План

1. Введение (тип как в пропоузале)

- Задачи, стоящие перед 3 поколением, и конкретно решаемая здесь: ребалансировка

- Кратко описанный новый подход к прогнозированию для ребалансировки, предлагаемый этим дипломом. Кратко о используемом свежем датасете.

2. Основная часть

- Положительное влияние велопроката, обуславливающее его популярность

- История поколений городского велопроката. Почему 4 поколение не вытеснит 3, и работать над улучшением 3 всё ещё актуально

Обзор литературы.

- История исследований загруженности велопроката. Кто и что применял.

Статьи про прогнозирование загруженности велопроката. Акцент на том, что всегда в этих статьях раньше юзался другой игрек. Максимум – 0 vs 1 для определения вообще ситуации переполненности. **Составить какую-нибудь большую таблицу, обозревающую методы, результаты, выявленные зависимости, точность?**

- Теоретическое моделирование исходящих и входящих потоков велосипедов как Пуассоновских процессов. Кусочно-линейный/ступенчатый поток. ???

- Возможность, причины и полезность применения этого для ситуаций ненаблюдаемого спроса. Способы оценки полученных алгоритмов в сравнении с наблюдаемыми.

- Кейс стади: Дублин

- Отбор признаков, описание данных

- Теоретическое описание использованных алгоритмов

- Построение моделей

- Обзор результатов, выбор лучшей модели для наблюдаемого спроса, для овердеманда, overall

- Заключение

-

Внедрение автоматизированных систем городского велопроката является распространённой практикой решения проблемы перегруженности улично-дорожной сети в крупных городах по всему миру.

Одной из первостепенных проблем, затрагиваемых в современной практике развития городской среды, является снижение нагрузки на улично-дорожную сеть

#транспортная политика мегаполисов

**Abstract** Bicycle sharing systems are eco-friendly transportation systems that have found wide application in Smart urban environments. Monitoring and analyzing the occupancy levels of the system’s stations is crucial for guaranteeing the quality of the offered service. Advanced data mining solutions tailored to bicycle sharing data analysis are needed to support system managers to react to critical situations (e.g., lack of parked bicycles at a station) that can yield service disruption.

**Введение.**

В XXI веке приоритетной стратегией в транспортной политике многих крупных городов является развитие велоинфраструктуры и её интеграция в улично-дорожную сеть и в систему общественного транспорта, в частности, внедрение автоматизированных систем городского велопроката, или велошеринга.

Велопрокат может как составлять конкуренцию существующей системе общественного транспорта, так и дополнять её как средство преодоления «последней мили». В центральной части города велосипед чаще становится заменой публичному транспорту, в то время как в периферийных районах использование общественного транспорта растёт в ответ на запуск велопроката (Shaheen et al, 2011). Тем не менее опыт многих американских и китайских городов показывает, что транспортная сеть в целом разгружается: внедрение системы велопроката напрямую снижает использование личных автомобилей и такси (Martin, Shaheen, 2014). Так, по оценкам Hamilton, Wichman (2018), открытие городского велопроката позволил сократить дорожный трафик в Вашингтоне на 4%. Wang, Zhou (2017) оценивают эффект от введения городского велопроката методом «разность разностей» на панельных данных за 2005-2014 года по 96 городам США. Полученные авторами статьи результаты говорят о том, что запуск общественного велопроката сокращает загруженность городской сети, причём этот эффект сильнее в крупных городах.

Система показала свою жизнеспособность и положила начало популяризации общественного велопроката: количество общественных прокатов велосипедов выросло с 13 в 2004 году до более 850 в 2014, а к концу 2017 года в мире действует около 1500 велопрокатов (Meddin, DeMaio, 2017).

**Глава 1.**

**История развития велопроката.**



<https://www.citylab.com/city-makers-connections/bike-share/> хороший таймлайн велопрокатов

В истории развития велопрокатов можно выделить четыре основных этапа (Shaheen et al. (2010)). Пилотная программа общественного проката велосипедов, появившаяся в Амстердаме в 1960-е годы, была запущена активистами движения Прово и предполагала возможность бесплатно воспользоваться велосипедом и оставить его в любом месте города, причём велосипеды не имели замка. Такую схему пытались адаптировать и другие города, например, Кембридж в 1993 году, но впоследствии отказывались от неё из-за воровства велосипедов.

В начале 1990-х в Дании появилось второе поколение велопроката, позволявшее бесплатно, но под залог небольшой суммы брать велосипеды на специальных станциях. Внесение залога было необходимо для разблокировки замка на велосипеде. Из-за того, что пользоваться системой можно было анонимно, количество краж велосипедов всё ещё было очень велико.

Решение этой проблемы было предложено в третьем и наиболее популярном в настоящее время поколении городского велопроката. Его отличают деанонимизация пользователей, наличие станций проката и использование информационных технологий на всех этапах работы. Система показала свою жизнеспособность и положила начало популяризации общественного велопроката: количество общественных прокатов велосипедов выросло с 13 в 2004 году до более 850 в 2014, а к концу 2017 года в мире действует около 1500 велопрокатов (Meddin, DeMaio, 2017). Современные велопрокаты третьего поколения оснащены автоматическими велостанциями, работающими по принципу самообслуживания, позволяют отслеживать расположение и загруженность велостанций в реальном времени и оплачивать поездки или абонементы с помощью смартфона или карты.

Велопрокат четвёртого поколения появился в Китае в 2014 году, и несмотря на то, что новое поколение бесстанционного велошеринга набирает популярность в Азии, концепция пока плохо приживается во многих европейских городах, вынуждая успешные азиатские компании сворачивать свои проекты в Европе. Бесстанционный велопрокат не требует больших стартовых вложений в строительство, поддержание велостанций и согласование и оплату их размещения. Кроме того, он даёт пользователям возможность оставить велосипед там, где им удобно, и решает проблему перегруженности велостанций, при которой велосипедисту может потребоваться искать станцию со свободными парковочными местами. Если не рассматривать затраты на обслуживание и ремонт велосипедов, находящихся в эксплуатации, то бесстанционный велошеринг дешевле запускать и обслуживать, поэтому такие системы не нуждаются в государственном субсидировании.

Тем не менее бесстанционные велопрокаты не освобождены от затрат на перераспределение велосипедов. Так, регулярные жалобы на хаотичную парковку велосипедов, затрудняющую перемещение пешеходов в Далласе, повлекли вмешательство муниципальных властей: городской управляющий Далласа Т. Броднакс в обращении к пяти действующим в городе операторам бесстанционного велопроката, опубликованном 18 января 2018 года, потребовал от них освобождать пешеходные зоны от скоплений велосипедов. Кроме того, бесстанционный велопрокат более подвержен вандализму. Это связано с тем, что в такой модели велошеринга каждый велосипед оснащён GPS-трекером и электронным замком, а значит, может представлять больший интерес для кражи. В то время как станции велопроката обычно расположены в людных местах и находятся под видеонаблюдением, велосипеды в бесстанционном прокате часто оставляют во дворах и других местах, которые не находятся под видеонаблюдением, что повышает вероятность вандализма, кражи деталей или самого велосипеда. Например, в феврале 2018 года гонконгский стартап Gobee.bike прекратил свою работу во Франции после того, как представители компании сообщили о краже или выводе из строя 60% их велосипедов за первые 4 месяца работы на французском рынке. Ранее в этом году Gobee.bike покинули Милан, Рим, Брюссель и другие города по этой же причине.

Вандализм и кражи в велопрокатах третьего поколения тоже неизбежны. Ущерб сильно варьируется между городами и зависит как от того, насколько развита система интегрированной безопасности в конкретной городской среде, так и от мер оператора велопроката по предупреждению вандализма. Например, в 2012 году было испорчено или украдено более 37% велосипедов парижского станционного проката Vélib. Эта доля значительно ниже ущерба, нанесённого компании Gobee.bike. Сообщается, что инциденты в Париже кластеризуются в бедных районах, где отсутствует уличное видеонаблюдение. В то же время представители велопроката в Белфасте сообщают о том, что в среднем 15% их велосипедов ломают или воруют каждый год со времени их запуска в 2015 году. Представители оператора лондонского транспорта Transport for London уверены, что решение сделать обязательную привязку банковской карты к аренде велосипеда и отсутствие дополнительного замка на велосипеде позволили значительно сократить число краж. Значительно меньшему уровню краж в Лондоне способствует и большее число камер видеонаблюдения.

Описанные трудности, с которым бесстанционный велопрокат сталкивается в настоящее время, позволяют предположить, что он не сможет быстро вытеснить станционный прокат, и в ближайшем будущем два поколения проката будут сосуществовать. Таким образом, исследуемая в этой работе оптимизация работы станционного проката не теряет актуальности.

**Оптимизация функционирования станционного велопроката.**

Современный станционный велопрокат позволяет брать и возвращать велосипеды на любой из расположенных в городе автоматизированных велостанций. Поскольку количество слотов для парковки велосипеда на станции ограничено, предложение велосипедов и пустых парковочных мест на станции не всегда может удовлетворить спросу на них. Если станция пустая, то есть спрос на велосипеды превышает их предложение, то пользователь не совершит поездку. Если станция переполнена, то велосипедист, который хочет припарковать велосипед, вынужден будет искать другую станцию, чтобы припарковать велосипед. Это может стоить ему дополнительных денег из-за большей длительности поездки, а также вынуждает его выбрать неоптимальную для него точку окончания поездки.

Ребалансировка необходима и для удержания регулярных клиентов велопроката. Значительная часть клиентов покупает полугодовой или годовой абонемент (цифры и источники?) и использует велосипед для преодоления «последней мили» по дороге на работу или домой. Если станция подвержена

В мировой практике велопрокатов поддержание оптимального количества велосипедов и парковочных мест на станции осуществляется различными способами. В зависимости от того, как быстро наступает эффект от тех или иных мер, исследователи разделяют их на три основные группы по горизонту планирования (Vogel et al., 2011):  
• В краткосрочном периоде – оперативная ребалансировка станций: перевоз велосипедов вечером и в час пик с тех станций, куда повышен приток велосипедистов, на те, с которых повышен отток велосипедистов). Так, в 2009 году стоимость перевоза одного велосипеда французского оператора Vélib оценивалась в 3 доллара (DeMaio, 2009).  
• В долгосрочном периоде – увеличение вмещаемости станций или открытие новых поблизости. В некоторых случаях открытие новых станций может увеличить встречный трафик.   
• За счёт искусственного изменения перемещения велосипедистов: например, система Citi Bike Angels в Нью-Йорке стимулирует пользователей сесть на велосипед на станции A, на которой слишком много велосипедов, и доехать до станции B, на которой слишком мало велосипедов. За это пользователям даются бонусы, например, дополнительная поездка или скидка на поездку. Парижский прокат Vélib стимулировал пользователей возвращать велосипеды на станции, расположенные на возвышении, давая им 15 дополнительных минут езды (DeMaio, 2009).

Наиболее популярной мерой является регулярное распределение велосипедов. Необходимость произвести наиболее эффективную с точки зрения неупущенных поездок при минимальных затратах рождает ряд оптимизационных проблем. Проблема поиска оптимального или близкого к оптимальному маршрута и числа грузовых транспортных средств часто затрагивается в академической литературе (источники).

Данная работа фокусируется на решении предваряющей поиск оптимального маршрута задачи ­– предсказании спроса на велосипеды и слоты на каждой станции. Очевидно, что более точная оценка количества велосипедов и парковочных мест, которых будет недоставать на каждой станции в определённое время в будущем, позволит построить более эффективный с точки зрения затрат или другой оптимизируемой в модели величины маршрут перераспределения велосипедов. Решение этой задачи осложняется конечным количеством парковочных мест на велостанциях, вследствие чего реальный спрос на велосипеды и парковочные места не всегда наблюдаем. Изучив корпус академических публикаций, в которых строятся предсказательные модели для станционных велопрокатов, авторы данной работы не обнаружили прямого подхода к решению этой проблемы; предлагаемый в работе подход позволит впервые оценить реальные объёмы спроса на отдельных станциях велопроката на основе не использованных ранее данных дублинского городского проката dublinbikes.

**Обзор литературы.**

**Прогнозирование загруженности станций велопроката.**

**Тут написать в конце этой подглавы про то, что уже есть системы, прогнозирующие количество великов через х минут, но они игнорируют ненаблюдаемый спрос. И ещё что есть предсказания бинарной переменной будет загружено – не будет загружено, но в этом не много смысла. Найти к этому статьи мб.**

Деанонимизация пользователей и внедрение информационных технологий в системы велопроката к концу 2000-х годов позволили операторам проката собирать информацию о клиентах и совершаемых ими поездках, что привлекло внимание к байкшерингу в академической литературе. На настоящий момент опубликовано значительное количество статей, посвящённых предсказанию загруженности велостанций. Публикация под авторством Froehlich, Neumann, Oliver (2009) одной из первых обращает внимание на потенциал использования цифровой информации для выявления закономерностей в использовании велостанций на примере велопроката Bicing, запущенного в Барселоне в 2007 году. Авторы применяют иерархическую кластеризацию для группировки станций на основе их загруженности в течение дня. Пространственный анализ результатов показывает, что паттерны использования велопроката различны для центральной части города и его периферии. Результаты применённой в этой статье для предсказания числа велосипедов на станциях Байесовская сеть превосходят по точности прогнозирование историческим средним значением.

Значительная часть ранних исследований, опубликованных после этой пилотной статьи, использует модели временных рядов. Bognat et al. (2009) моделирует отток велосипедистов со станции как Bognat, Abry, Flandrin (2011) conducted clusterization similar to an already mentioned work of Froehlich et al., but using a rich dataset published by Vélo’v, bike-sharing service in Lyon, France.

С появлением большего количества открытых данных велопрокатов набирают популярность эконометрическое моделирование и методы машинного обучения для предсказания спроса на велошеринг. Так, Mauer (2011) отбирает ряд экономических и демографических характеристик местоположения разных велостанций, строит логлинейную регрессионную модель и делает выводы, что доход, плотность числа рабочих мест, количество автомобилей на число жителей, вместительность велостанции и другие факторы влияют на количество использований велопроката. Rixey (2013) использует многомерную линейную регрессию для объяснения того, как средний уровень образования, доход и число станций в районах влияет на ежемесячный объём спроса на велосипеды. Качественно новый подход одним из первых изложен в A Faghih-Imani (2014), где авторами построены отдельные для каждой станции смешанные линейные модели и использованы время, погодные признаки и признаки, характеризующие использование окружающей станцию территории. В ряде статей показано значимое влияние наличия и концентрации развлекательных учреждений и популярных среди туристов мест на частоту использования станции (Hampshire, 2011).

Значительный массив публикаций затрагивает проблему избытка спроса, и для моделирования таких ситуаций предлагаются как вероятностные модели, так и использование методов машинного обучения. Тем не менее эти работы фокусируются на предсказании факта избытка спроса, а не на величине этого избытка. Таким образом, преимущественно используются бинарные модели, которые выявляют ситуацию, когда на станции нет свободных парковочных мест, против ситуации, когда на станции есть свободные места. Аналогичная модель строится для того, чтобы научиться предсказывать, когда на станции не останется велосипедов. Так, публикация Chen et al. (2016) предлагает динамическую кластеризационную модель для предсказания избытка спроса и сравнивает её результаты с базовыми моделями временных рядов и машинного обучения: ARIMA, Байесовская сеть, искусственные нейросети.

Ряд преимущественно теоретических статей использует вероятностный подход к моделированию частоты притока и оттока велосипедов со станции. Так, Feng et al. (2016) доказывают, что независимый Пуассоновский процесс может быть использован для моделирования притока и оттока. Исследование Gast, Massonnet et al. (2015) подтверждает на исторических данных, что эти два параметра достаточно точно описываются неоднородными во времени Пуассоновскими процессами.

Список литературы:

Meddin, R. and DeMaio, P. The bike-sharing world map. *URL http://*www.bikesharingmap.com

*[Accessed April 12, 2018.]*.

Shaheen, S., Zhang, H., Martin, E. and Guzman, S., 2011. China's Hangzhou public bicycle: understanding early adoption and behavioral response to bikesharing. *Transportation Research Record: Journal of the Transportation Research Board*, (2247), pp.33-41.

Shaheen, S., Guzman, S. and Zhang, H., 2010. Bikesharing in Europe, the Americas, and Asia: past, present, and future. *Transportation Research Record: Journal of the Transportation Research Board*, (2143), pp.159-167.

Martin, E.W. and Shaheen, S.A., 2014. Evaluating public transit modal shift dynamics in response to bikesharing: a tale of two US cities. *Journal of Transport Geography*, *41*, pp.315-324.

DeMaio, P., 2009. Bike-sharing: History, impacts, models of provision, and future. *Journal of public transportation*, *12*(4), p.3.

Hamilton, T.L. and Wichman, C.J., 2018. Bicycle infrastructure and traffic congestion: evidence from DC's capital Bikeshare. *Journal of Environmental Economics and Management*, *87*, pp.72-93.

Broadnax T.C. Memorandum: Letter to Dockless Bike Share Operators, January 18, 2018. URL http://dallascityhall.com/government/Council%20Meeting%20Documents/msis\_3\_bike-share-program-update\_combined\_022618.pdf