

Федеральное государственное автономное образовательное
учреждение высшего образования
«Национальный исследовательский университет
«Высшая школа экономики»

на правах рукописи

Курмуков Анвар Илдарович

Иерархическая структура коннектомов
ГОЛОВНОГО МОЗГА

РЕЗЮМЕ

диссертации на соискание учёной степени
кандидата компьютерных наук

Москва - 2021

Диссертационная работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Национальный исследовательский университет «Высшая школа экономики».

Научный руководитель: Жуков Леонид Евгеньевич, PhD, профессор, Национальный исследовательский университет «Высшая школа экономики»

Тема диссертации

Актуальность исследования

Научная проблема, на решение которой направлено диссертационное исследование, состоит в практически полном отсутствии клинически информативных нейровизуализационных биомаркеров заболеваний мозга (как нейродегенеративных, так и психиатрических), которые позволили бы своевременно идентифицировать патологический процесс, построить прогноз индивидуальной траектории течения заболевания и оценить результативность терапии, в том числе фармакологического воздействия. Такие биомаркеры остро необходимы для решения задач персонализированной диагностики, эффективного лечения и предупреждения развития заболеваний мозга.

Обращение к данным неинвазивной нейровизуализации при поиске таких биомаркеров выглядит вполне естественным [1]. Задача, однако, существенно осложняется системной природой заболеваний мозга, в первую очередь психиатрических, когда в большинстве случаев невозможно четко локализовать патологический субстрат и однозначно связать конкретные изменения с течением болезни. По этой причине современные исследования стремятся анализировать патологические паттерны на уровне сетевых структур мозга, репрезентирующих не только отдельные регионы мозга, но и их структурные и функциональные связи [14]. Построение биомаркеров заболеваний мозга на этом уровне требует работы на стыке обработки изображений (анализа данных нейровизуализации), современного предсказательного моделирования и анализа графов и сетевых структур.

В рамках диссертационного исследования были разработаны алгоритмы анализа модулярной структуры сетей головного мозга [10],

которые могут выступать в качестве потенциальных биомаркеров патологических процессов. Предсказательные модели на основе разработанных алгоритмов в перспективе позволят строить персонализированный прогноз развития заболеваний мозга с помощью автоматизированного анализа данных неинвазивной нейровизуализации.

Актуальность настоящего исследования обусловлена важностью проблемы раннего прогнозирования нейродегенеративных заболеваний, таких как болезнь Альцгеймера и болезнь Паркинсона. В соответствии с мета анализом проведенным международной организацией Alzheimer's Disease International (ADI) по состоянию на 2015 год число людей живущих с Болезнью Альцгеймера достигает 46.8 миллионов. Каждые 20 лет это число увеличивается в 2 раза, таким образом ожидается что к 2050 году оно достигнет 131.5 миллиона человек [35]. Согласно тому же докладу общемировые расходы связанные с деменцией достигают 818 миллиардов долларов (США) только в 2015 году, в 2018 году эти расходы составили уже 1 триллион долларов (США) [36].

В последнее время считается что использование различных биомаркеров полученных из данных нейровизуализации, таких как МРТ и ПЭТ (позитронно эмиссионная томография) должно играть важную роль в диагностике деменции в целом, и болезни Альцгеймера в частности [25]. Так например было показано, что использование такого простого биомаркера как атрофия мозга (brain atrophy), измеренного по данным структурного МРТ, служит очень точным предиктором для ранней диагностики заболевания Альцгеймера [29]. Так например большое количество работ ставят перед собой следующую задачу: предсказать произойдет ли переход от более ранней стадии заболевания к более поздней, на некотором временном отрезке. Такая постановка позволяет

использовать современные методы машинного обучения и решать, так называемую, задачу классификации (в случае если временной промежуток зафиксирован) [24], [40] или регрессии (если не зафиксирован). Для решения этой задачи необходимо наличие достаточного количества размеченных лонгитюдных данных. Разметка должна содержать точную информацию о стадии заболевания для каждого временного наблюдения. Отметим что такая постановка отличается от простой задачи классификации патологии от нормы (например Альцгеймер - не Альцгеймер) тем, что предсказание делается о состоянии пациента в будущем, а не в настоящий момент. Таким образом решение этой задачи может быть целесообразно не только с научной, но и с клинической точки зрения.

Исследование развития (disease developing, disease progression) нейродегенеративных заболеваний играет важнейшую роль в вопросах ранней диагностики и стратегии лечения [15], [32].

Цель и задачи исследования

Задача диссертационного исследования состоит в разработке алгоритмов интеллектуального анализа данных неинвазивной нейровизуализации различных модальностей, позволяющих на основе реконструкции анатомических и функциональных сетевых структур мозга [41] оценивать их иерархическую модулярную организацию [3] и строить предсказательные модели в области диагностики и прогноза течения нейродегенеративных и психиатрических заболеваний. С практической точки зрения, речь идет о построении клинически информативных биомаркеров для решения задач персонализированной диагностики, планирования терапии и предупреждения развития заболеваний мозга. С научной точки зрения, были получены новые результаты, развивающие

наиболее перспективные подходы к предсказательному моделированию на мультимодальных данных нейровизуализации.

Следуя общепринятому определению, под биомаркером будем понимать объективно измеримую характеристику, которая может рассматриваться в качестве индикатора нормальных биологических процессов, патогенетических процессов (в данном случае - заболеваний мозга), или фармакологических ответов на терапевтические вмешательства [31]. Внимание данной работы именно к сетевым структурам мозга обусловлено системным характером нарушений структуры и функционирования мозга, характерным для многих патологий, в первую очередь в психиатрии. По этой причине в современных исследованиях всё больший интерес вызывает коннектомный подход [5], в рамках которого мозг человека анализируется как целостная сеть с определенной структурой. Коннектом - это граф, в котором вершины соответствуют регионам мозга, а связи указывают либо на наличие анатомических трактов между соответствующими зонами, либо на их тенденцию совместно активироваться в процессе функционирования. На Рис. 1 представлен пример коннектома человека.

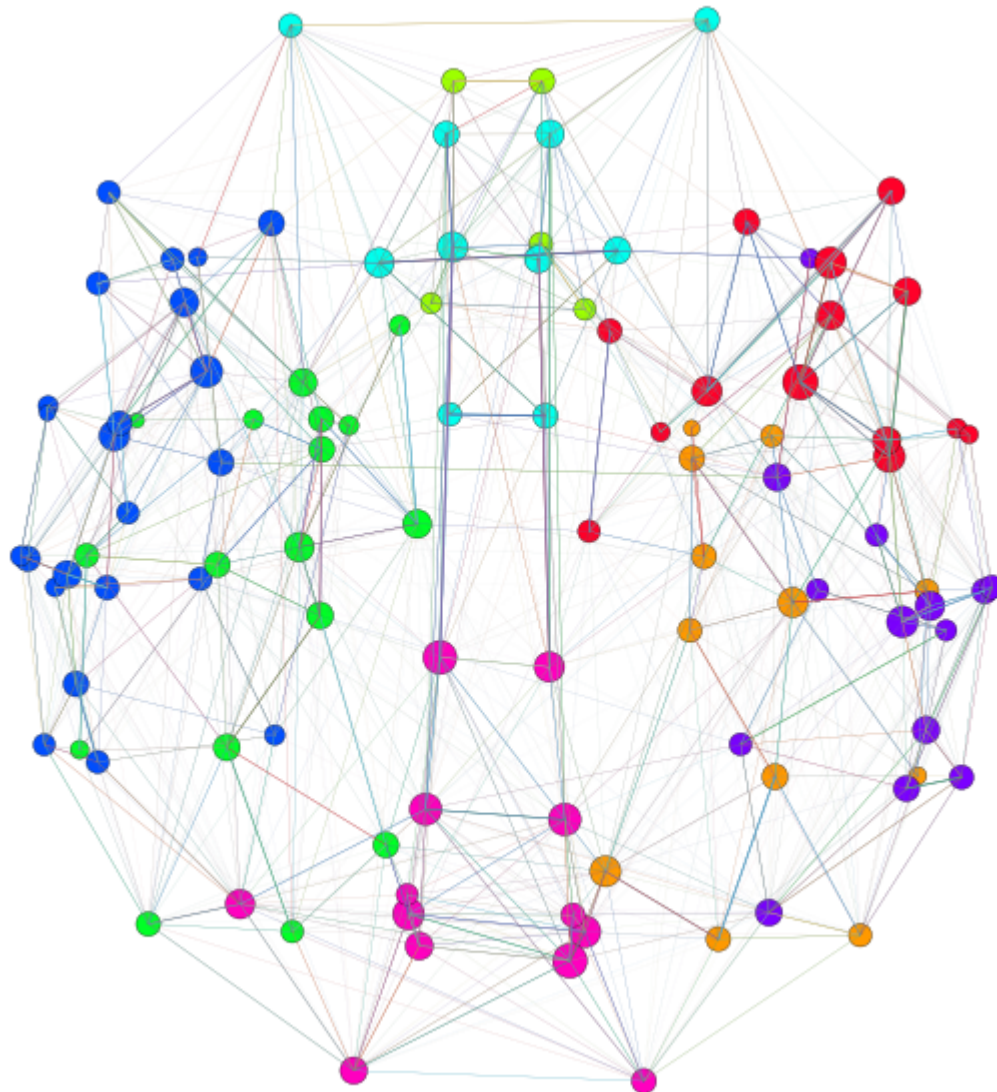


Рисунок 1. Граф коннектома мозга человека. В качестве вершин используются регионы атласа Десикан. Цвета вершин соответствуют сетевым сообществам, полученным с помощью алгоритма Louvain modularity.

Ожидается, что применение современных методов предсказательного моделирования к такого рода данным, реконструированным на основе нейровизуализации, позволит повысить точность диагностических и прогностических моделей заболеваний, связанных с патологиями мозга.

С математической точки зрения задача предсказательного моделирования на сетевых структурах мозга выглядит следующим образом. Мы рассматриваем выборку данных как набор пар, состоящих из входного графа, репрезентирующего сеть головного мозга, восстановленную на основе данных нейровизуализации, и выходной категориальной переменной, принимающей малое число различных значений. Например, целевая переменная может принимать значения от 0 до 2, кодируя три состояния: испытуемый здоров (0), продромальная фаза шизофрении (1) и собственно шизофрения (2). Задача сводится к классификации, то есть построению математической модели, которая для каждого нового примера входного графа будет выдавать наиболее вероятное значение целевой переменной. Для оценки качества модели могут использоваться подходящие для конкретной прикладной задачи стандартные метрики качества.

Заметим, что специфика предметной области определяет ряд ключевых особенностей анализируемых входных графов. Речь идет об относительно небольших (порядка сотни вершин) неориентированных графах с единственной компонентой связности. Каждая вершина в таком графе помечена уникальным образом в соответствии с конкретной зоной мозга, которую она представляет, и набор таких меток одинаков для всех графов в рамках решаемой прикладной задачи. Вершины имеют набор атрибутов, задающих их координаты в трехмерном пространстве и физические параметры соответствующих зон (толщины и площади поверхности). Граф является взвешенным, причем в случае мультимодальных данных для сетевой структуры одного мозга может существовать два набора весов, один из которых включает оценки связности регионов мозга посредством анатомических трактов, а другой указывает на интенсивность совместной активации зон по данным

функциональной диагностики. Таким образом, стандартные методы классификации не могут быть использованы для решения поставленной задачи в силу специфики входных данных, имеющих описанную структуру. При построении предсказательных моделей на таких данных важно наиболее полно сохранить информацию о ключевых особенностях структуры сетей головного мозга, из которых наиболее важная для нас - их модулярная организация [26]. Модулярная организация сети означает, что ее вершины имеют тенденцию объединяться в модули, или сообщества, с очень плотными внутригрупповыми и разреженными межгрупповыми связями. В работе [2] впервые продемонстрировано, что именно модулярная организация сетей мозга меняется в случае системных психиатрических заболеваний (анализировались топологические особенности функциональных сетей у больных с диагнозом ранняя детская шизофрения).

Результаты, полученные в рамках работы, позволяют говорить о том, что такой подход к построению предсказательных моделей на основе данных нейровизуализации может быть продуктивен при решении задачи классификации психиатрических заболеваний (в частности, расстройств аутистического спектра [20]) и нейродегенеративных заболеваний (группы риска болезни Альцгеймера [21]).

Это позволяет предположить, что именно структура сообществ регионов мозга в коннектоме в полной мере отражает топологическую организацию исходной сети и, таким образом, содержит достаточно информации для построения предсказательных моделей с целью различения клинически значимых диагностических групп. Каждый коннектом может быть представлен вектором длины, равной количеству регионов мозга (вершин графа), в котором каждое значение кодирует

принадлежность соответствующего региона мозга к одному из сообществ в оптимальном в том или ином смысле разбиении исходной сети на непересекающиеся сообщества. Такое представление позволяет оценивать расстояния между исходными графами как расстояния между соответствующими разбиениями и, на следующем шаге, строить алгоритмы классификации нормальных и патологических сетевых структур с использованием ядерных методов, принимающих на вход матрицу таких попарных расстояний между исходными объектами.

В рамках диссертационного исследования были поставлены и решены следующие задачи:

1. Были предложены методы классификации графов коннектомов, позволяющий учитывать пересекающуюся структуру сообществ и их иерархическую организацию.
2. Был предложен способ построения атласа головного мозга, на основе данных диффузионного МРТ с использованием непрерывной модели коннектома. Атлас был получен на основе data-driven подхода, а не на основе анатомической или функциональной структуры головного мозга, однако в качестве регионов успешно выделяет анатомические структуры и при этом может быть скорректирован для работы с конкретным набором данных.
3. Была предложена модель прогрессии заболевания на основе мультимодальных данных нейровизуализации. Предложенная вероятностная модель учитывает как морфометрические характеристики (такие как толщина серого вещества головного мозга), так и структурные (связи между отдельными регионами).

Важно подчеркнуть, что разрабатываемые в рамках диссертационного исследования алгоритмы оценивались и сравнивались с

точки зрения их информативности и полезности в контексте предсказательного моделирования заболеваний, связанных с патологией мозга. Модулярная организация сетевых структур, реконструированных на основе данных нейровизуализации, рассматривалась в качестве потенциального биомаркера, информативного с точки зрения ранней диагностики заболевания и персонализированного прогноза его течения.

Основные результаты

В Главе 1 изложена математическая модель структурного коннектома и детали восстановления таких графов на основе данных магнитно-резонансной томографии (МРТ). В Главе 2 описаны список используемые методы поиска сообществ вершин, методы сравнения отдельных кластеризаций и способы их усреднения. В Главе 3 приводятся существующие, и предлагаются некоторые новые способы построения ядер, на основе различных характеристик графов. Мы демонстрируем эффективность описанных ядер для решения задач бинарной классификации различных фенотипов, патологии и нормы. В Главе 4 излагается расширение модели ЕВМ, которое позволяет использовать данные о связности регионов (коннектомы) для генерации априорного распределения для порядка дегенерации регионов мозга. Наконец, Глава 5 посвящена исследованию иерархической организации графов коннектомов. Демонстрируется, что группы вершин графов коннектомов, образующие плотные сообщества часто оказываются анатомически близкими. Наконец в Главе 6 предлагается способ построения анатомического атласа, основанного на структурной (физической, посредством аксональных соединений) связности регионов, который по целому ряду показателей превосходит существующие анатомические атласы.

Диссертация состоит из введения, 6 глав, заключения, списка литературы, иллюстраций и таблиц. Общий объем диссертации составляет 177 страниц.

Основные результаты исследования и положения, выносимые на защиту.

В рамках диссертационного исследования было продемонстрировано что:

1. Предложен метод решения задачи классификации объектов представимых в виде графов определенных на одном множестве вершин (с отличающимся множеством ребер). Метод основан на сравнении пересекающихся и непересекающихся кластерных структур таких графов, что позволяет эффективно снизить их признаковое описание.
2. Предложена модель прогрессии заболевания на основе мультимодальных данных нейровизуализации. Предложенная вероятностная модель учитывает как морфометрические характеристики (такие как толщина серого вещества головного мозга), так и структурные (связи между отдельными регионами).
3. Предложен метод построения анатомического атласа коры головного мозга на основе данных структурной связности с использованием модели непрерывного коннектома.
4. Изучены и описаны иерархические свойства графов коннектомов выводимые на основе данных. Продемонстрировано что в графах коннектомах сообщества вершин часто образуют анатомически близкие регионы.

Апробация результатов исследования

Публикации повышенного уровня:

1. Kurmukov A., Mussabaeva A., Denisova Y., Moyer D., Neda J., Thompson P. M., Gutman B. A. Optimizing connectivity-driven brain parcellation using ensemble clustering // *Brain Connectivity*. 2020. Vol. 10. No. 4. P. 183-194.
2. Kurmukov, A., Musabaeva, A., Denisova, Y., Moyer, D. and Gutman, B., 2018, September. Connectivity-driven brain parcellation via consensus clustering. In *International Workshop on Connectomics in Neuroimaging* (pp. 117-126). Springer, Cham.
3. Kurmukov A., Zhao Y., Mussabaeva A., Gutman B. Constraining Disease Progression Models Using Subject Specific Connectivity Priors, in: *Lecture Notes in Computer Science book series Issue 11848*. Springer, 2019. P. 106-116.
4. Kurmukov, A., Ananyeva, M., Dodonova, Y., Gutman, B., Faskowitz, J., Jahanshad, N., Thompson, P. and Zhukov, L., 2017. Classifying phenotypes based on the community structure of human brain networks. In *Graphs in Biomedical Image Analysis, Computational Anatomy and Imaging Genetics* (pp. 3-11). Springer, Cham.

Публикации стандартного уровня:

1. Kurmukov A., Dodonova Y., Burova M., Mussabayeva A., Petrov D., Faskowitz J., Zhukov L. E. Topological modules of human brain networks are anatomically embedded: evidence from modularity analysis at multiple scales, in: *Computational Aspects and Applications in Large-Scale Networks*. Springer Proceedings in Mathematics and Statistics Vol. 247. Springer, 2018., P. 299-308.
2. Kurmukov A., Dodonova Y., Zhukov L. E. Machine learning application to human brain network studies: a kernel approach, in: *Models, Algorithms, and Technologies for Network Analysis*. Springer Proceedings in Mathematics and Statistics / Ed. by V. A. Kalyagin, A. I. Nikolaev, P. M. Pardalos, O. Prokopyev. Vol. 197. Springer, 2017. doi P. 229-249.
3. Kurmukov A., Dodonova Y., Zhukov Classification of normal and pathological brain networks based on similarity in graph partitions.

IEEE 16th International Conference on Data Mining Workshops,
ICDMW. Tom 0, Pages 107 – 112, July 2016.

Содержание работы

Задача: *”Классификация нормы и патологии на основе кластерной структуры графов головного мозга человека.”*

Предполагается, что внутри коннектомов возможно информативным образом выделить сообщества вершин, представляющих регионы мозга. Информативность означает для нас возможность определить на основе получаемых разбиений расстояния между графами таким образом, чтобы близкими оказались ожидаемо похожие графы – например, принадлежащие людям с конкретным нейродегенеративным заболеванием.

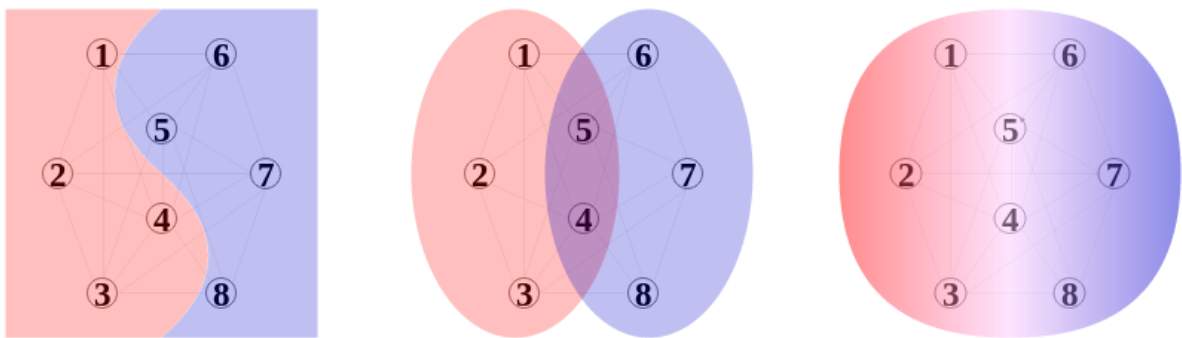


Рисунок 2. Различные виды кластеризации (вершин графа). Слева направо: непересекающиеся кластеры, пересекающиеся кластеры, нечеткие кластеры.

На основе этой идеи был решен ряд задач различения пациентов с различными заболеваниями, такими как: расстройства аутистического спектра, болезнь паркинсона и болезнь альцгеймера и людей без патологии на основе графов структурных связей головного мозга (коннектомов). Для этого были предложены различные модели использующие информацию о различиях в разбиениях графов на подграфы, характерные для коннектомов групп нормы и патологии. Предложенные подходы различаются тем, допускается ли одновременная принадлежность вершины к нескольким сообществам и возможна ли различная степень такой принадлежности. В случае если рассматривается разбиение графов

коннектомов на непересекающиеся сообщества, эта информация может быть закодирована вектором длины равной числу вершин графа. Каждый элемент такого вектора является индикатором принадлежности вершины к сообществу. В противном случае если рассматривается разбиение графов на пересекающиеся сообщества [37], оно будет кодироваться уже не вектором, а матрицей, в которой конкретное значение соответствует вероятности принадлежности региона мозга к одному из сообществ, или набором матриц, если речь идет о разных иерархических уровнях анализа. На Рис. 2 представлено схематическое изображение разбиения вершин графа на сообщества; на пересекающиеся кластеры; на нечеткие кластеры. При этом возникает задача корректного оценивания расстояний между такими объектами для последующего построения алгоритмов предсказательного моделирования на них.

Задача: *“Иерархическая структура графов коннектомов”*

Сети связей человеческого мозга демонстрируют модулярную организацию: кортикальные регионы, как правило, образуют плотно связанные кластеры с относительно небольшим количеством межкластерных связей.

Однако мало что известно о том, является ли модулярная структура мозговых сетей надежной с точки зрения воспроизводимости, и, самое главное, в какой степени эти кластеры вершин связаны с анатомией. Для решения этих вопросов мы используем данные МРТ одних и тех же людей, отсканированных с интервалом в несколько недель (лонгитюдные данные), и реконструируем структурные сети мозга в нескольких масштабах, разделяем их на сообщества и оцениваем сходство полученных кластеризацией. Мы проверяем устойчивость модулярной структуры двумя способами, во-первых, сравниваем кластерные структуры графов, полученных от одного и того же человека и разных людей, во-вторых, сравниваем получаемые графовые кластеры с анатомически близкими регионами. Наши результаты показывают, что модулярная структура сетей мозга хорошо воспроизводится в рамках повторных экспериментов (англ., *test-retest reliability*). Кроме этого, полученные результаты подтверждают теоретически обоснованную гипотезу о том, что области мозга, соседствующие в анатомическом пространстве, также, как правило, принадлежат к одним и тем же сетевым кластерам вершин.

Для сетей с высоким разрешением мы сравнили два подхода к их разбиению их вершин на сообщества. Первый подход основан исключительно на структуре топологических связей графа. Второй подход в первую очередь основан на анатомии, потому что вершины графа высокого разрешения были помещены в сообщества, на основе того, что они были анатомически соседними и принадлежали к одной и той же области коры в анатомическом атласе (и, следовательно, к одному и тому

же родительскому узлу в сети с низким разрешением). Мы продемонстрировали, что модулярность этих последних раскрасок вершин все еще была очень высокой и лишь немного ниже, чем модулярность, оцененная по отношению к топологически оптимальным разбиениям.

Кроме того, мы продемонстрировали, что сходство между топологически оптимальными и анатомическими раскрасками оказалось очень высоким. Топологические модули в значительной степени напоминали анатомическую группировку соседних областей коры и иерархически встраивались в структуру анатомических сообществ. Используя многомасштабный анализ (графов в разных разрешениях) и алгоритмы разбиения сетей на сообщества, мы нашли новые доказательства в поддержку теоретически обоснованной гипотезы о том, что области мозга, соседствующие в анатомическом пространстве, также, как правило, принадлежат к одним и тем же иерархически вложенным топологическим модулям

Рис. 3.

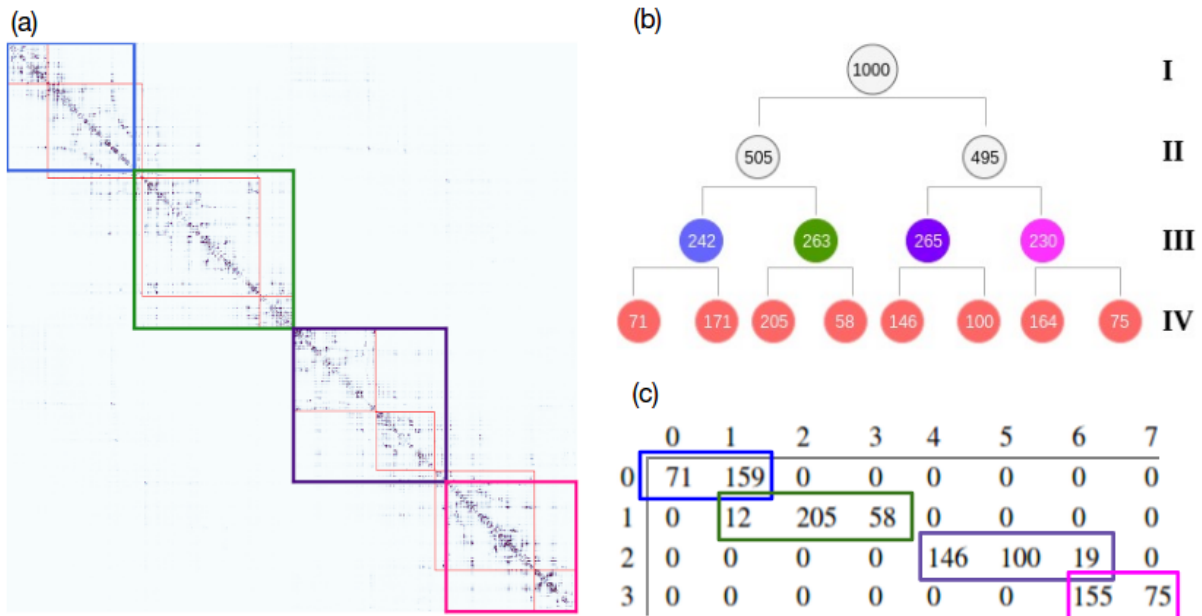


Рисунок 3. Три способа продемонстрировать иерархическую вложенность кластеров сетей высокого разрешения в сообщества образованные в сетях низкого разрешения. (a) Матрица смежности сети

высокого разрешения с восемью топологическими модулями с наилучшим разбиением на сообщества, (светло-красные квадраты), и четырьмя анатомическими сообществами (синий, зеленый, фиолетовый и розовый квадраты); сообщества первого уровня почти строго встроены в сообщества второго. (b) Иерархическое дерево с 4 уровнями: I - весь 1000-вершинный коннектор, II - два полушария, III - четыре кластера, унаследованные от анатомически отображенного разбиения сети низкого разрешения, IV-восемь кластеров, полученных путем разбиения сети с высоким разрешением по Лёвену. Перекрестная таблица совместного появления меток кластеров в сети высокого разрешения (восемь кластеров, по горизонтали) и четырех сообществ анатомического разбиения (по вертикали).

Задача: *“Построение атласа головного мозга человека на основании данных диффузионного МРТ и граф-теоретических моделей.”*

Был предложен способ построения атласа головного мозга, на основе данных диффузионного МРТ с использованием непрерывной модели коннектома (continuous connectivity model [27]). Для вычислительных экспериментов были использованы данные Human Connectome Project [33]. Атлас был построен по данным 425 участников НСР (167 мужчин, 258 женщин).

Для восстановления индивидуальных кластеров был использован алгоритм Louvain Modularity, реализация igraph [6]. Для построения группового атласа был предложен использован мета алгоритм усреднения индивидуальных парцелляций. Было проведено сравнение разных подходов к усреднению: с использованием мета-графа, с использованием среднего графа, жадная оптимизация раскраски вершин графа таким образом чтобы минимизировать среднее Карчера (Karcher means):

$$C^* = \arg \min_C \sum_{i=1}^N d(C_i, C)^2,$$

где C_i - индивидуальные парцелляции полученные путем кластеризации непрерывного коннектома, а C^* - оптимальный групповой атлас.

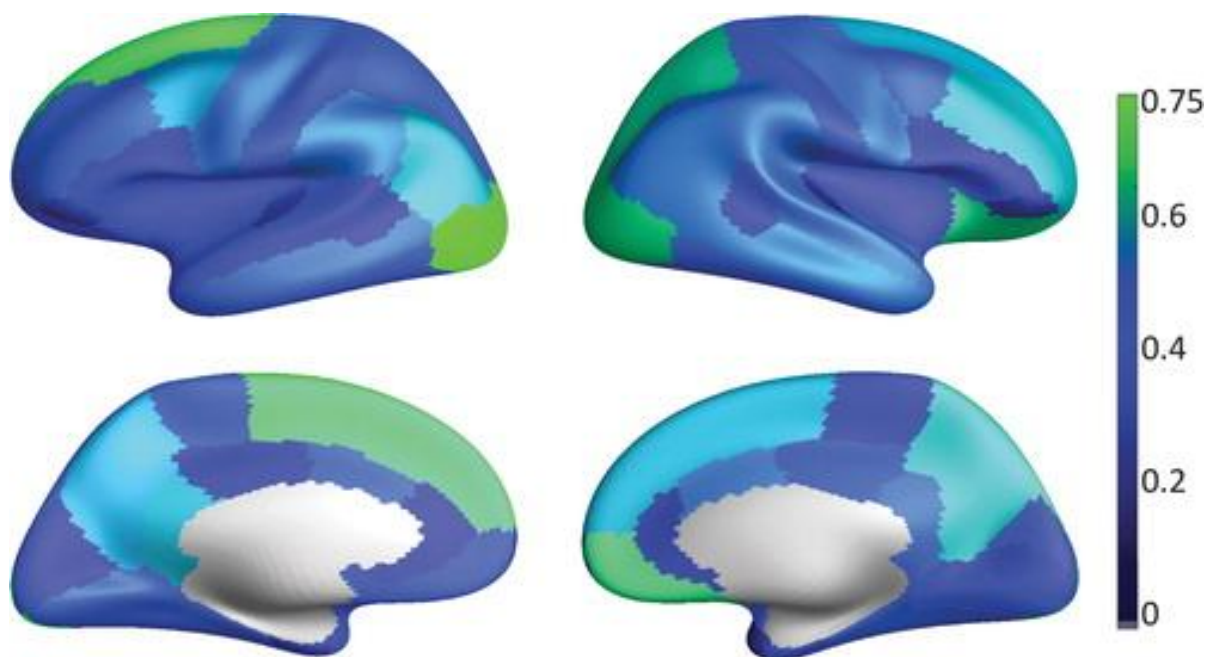


Рисунок 4. Совпадения регионов анатомического атласа Десикана и атласа построенного на основе топологических связей непрерывного коннектома.

Результаты показали, что атлас полученный таким способом может эффективно описывать данные о связности регионов головного мозга, превосходя по целому ряду параметров существующие атласы (например [7], [8], [9]) основанные на анатомии.

Задача: *“Моделирование и предсказания нейродегенеративных заболеваний без использования лонгитюдных данных.”*

Существует несколько подходов к моделированию прогрессии нейродегенеративных заболеваний, так например целый ряд работ основаны на гипотезе о том, что прогрессия биомаркеров, следует сигмоидальной траектории [17] в зависимости от стадии прогрессии, например толщина коры [11] головного мозга уменьшается, а CDR (Clinical Dementia Rating) маркер растет. Эта группа работ характеризуется различными способами подбора параметров сигмоиды на основе реальных данных. Например в работе [18] предложен способ подбора индивидуальных параметров прогрессии с использованием не более чем 4 временных наблюдений, в качестве модели прогрессии (в зависимости от времени) авторы используют линейную модель, а значения биомаркеров (таких как толщина коры, когнитивные показатели и др.) следует сигмоидальной траектории (в зависимости от стадии прогрессии).

Еще одним популярным подходом считается, так называемый Event-based modelling (EBM) [12]. Эта генеративная модель основана на следующей гипотезе: заболевание поражает различные участки головного мозга в определенном порядке. Сначала поражаются одни области, затем другие. Изменение участков головного мозга из конца этого списка соответствует более поздним стадиям развития болезни. В случае если порядок поражения областей может быть восстановлен, на его основе можно делать предсказания о текущем состоянии пациента даже не имея лонгитюдных данных. Кроме того в работе [39] было продемонстрировано что для восстановления этого порядка можно обойтись перекрестными данными (то есть данными в которых каждому пациенту соответствует только одно наблюдение). Работы основанные на этой гипотезе отличаются способом моделирования последовательности событий.

Несмотря на свою популярность такая модель обладает рядом недостатков, например предполагается что каждый конкретный биомаркер следует единой траектории в процессе прогрессии заболевания для всех пациентов. В работе [38] авторы постарались устранить этот недостаток путем разделения пациентов на гомогенные группы, внутри которых поведение биомаркеров отличается лишь незначительно. Другим значительным недостатком является то, что количество рассматриваемых областей в оригинальной работе было мало: использовался стандартный атлас поверхности головного мозга, разделяющий ее на 70 областей. В работе [23] авторы перешли от региональных признаков к признакам измеренным на уровне вершин полигональной сетки поверхности головного мозга (порядка сотен тысяч вершинных признаков). Более того, авторы использовали предположение о сигмоидальной траектории прогрессии биомаркера, таким образом их работа может быть рассмотрена как попытка преодолеть еще один недостаток ЕВМ подхода - а именно детерминированность последовательности пораженных регионов. В качестве еще одного недостатка отметим что информация о диагнозе используется в моделировании очень слабо или не используется вовсе, то есть в основном используются параметрические модели обучения без учителя такие как Gaussian Mixture Model (и а также смеси других параметрических распределений), поэтому в работе [34] авторами был использован комбинированный подход заключающийся в последовательном применении классификационной модели (этот этап помогает отделить ранние стадии от поздних, на основе данных), а затем - генеративной модели (на этом этапе выводится конкретный порядок поражения областей головного мозга). Последним значительным недостатком который, насколько нам известно не был затронут ни в одной из существующих работ, является соотношение между количеством подбираемых параметров M и величиной выборки N . Эффективность

современных методов машинного обучения во многом обусловлена увеличивающимися размерами обучающих выборок. Дело в том что число параметров прогностической модели обычно никак не зависит от величины выборки, таким образом отношение M/N стремится к 0 с увеличением обучающего набора данных. В последних модификациях EBM модели, проводится подбор индивидуальных параметров, таким образом величина M линейно растет с ростом N , что значительно ограничивает обобщаемость данного подхода.

Наконец последняя группа работ, это работы в которых прогрессия заболевания описывается в терминах распространения по “сетям” головного мозга - коннектомам [30], [16]. Такие сети могут быть построены на основе диффузионных или функциональных МРТ данных. Работы этого блока отличаются друг от друга характером (законом) распространения (прогрессии) заболевания по сети головного мозга, а также предположениями о том какие группы вершин (областей) или ребер (связей между областями) подвержены поражению в первую очередь. Так например в работе [4] рассматривается модель в соответствии с которой наибольшему риску подвержены “хабы”, то есть участки сети с наибольшим количеством связей (“nodal stress”), идея того что болезнь распространяется вдоль пучков нейронов (“trans neuronal spread”) рассмотрена в работе [19]. Согласно недавнему обзору [28] сетевой подход пока что не позволяет строить предсказаний относительно стадии заболевания или скорости прогрессии, поэтому не применим с клинической точки зрения.

В рамках решения данной задачи была предложена прогностическая модель прогрессии заболевания Болезнь Альцгеймера [22], которая учитывает не только статистические параметры популяции, но и индивидуальные факторы развития. При этом модель можно строить без использования лонгитюдных данных.

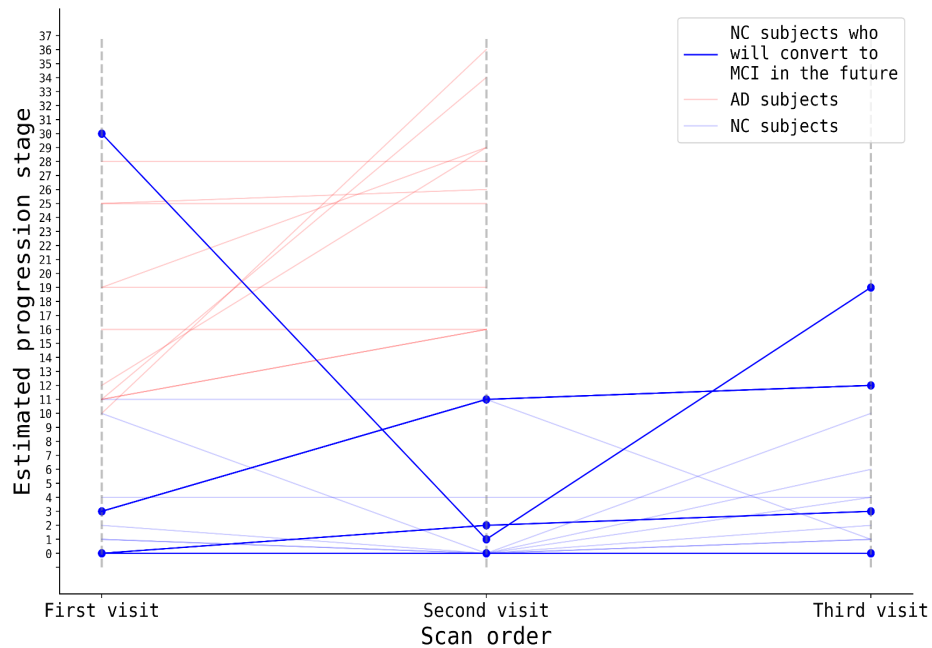


Рисунок 5. Несколько отобранных испытуемых. Стадии заболевания посчитанные с помощью модели прогрессии с индивидуальным коннектомным порядком на перекрестных данных. На изображении приведены стадии для наблюдений из тестового набора данных, каждая линия соответствует одному пациенту, красные линии соответствуют больным болезнью Альцгеймера, бледно-синие - здоровым людям, ярко-синие, людям здоровым во время первого визита, для которых известно что в будущем начали проявляться симптомы деменции.

Конкретно мы предложили расширение событийной (Event-based model) модели [13], которое позволяет инкорпорировать индивидуальные данные о связности областей головного мозга. Событийная модель прогрессии заболевания использует предположение о том что дегенерация отдельных областей происходит последовательно (разные области могут деградировать с разной скоростью), и своей задачей ставит восстановление этой последовательности. Для этого используется байесовская модель с

неинформативным априорным распределением. В нашей работе мы предложили информативное априорное распределение основанное на анатомической связи отдельных регионов. Вычислительные эксперименты показали, что использование данных диффузионного МРТ (вместе с биомаркерами полученными из структурного МРТ) может улучшать прогностические показатели, по сравнению с использованием только структурного МРТ. На Рис. 5 приведен пример предсказания оценок прогрессии для тестовой части пациентов. Пациенты с болезнью Альцгеймера получают большие оценки степени прогрессии по сравнению со здоровыми пациентами.

Список использованных источников и литературы

1. Abi-Dargham, A., Horga, G. (2016) The search for imaging biomarkers in psychiatric disorders. *Nature Medicine*, 22, 1248-1255.
2. Alexander-Bloch, A.F., Gogtay, N., Meunier, D., Birn, R., et al. (2010) Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Frontiers in Systems Neuroscience*, 4, 147, 1-17.
3. Betzel, R.F., et al. (2013) Multi-scale community organization of the human structural connectome and its relationship with resting-state functional connectivity. *Network Science* 1 (3), 353–373.
4. Buckner R. L. et al. Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease // *Journal of neuroscience*. – 2009. – Т. 29. – №. 6. – С. 1860-1873.
5. Craddock R.C., Jbabdi S., Yan C.G., Vogelstein J.T. (2013) Imaging human connectomes at the macroscale. *Nature Methods*, 10(6), 524-539.

6. Csardi, G. and Nepusz, T., 2006. The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695(5), pp.1-9.
7. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T. and Albert, M.S., 2006.
8. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, 31(3), pp.968-980.
9. Destrieux, C., Fischl, B., Dale, A. and Halgren, E., 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage*, 53(1), pp.1-15.
10. Meunier, D., Lambiotte, R., Bullmore, E.T. (2010) Modular and hierarchically modular organization of brain networks. *Frontiers of Neuroscience*, 4, 200, 1-11.
11. Fischl, B., Sereno, M.I. and Dale, A.M., 1999. Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. *Neuroimage*, 9(2), pp.195-207.
12. Fonteijn H. M. et al. An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease // *NeuroImage*. – 2012. – T. 60. – №. 3. – C. 1880-1889.
13. Fonteijn, H.M., Modat, M., Clarkson, M.J., Barnes, J., Lehmann, M., Hobbs, N.Z., Scahill, R.I., Tabrizi, S.J., Ourselin, S., Fox, N.C. and Alexander, D.C., 2012. An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. *NeuroImage*, 60(3), pp.1880-1889.
14. Fornito, A., Zalesky, A., Breakspear, M. (2015) The connectomics of brain disorders. *Nature Reviews Neuroscience*, 16, 159-172.

15. Frisoni G. B. et al. The clinical use of structural MRI in Alzheimer disease //Nature Reviews Neurology. – 2010. – T. 6. – №. 2. – C. 67.
16. Iturria-Medina Y. et al. Multifactorial causal model of brain (dis) organization and therapeutic intervention: application to Alzheimer's disease //Neuroimage. – 2017. – T. 152. – C. 60-77.
17. Jack Jr C. R. et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade //The Lancet Neurology. – 2010. – T. 9. – №. 1. – C. 119-128.
18. Jedynak B. M. et al. A computational neurodegenerative disease progression score: method and results with the Alzheimer's disease Neuroimaging Initiative cohort //Neuroimage. – 2012. – T. 63. – №. 3. – C. 1478-1486.
19. Jucker M., Walker L. C. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases //Nature. – 2013. – T. 501. – №. 7465. – C. 45.
20. Kurmukov A., Dodonova Y., Zhukov L. (2017) Classification of normal and pathological brain networks based on similarity in graph partitions. IEEE International Conference on Data Mining Workshops, ICDMW. <http://dx.doi.org/10.1109/ICDMW.2016.0023>.
21. Kurmukov A., Dodonova Y., Zhukov L. Machine learning application to human brain network studies: a kernel approach, in: Models, Algorithms, and Technologies for Network Analysis, Springer Proceedings in Mathematics and Statistics.
22. Kurmukov, A., Zhao, Y., Mussabaeva, A. and Gutman, B., 2019, October. Constraining Disease Progression Models Using Subject Specific Connectivity Priors. In International Workshop on Connectomics in Neuroimaging (pp. 106-116). Springer, Cham.
23. Marinescu R. V. et al. A vertex clustering model for disease progression: application to cortical thickness images //International Conference on

- Information Processing in Medical Imaging. – Springer, Cham, 2017. – C. 134-145.
24. Marinescu R. V. et al. TADPOLE Challenge: Prediction of Longitudinal Evolution in Alzheimer's Disease //arXiv preprint arXiv:1805.03909. – 2018.
 25. McKhann G. M. et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease //Alzheimer's & dementia. – 2011. – T. 7. – №. 3. – C. 263-269.
 26. Meunier, D., Lambiotte, R., Bullmore, E.T. (2010) Modular and hierarchically modular organization of brain networks. *Frontiers of Neuroscience*, 4, 200, 1-11.
 27. Moyer, D., Gutman, B.A., Faskowitz, J., Jahanshad, N. and Thompson, P.M., 2017. Continuous representations of brain connectivity using spatial point processes. *Medical image analysis*, 41, pp.32-39.
 28. Oxtoby N. P., Alexander D. C. Imaging plus X: multimodal models of neurodegenerative disease //Current opinion in neurology. – 2017. – T. 30. – №. 4. – C. 371.
 29. Querbes O. et al. Early diagnosis of Alzheimer's disease using cortical thickness: impact of cognitive reserve //Brain. – 2009. – T. 132. – №. 8. – C. 2036-2047.
 30. Seeley W. W. et al. Neurodegenerative diseases target large-scale human brain networks //Neuron. – 2009. – T. 62. – №. 1. – C. 42-52.
 31. Strimbu, Kyle; Jorge, Tavel (2010). "What are Biomarkers?". *Current Opinion in HIV and AIDS*. 5 (6): 463–466.
 32. Teipel S. et al. Multimodal imaging in Alzheimer's disease: validity and usefulness for early detection //The Lancet Neurology. – 2015. – T. 14. – №. 10. – C. 1037-1053.

33. Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K. and Wu-Minn HCP Consortium, 2013. The WU-Minn human connectome project: an overview. *Neuroimage*, 80, pp.62-79.
34. Venkatraghavan V. et al. Disease progression timeline estimation for Alzheimer's disease using discriminative event based modeling //NeuroImage. – 2019. – T. 186. – C. 518-532.
35. World Alzheimer Report 2015: The Global Impact of Dementia, <https://www.alz.co.uk/research/world-report-2015>
36. World Alzheimer Report 2018 The state of the art of dementia research: New frontiers, <https://www.alz.co.uk/research/world-report-2018>
37. Wu, K., et al. (2011) The overlapping community structure of structural brain network in young healthy individuals. *PLoS One*, 6, 5, e19608, 1-14.
38. Young A. L. et al. A data-driven model of biomarker changes in sporadic Alzheimer's disease //Brain. – 2014. – T. 137. – №. 9. – C. 2564-2577.
39. Young A. L. et al. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference //Nature communications. – 2018. – T. 9. – №. 1. – C. 4273.
40. Young J. et al. Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment //NeuroImage: Clinical. – 2013. – T. 2. – C. 735-745.
41. Sporns, Olaf, Giulio Tononi, and Rolf Kötter. "The human connectome: a structural description of the human brain." *PLoS computational biology* 1, no. 4 (2005): e42.