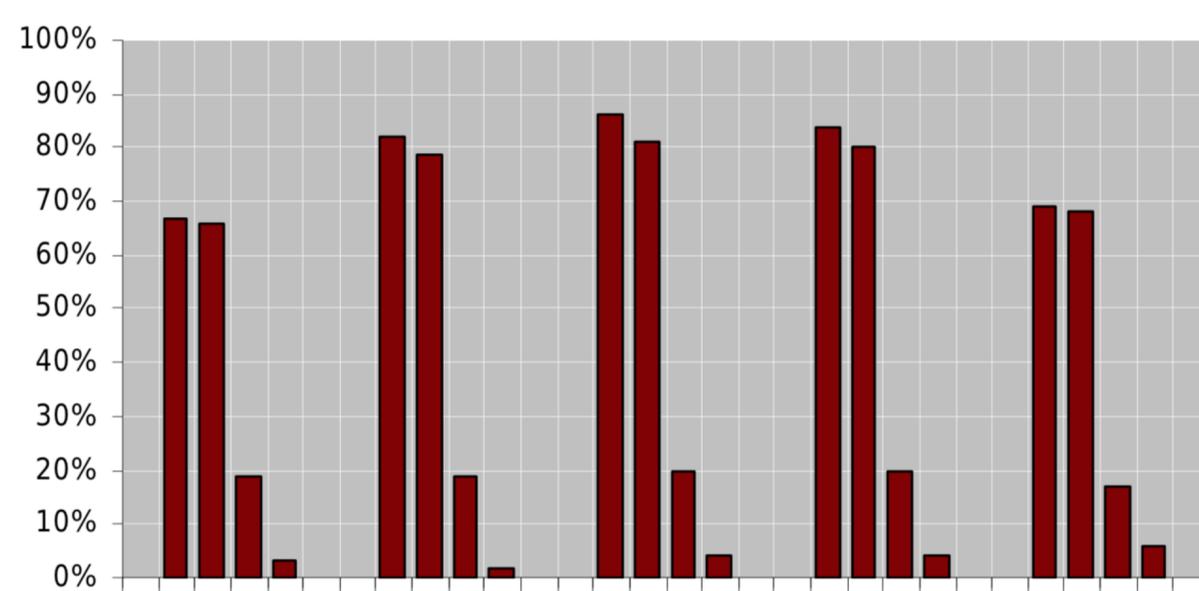
Исследование возможных шаблонов



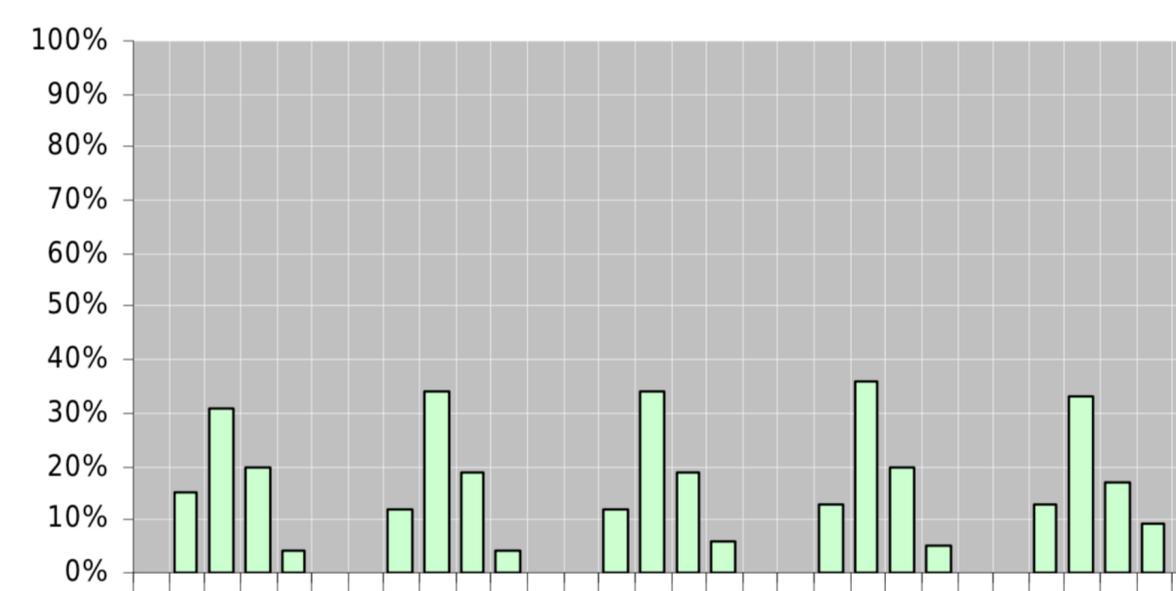
Группа 1: 85,6% взрослых, 13,9% обучающихся, 0,5% пожилых пассажиров



Группа 2: 62,4% взрослых, 36,1% обучающихся, 1,4% пожилых пассажиров

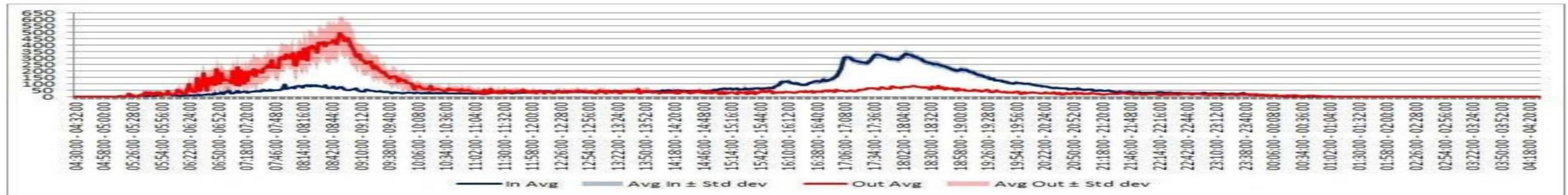


Группа 3: 42,7% взрослых, 55,4% обучающихся, 1,8% пожилых пассажиров

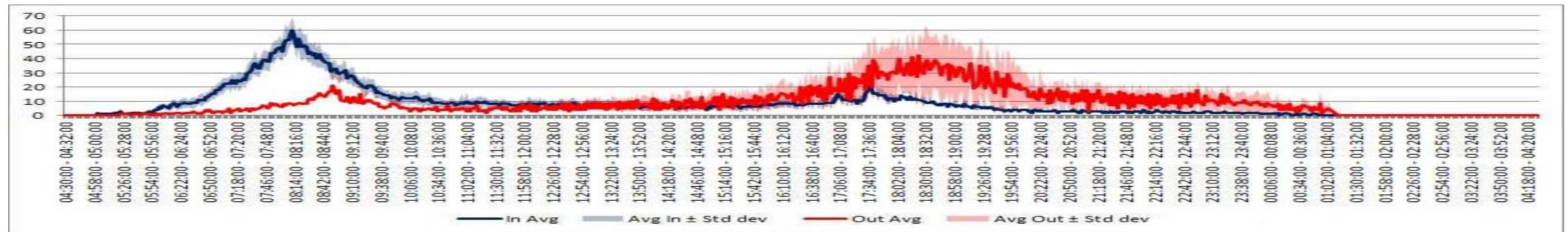


Группа 4: 47,7% взрослых, 41,7% обучающихся, 10,6% пожилых пассажиров

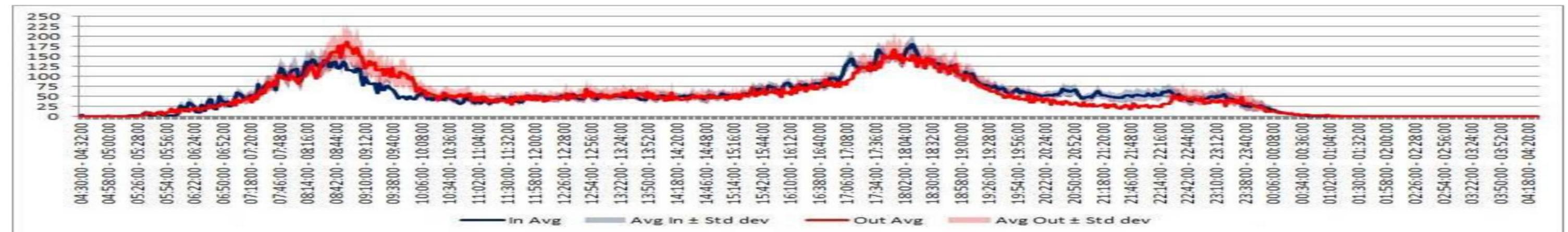
Исследование возможных шаблонов



Бизнес-центр

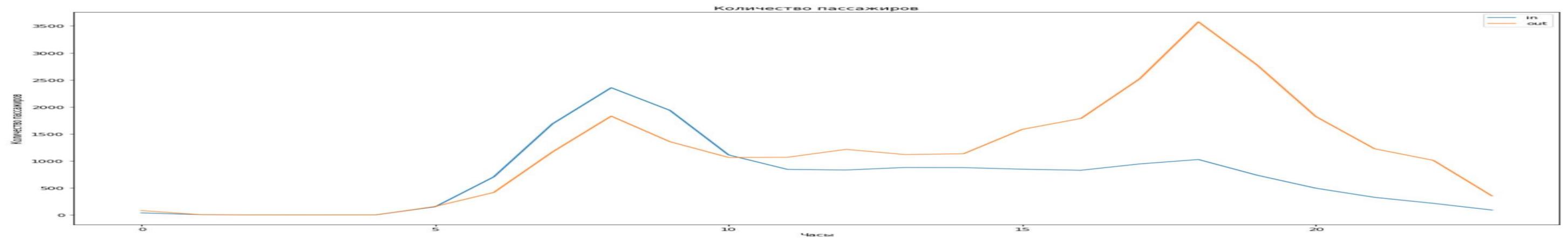


Спальный район

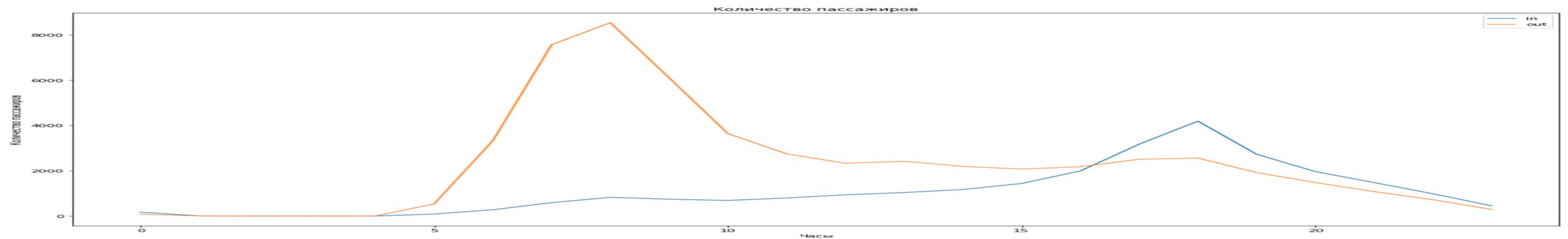


Пересадочный пункт

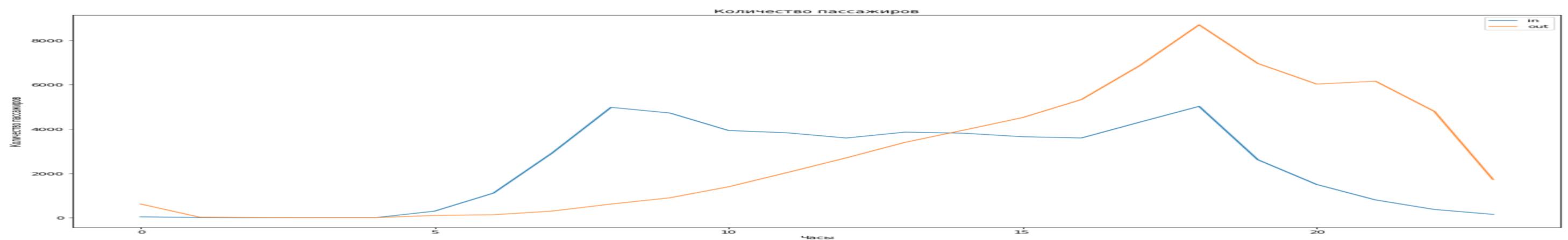
Исследование возможных шаблонов



Бизнес-центр (Выставочная)



Спальный район (Жулебино)



Пересадочный пункт (Охотный ряд)

Временные ряды

Станции могут различаться по типу использования. Количество шаблонов поведения конечно.

Имеем данные входов и выходов пассажиров на каждой станции Московского метрополитена. Считаем, что они полностью описывают модель использования станции.

На каждой станции можем построить временной ряд входов и выходов пассажиров по часам.

Кластеризация временных рядов поможет выявить основные шаблоны использования станций.

Анализ

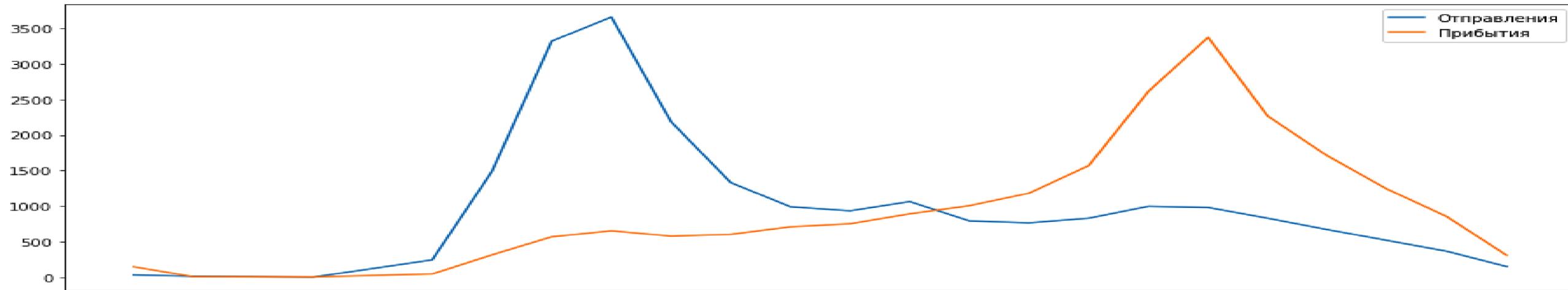
Был реализован алгоритм кластеризации временных рядов k-mean.

С помощью него кластеризовали станции по параметрам входов и выходов. Этот вариант может дать любое количество кластеров, которые будут отличаться только объемами.

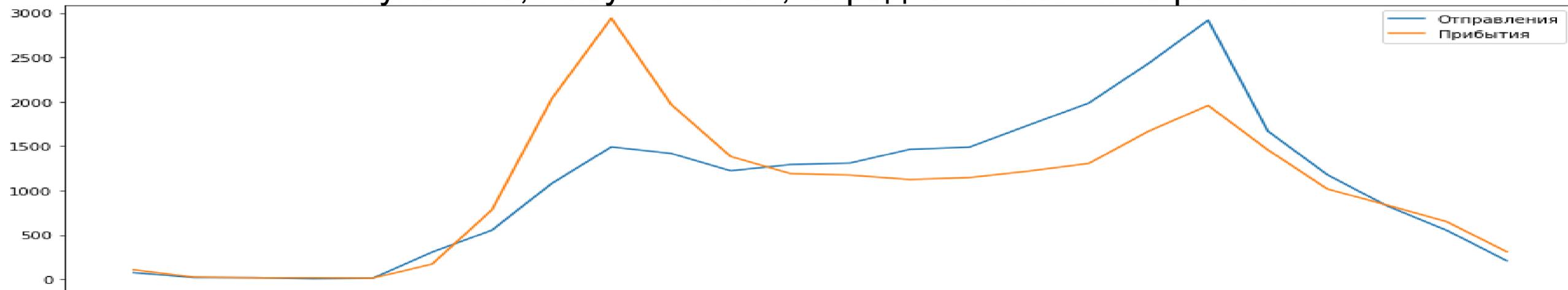
Кластеризовали станции по соотношению входов к выходам.

Получилось 5 основных кластеров. Изучили их наполнение.

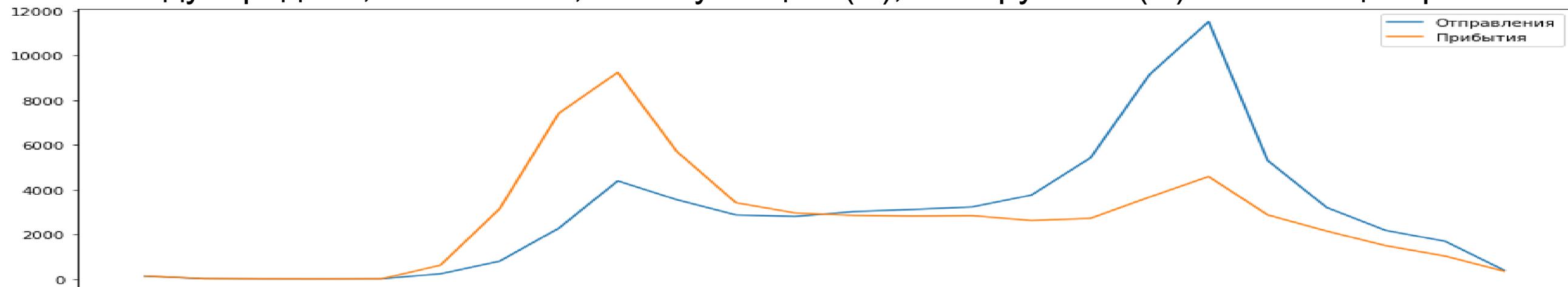
Полученные кластеры



Жулебино, Бабушкинская, Отрадная - спальные районы

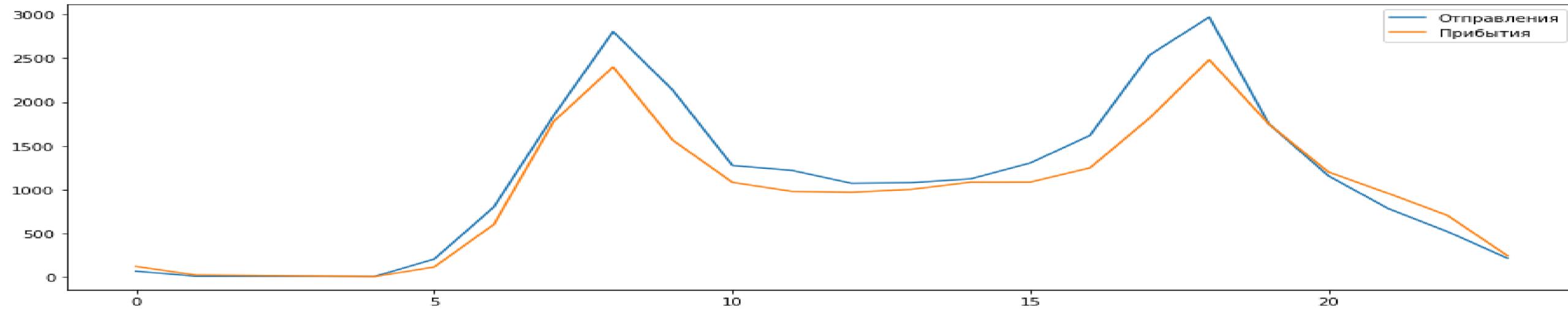


Международная, Войковская, Новокузнецкая(з.), Белорусская (з.) - бизнес-центры

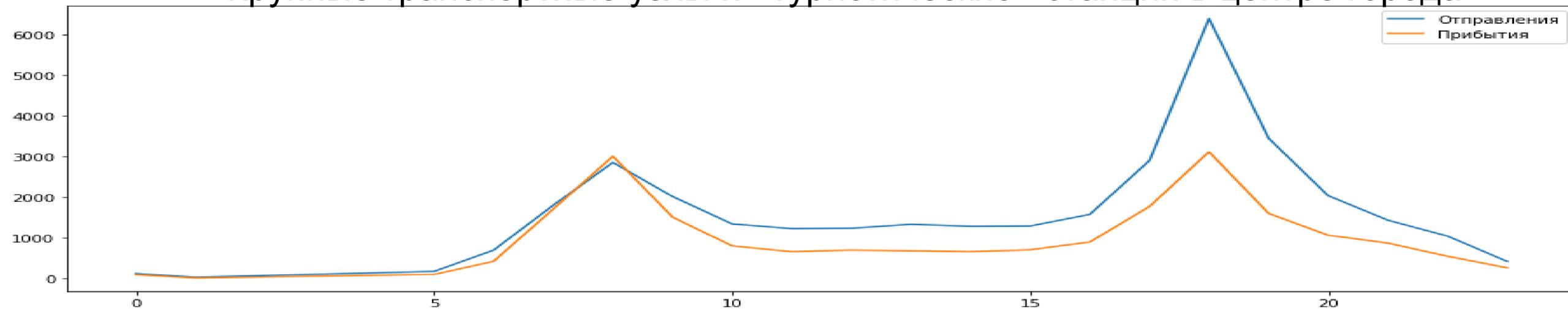


Бауманская, Университет, Каширская, Проспект Вернадского - «университетские» станции

Полученные кластеры



Крупные транспортные узлы и «туристические» станции в центре города



Водный стадион, Теплый стан, Юго-Западная - автовокзалы и окраинные станции

Итоги работы

- На основе анализа литературы и проведенных исследований предложены методы определения шаблонов использования станций.
- Сформирован ряд факторов, которые влияют на соотношение входов и выходов.
- Исследованы и определены основные шаблоны перемещения пассажиров по данным Московского метрополитена. Аномалиями будут являться отклонения от основных шаблонов.
- Работа представлена для ежегодной научной конференции «Ломоносовские чтения-2020».