

Разработка вычислительного пакета для оценки сложности структуры и расчёта индекса Тонони (когнитивности) по фМРТ данным

З.А. Носовец^{1,2}

¹Московский физико-технический институт (государственный университет)

²Национальный исследовательский центр «Курчатовский институт»

В данном докладе рассматривается индекс Тонони, введённый для количественной оценки уровня сознания. Такая оценка может применяться, например, при проведении экспериментов для определения, в сознании ли человек или нет. Эта оценка не зависит от того, способен ли человек взаимодействовать с внешней средой. Для вычисления индекса требуется простое неинвазивное измерение. Метод вычисления индекса Тонони был разработан для данных ЭЭГ. В этом методе используется решение обратной задачи для определения источников активности и вычисление сложности Лемпеля-Зива на основе матрицы значимых источников. Согласно теории, разработанной Тонони, уровень сознания зависит от информационного содержания активности зон головного мозга. Преимущество данных фМРТ для вычисления индекса Тонони состоит в возможности наблюдать активные регионы без необходимости решения обратной задачи. К тому же, можно расширить задачу и учитывать межсетевые взаимодействия полушарий.

Основная задача работы состояла в разработке метода вычисления индекса Тонони для данных фМРТ и реализации этого метода в виде программы с пользовательским графическим интерфейсом. Программа позволяет не только вычислять индекс Тонони, но и предоставляет возможность для исследования зависимостей этого индекса от различных параметров (уровня отсечки, времени, местонахождения зон).

Индекс Тонони для данных ЭЭГ определяется как нормализованная сложность Лемпеля-Зива пространственно-временной картины корковой активации, вызванной прямым TMS возмущением [1].

Теоретические соображения позволяют предположить, что сознание зависит от способности мозга поддерживать сложную структуру активности, которая одновременно интегрирована (распределена среди взаимодействующих зон коры головного мозга) и дифференцирована в пространстве и времени (информационно богата) [2]. Индекс Тонони для данных ЭЭГ рассчитывается путём (I) возмущения коры головного мозга транскраниальной магнитной стимуляцией (TMS) для вовлечения распределённых взаимодействий в мозге (интеграция) и (II) сжатия пространственно-временной структуры этих взаимодействий для оценки её алгоритмической сложности (информация).

Последний шаг в вычислении индекса Тонони для данных ЭЭГ заключается в нахождении нормализованной сложности Лемпеля-Зива матрицы значимых источников с целью определения её информационного содержания [1]. Для фМРТ данных предлагается заменить матрицу источников на матрицу корреляций сигнала из всех областей головного мозга, и в зависимости от коэффициента корреляции считать взаимодействие значимым или нет.

Теория интегрированной информации Тонони утверждает, что сознание есть интегрированная информация [2]. А именно количество сознания соотносится с количеством интегрированной информации, созданным совокупностью элементов. Интегрированная информация – информация, генерируемая системой через причинное взаимодействие между её элементами. Эти элементы работают вместе как интегрированная система, поэтому их производительность ухудшается, если они разъединены. В качестве этих элементов системы рассматриваются области головного мозга.

Функциональная интеграция – изучение связанных процессов. Одним из методов функциональной интеграции является функциональная коннективность, то есть нахождение статистических связей или зависимостей между двумя или более анатомически отделёнными временными рядами [3]. В качестве этих зависимостей берутся корреляции временных рядов. Тулбокс MATLAB MarsBaR извлекает сигнал для каждого вокселя в пределах ROI (ROI определяет воксели в пределах изображения, которые представляет интерес) для каждого изображения, таким образом предоставляя воксельные временные ряды. Затем вычисляет новый

временной ряд для каждого ROI, усредняя все значения для вокселей, попадающих внутрь ROI. Связанность между ROI получается с помощью коэффициента корреляции Пирсона между извлечёнными временными рядами в каждом ROI, при этом создаётся набор матриц корреляции для каждого субъекта. Коэффициенты корреляции рассчитываются не на полных временных рядах, а на нарезанных на окна по времени. Формула для коэффициента корреляции Пирсона:

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

Это может быть переписано на основании оценок ковариации и стандартных отклонений:

$$\text{corr}(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1) \sqrt{\frac{\sum_{j=1}^n x_j^2 - n\bar{x}}{n-1}} \sqrt{\frac{\sum_{j=1}^n y_j^2 - n\bar{y}}{n-1}}} = \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sqrt{\sum_{j=1}^n x_j^2 - n\bar{x}}} \frac{y_i - \bar{y}}{\sqrt{\sum_{j=1}^n y_j^2 - n\bar{y}}} \right)$$

Для создания матрицы значимых источников (или взаимодействий) и для последующего применения алгоритма Лемпеля-Зива для бинарной последовательности, нужно эту последовательность сначала соорудить. Коэффициенты корреляционной матрицы принимают вещественные значения от -1 до 1, следовательно, нужно применить к матрице некоторое пороговое значение (отсечку) для разделения на значимые и незначимые корреляции. Корреляция считается значимой, если её коэффициент больше, чем значение отсечки, и незначимой в другом случае. Таким образом, рассматриваются только высокие корреляции. От значения этой отсечки сам индекс Тонони тоже зависит, поэтому можно провести исследования индекса от отсечки. Если интересует определённое число зон, то можно построить отсечку автоматически, предварительно отсортировав матрицы корреляций.

Сначала нужно получить бинарную матрицу пространственно-временной картины активации. Для этого следует развернуть корреляционные матрицы для каждого значения окна в столбец, перед этим убрав повторяющиеся из неё повторяющиеся коэффициенты (корреляционная матрица симметрична). Назовём эту матрицу SC .

Алгоритмическая сложность одномерной строки была изначально определена Андреем Колмогоровым и другими как длина самой короткой компьютерной программы, которая может генерировать строку [4]. Поскольку алгоритмическая сложность в общем случае не вычислима, были разработаны меры приближения для её оценки. Здесь используется одна из наиболее популярных мер сложности класса Колмогорова, введённая Лемпелем и Зивом [5]. Эта мера приближает количество содержащейся в строке избыточной информации, оценивая минимальный размер "словаря", необходимого для описания строки. Строки с высокой сложностью Лемпеля-Зива требуют большого количества различных паттернов ("слов") для их воспроизведения, а строки с низкой сложностью Лемпеля-Зива могут быть в значительной степени сжаты, и для устранения их избыточности без потери информации используется несколько "слов". Простая процедура может быть легко реализована для вычисления меры Лемпеля-Зива для двоичной последовательности произвольной длины L . Алгоритм сканирует последовательность и ищет подпоследовательности последовательных символов или "слов". Сложность Лемпеля-Зива C_L составляет число, описывающее сколько раз встречается новая подпоследовательность. В данной работе реализуется прямое обобщение этого алгоритма, который также может быть применён к бинарным матрицам. В этом случае алгоритм проходит по первому столбцу входной матрицы с поиском паттернов; этот поиск повторяется для каждого последующего столбца, сохраняя при этом паттерны, встречающиеся в предыдущих столбцах. Асимптотика C_L для случайных строк – $LH(L)/\log_2 L$, где $H(L)$ является энтропией источника,

$$H(L) = -p_1 \log_2 L - (1-p_1) \log_2 (1-L)$$

и p_1 есть доля "1", содержащаяся в двоичной строке длины L . Так как максимальная сложность чрезвычайно чувствительна к энтропии источника, как правило, принимается во внимание

нормированная сложность Лемпеля-Зива $\overline{c_L}$, где $\overline{c_L} = c_L \frac{\log_2 L}{LH(L)}$. Для строго случайных последовательностей асимптотически в $L \overline{c_L} = 1$.

Определим индекс Тонони как нормированную сложность Лемпеля-Зива ($\overline{c_L}$) пространственно-временной структуры корковой активации SC . Пусть c и t – дискретные индексы пространственных (корреляции) и временных измерений соответственно, $c = 1 \dots L_1$ и $t = 1 \dots L_2$, пусть L – общее число пространственно-временных отсчётов SC , $L = L_1 \times L_2$, и пусть $l(t)$ – количество пространственно-временных выборок SC до времени t , $l(t) = L_1 \times t$. Имея SC в качестве входных данных, для каждого отсчёта времени t алгоритм вычислит число $c_l(t)$, которое соответствует сложности Лемпеля-Зива двумерной последовательности длины $l(t)$. Индекс Тонони затем приравнивается к сложности Лемпеля-Зива полной двумерной последовательности длины L , $c_L = c_l(t = L_2)$, нормированный на $LH(L) / \log_2 L$:

$$TI = c_l(t = L_2) \frac{\log_2 L}{LH(L)}$$

Из определения следует, что индекс Тонони является мерой, которая увеличивается с числом различных пространственных паттернов бинарной значительной корковой активности, происходящей на данном временном отсчёте, а не на предыдущих. Действительно, оценивая алгоритмическую сложность пространственных структур, а не временных, алгоритм способен обеспечить динамическую меру сложности,

$TI(t) = c_l(t) \frac{\log_2 L}{LH(L)}$, которая монотонно возрастает с появлением во времени избыточных пространственных связей значительной корковой активации.

Данный метод был реализован с помощью пакета MATLAB 2013b. Программа позволяет выбрать формат входных данных (изображения, временные ряды или матрицы корреляции), вычислить индекс Тонони, а также найти зависимость индекса от времени и от отсечки с учётом местонахождения интересующих регионов в требуемом полушарии, задаваемом в качестве параметров.

Для апробации разработанного метода и приложения использовались данные, предоставленные лабораторией когнитивных технологий. Пациенты с нарушением языковых функций (повреждено левое полушарие). Результаты работы программы:

- Индексы Тонони без учёта местонахождения ROI по регионам показали более низкие значения для пациентов (от 0.26 до 0.38), чем для нормы (от 0.47 до 0.59) при уровне отсечки 0.5
- При возрастании отсечки индекс Тонони убывает (не всегда линейно (рис. 2))
- У пациентов с повреждённым левым полушарием индекс Тонони низкий, если его рассчитывать только по активности внутри левого полушария, а также по активности между левым и правым полушарием. Для правого полушария индекс выше. у здоровых индекс Тонони в целом выше
- При учёте местонахождения ROI у пациентов при отсечке 0.5 индекс Тонони для активности внутри левого полушария ниже, чем внутри правого (рис. 1)
- Активность между левым и правым полушарием самая низкая для пациентов и самая высокая для здоровых, о чём свидетельствует индекс Тонони
- При отсечке, близкой к 0.7, индексы смешиваются из-за сильной разреженности матрицы, очень мало зон учитывается для вычисления сложности Лемпеля-Зива

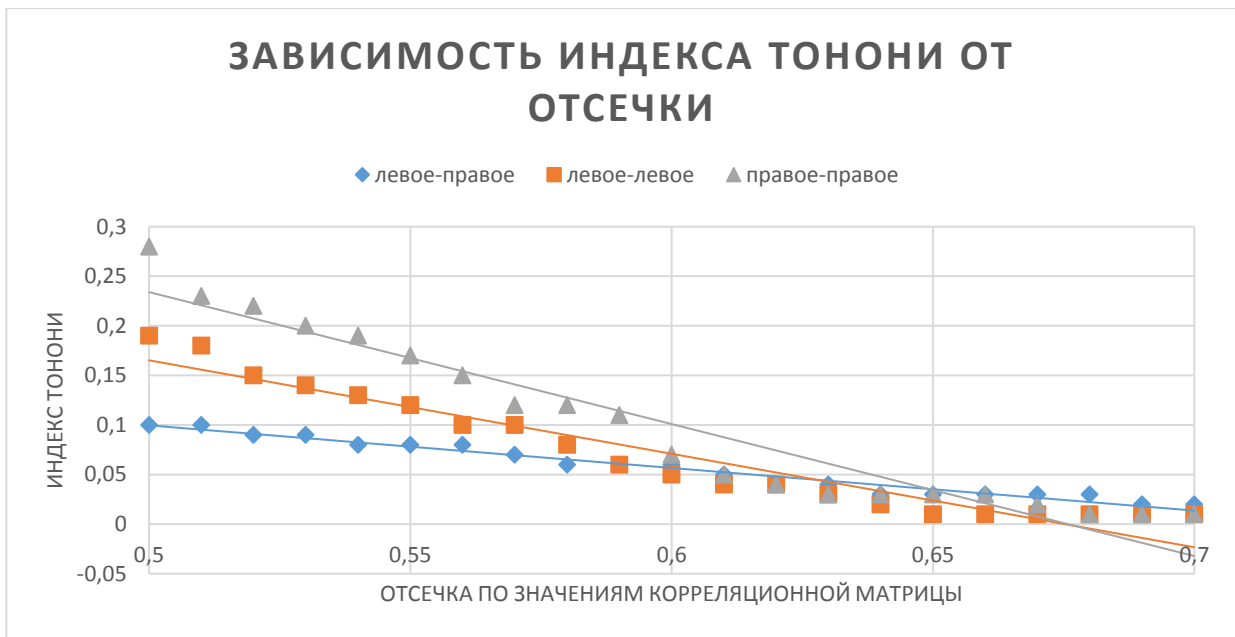


Рис. 1. Зависимость индекса Тонони от отсечки для больных

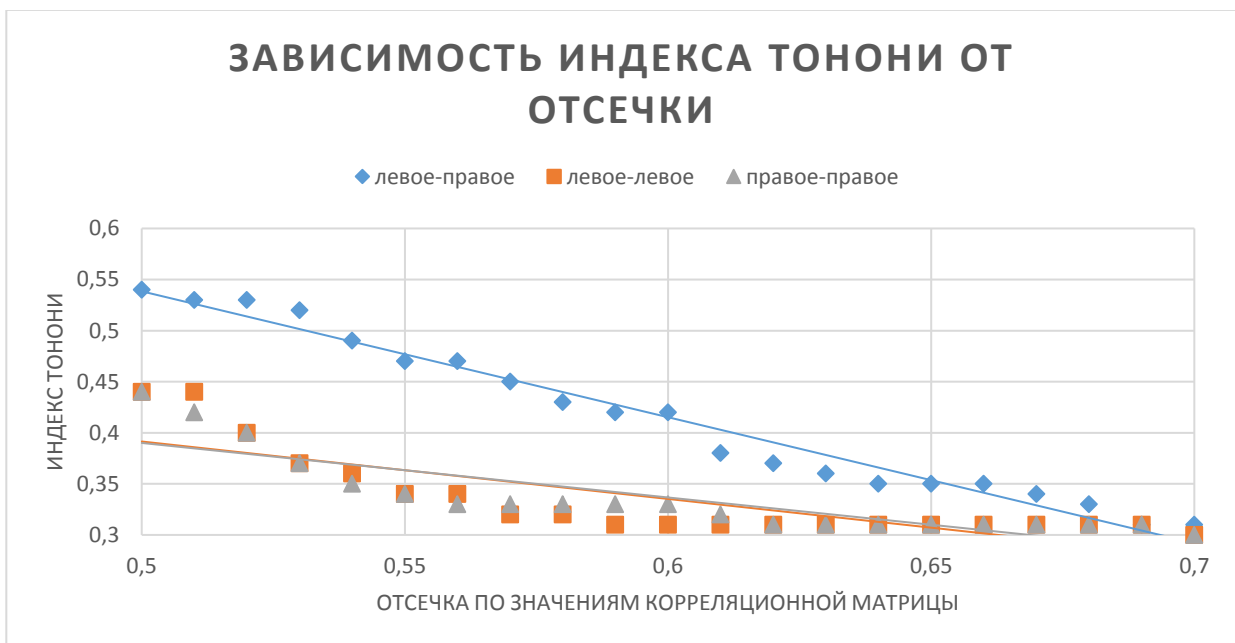


Рис. 2. Зависимость индекса Тонони от отсечки для здоровых

Литература

1. Casali, Adenauer G.; Gosseries, Olivia; Rosanova, Mario; Boly, Mélanie; Sarasso, Simone; Casali, Karina R.; Casarotto, Silvia; Bruno, Marie-Aurélie; Laureys, Steven (2013-08-14). "A Theoretically Based Index of Consciousness Independent of Sensory Processing and Behavior". *Science Translational Medicine* 5 (198): 198ra105.
2. G. Tononi, An information integration theory of consciousness. *BMC Neurosci.* 5, 42 (2004).
3. A. Vanhaudenhuyse; Default network connectivity reflects the level of consciousness in non-communicative braindamaged patients. [Brain](#). 2010 Jan; 133(Pt 1):161-71.
4. A. N. Kolmogorov; Three approaches to the quantitative definition of information; *International Journal of Computer Mathematics* [Volume 2, Issue 1-4](#), 1968
5. A. Lempel, J. Ziv, On the complexity of finite sequences. *IEEE Trans. Inform. Theory* 22, 75–81 (1976).