

2017 Том 10 No. 55

Морозова О.А. Структурное сетевое моделирование в когнитивной науке



МОРОЗОВА О.А. СТРУКТУРНОЕ СЕТЕВОЕ МОДЕЛИРОВАНИЕ В КОГНИТИВНОЙ НАУКЕ
English version: [Morozova O.A. Structural network modelling in cognitive science](#)

Институт психологии Российской академии наук, Москва, Россия

[Сведения об авторе](#)
[Литература](#)
[Ссылка для цитирования](#)

Многие объекты исследования когнитивных наук целесообразно описывать как сеть. В сетевой модели узлы (клетки, индивиды, группы, слова, категории и т.п.) характеризуются через связи, которые они (не) имеют, устанавливают и теряют. Таким образом, сетевая модель смещает исследовательские акценты со свойств элементов на связи между ними, эволюцию этих связей и – как следствие – целостность системы. Традиционно сетевое моделирование развивалось в рамках концептуального подхода (символьные модели А.Коллинза и Э.Лофтус, Дж.Р.Андерсона, нейронные сети Д.Румельхарта, Дж.Хинтон и др.). Недостаток подхода заключается в том, что концептуальные модели являются по сути изложением идей автора о строении когнитивной системы и часто опираются на гипотетические конструкции (чанки, блоки, искусственные нейроны и т.п.). С начала XXI века в когнитивной науке начинает набирать популярность другой, структурный, подход к сетевому моделированию. В отличие от концептуальной, структурная модель является непосредственной визуализацией массива данных, описывающих систему. В качестве массива данных могут выступать результаты МРТ, орфографический словарь, база социальных контактов, лог входящих и исходящих пакетов информации, ассоциативный тезаурус и т.д. Топология получившейся сети затем анализируется математическим аппаратом науки о сетях (computational network science). Результаты анализа позволяют выдвигать гипотезы об эволюционных закономерностях, определивших наблюдаемую структуру, а также о процессуальных следствиях – о влиянии структуры системы на протекание психических процессов. В статье изложены основные принципы, понятия и цели структурного сетевого моделирования. Кратко изложена история математических сетевых моделей: от простых графов к сложным сетям. Рассмотрена специфика структурного моделирования в приложении к объектам исследования когнитивных наук.

Ключевые слова: сетевое моделирование, безмасштабная сеть, сеть малого мира, когнитивная наука, ассоцианизм, ментальный лексикон

В когнитивной науке, в которую входит экспериментальная психология познания, сеть является одним из наиболее распространенных типов информационных моделей. Популярность метода является следствием его выразительной мощности. Обычно сеть строится из двух составляющих – узлов как элементов сети и ребер, отражающих взаимодействие между элементами. Несмотря на базовую простоту, эти составляющие позволяют описывать большой спектр объектов разной природы и уровней сложности.

При этом сетевое моделирование не существует в едином наборе принципов, а развивается в рамках двух подходов, в зависимости от способа работы с данными. Для простоты обозначения назовем модели, создаваемые в рамках первого подхода, «концептуальными», а в рамках второго –

«структурными». Концептуальная сетевая модель создается ученым-когнитивистом для описания гипотетического строения и работы изучаемой системы. Такая модель призвана объяснить результаты экспериментальных исследований и часто опирается на гипотетические конструкции (чанки, искусственные нейроны и т.п.). Итоговое строение сети является отражением идей ее создателя о строении когнитивной системы.

Первыми концептуальными моделями в когнитивных науках были модели вербальных сетей, предложенные в психолингвистике, к примеру – для объяснения феноменов генерализации условных реакций [Cofer, Foley, 1942]. В дальнейшем опыт концептуального сетевого моделирования складывался в двух направлениях – символическом и коннекционистском. Классические символические модели принимают за объект моделирования семантическую память и описывают мышление как процесс оперирования элементами памяти – концептами или чанками информации. Семантическая сеть в символическом подходе является механизмом распространения активации и может дополняться описанием отдельной системы, отвечающей за правила вывода – основные операции с элементами памяти. Символические модели были впервые строго описаны на материале психолингвистики в работах М.Киллиана, а позже А.Коллинза и Э.Лофтуса [Collins, Loftus, 1975], с целью объяснения процессов восприятия и понимания речи. В более широком контексте подход был развит в концепции когнитивной архитектуры АСТ Дж.Р.Андерсона [Anderson, 1983].

Коннекционистское направление использует в своей основе метафору нейронной сети и принцип параллельной переработки информации. Сетевые модели коннекционизма, разрабатываемые на стыке психологии и компьютерных наук, получили название искусственных нейронных сетей. В таких моделях сеть является универсальным механизмом – она отвечает не только за распространение активации, но и за операции с элементами сети. Понятия не присваиваются отдельным элементам сети, а могут образовываться из их констелляций по аналогии с нейронами. Коннекционистский подход получил развитие в работах таких ученых, как Д.Румельхарт, Дж.МакКлелланд [Rumelhart, McClelland, 1986] и Дж.Хинтон [Hinton, 1986].

В рамках структурного подхода исследователь прибегает к сетевому моделированию не для объяснения, а для непосредственной визуализации массива данных, описывающих систему. Свойства данных сами задают структуру сети – без вмешательства автора. Поэтому топология итоговой сети понимается как прямое отражение строения изучаемой системы. Для извлечения ценных данных об этом строении топология сети подвергается анализу статистических инструментов. Полученные метрики позволяют исследователю перейти к постановке гипотез о закономерностях, приведших к образованию конкретной топологии, и о ее влиянии на работу системы. Эти гипотезы могут быть проверены в компьютерной модели роста сети и динамики на ее основе. Надежность и описательная мощь структурной модели в первую очередь обеспечиваются корректностью визуализации и статистической обработки данных.

Если история концептуальных моделей в когнитивных науках богата примерами, то про структурные модели такого сказать нельзя. Главная причина кроется в ограничениях, задаваемых уровнем развития аппарата анализа сетевых структур. Доступные математические модели сетей определяют спектр объектов, которые возможно и целесообразно моделировать этим методом.

История структурных сетевых моделей

Сетевой подход к изучению мира берет свое начало с 1735 года, когда Леонард Эйлер (1707–1783) заложил основы новой математической дисциплины – теории графов [Barabasi, 2002, p. 9]. Теория графов оказалась продуктивной в изучении упорядоченных объектов, таких как решетки атомов в графите, гексагональные решетки пчелиных ульев и ходы шахматных фигур. Сети, которые отражали строение подобных объектов, получили название регулярных / простых [Strogatz, 2001]. Примеры таких сетей представлены на рис. 1, а и б.

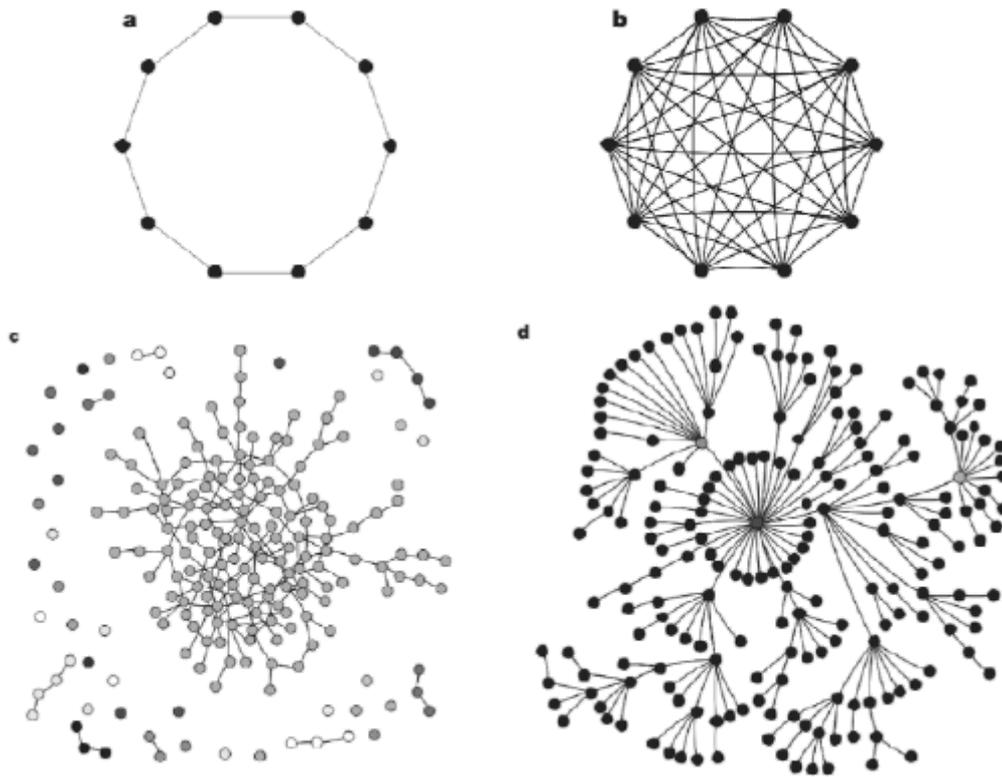


Рис. 1. Лог входящих и исходящих пакетов. Примеры сетевых моделей.

Примечания. а и b – регулярные графы, с – случайный граф, d – безмасштабный граф [Strogatz, 2001].

Важной вехой в развитии теории графов стала середина прошлого столетия. Тогда венгерские математики П.Эрдеш и А.Реньи признали, что большинство объектов, которые мы встречаем в природе, не регулярны, они неоднородны и сложны – а значит, должны быть описаны другой сетевой моделью [Barabasi, 2002, p. 21–22]. Они предложили простейший механизм, которому могла подчиняться природа в создании своих сетей. Это был механизм случайного соединения узлов (см. рис. 1 с).

Теория случайных графов, предложенная П.Эрдешем и А.Реньи, стала основным направлением изучения сетевых структур вплоть до конца XX века. Иногда эти сети называют пуассоновскими случайными сетями, потому что распределение по числу связей у узлов (степени) в таких сетях подчиняется закону С.Пуассона. К сожалению, попытки П.Эрдеша и А.Реньи описать реальные сложные системы не выдерживали критики.

В конце XX века интерес к сетевому моделированию усилился из-за открытия доступа к многочисленным базам данных: о сотрудничестве актеров кино, о соавторстве ученых в различных областях науки, о белковых взаимодействиях в живых клетках, о структуре нейронной сети нематоды *C. elegans*, о ссылках в World Wide Web и т.д. Эти массивы данных позволяли идти прямой дорогой от объекта к модели: простым алгоритмом можно было визуализировать структуру и функции реально существующей системы, а затем абстрагировать закономерности ее эволюции.

Одновременно идейный толчок поступил из социологии. Так, в мае 1973 года в *American Journal of Sociology* была издана статья Марка Грановеттера под названием «Сила слабых связей» [Granovetter, 1973]. Эта статья на данный момент считается одной из самых влиятельных статей, когда-либо написанных в этой области науки. В статье М.Грановеттер воспользовался аналогией с тем, как слабые водородные связи держат вместе огромные молекулы воды, чтобы проиллюстрировать значимость слабых социальных связей в нашей жизни, практически во всех социальных процессах (сам М.Грановеттер рассматривал процесс поиска работы). Если принять на вооружение идею М.Грановеттера, социум можно представить в виде нерегулярной сети, состоящей из набора законченных графов – маленьких кластеров, в которых каждый узел соединен с каждым другим. Эти кластеры, в свою очередь, соединены друг с другом сравнительно небольшим количеством слабых связей.

Идея Грановеттера иллюстрировала представление о том, что социальный мир является «малым миром», или, другими словами, что любые два человека в сети знакомств отделены друг от друга ничтожно маленьким количеством «рукопожатий». Гипотеза «малого мира» была предложена математиком Ф.Каринти еще в 1929 году и подробнее разработана М.Гуревичем в эмпирическом исследовании социальных сетей в 1961 году. Математик М.Кохен интерпретировал эти результаты, заявив, что два человека в мире разделены не более чем тремя контактами [Barabasi, 2002, p. 35–39]. В 1973 году он провел компьютерную симуляцию данных, собранных М.Гуревичем, учитывающую и сильные, и слабые связи. Результаты симуляции указали на расстояние в три «рукопожатия» между жителями США и поддержали экспериментальные находки С.Милграма [Milgram, 1967].

В 1998 году профессор прикладной математики из Корнельского университета Стивен Строгатц и его аспирант Данкан Воттс опубликовали в журнале Nature статью, в которой показали, что такие реальные сетевые структуры, как сеть нейронных связей нематоды, сеть сотрудничества актеров Голливуда и сеть из приблизительно 5000 электростанций США, также обладают свойством «малого мира» [Watts, Strogatz, 1998]. Так, они обнаружили, что нейронная сеть нематоды обладает высоким коэффициентом кластеризации: соседи нейрона в сети в пять раз более вероятно связаны друг с другом, чем если бы связи распределялись случайно.

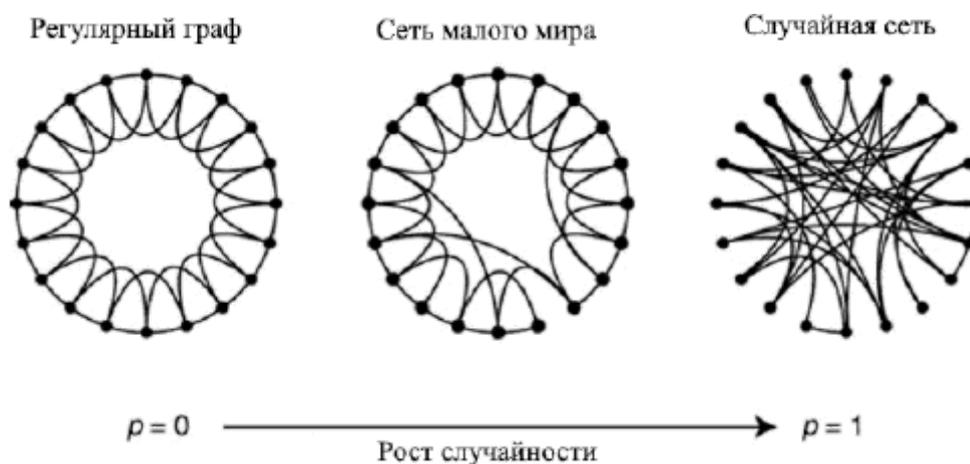


Рис. 2. Вклад случайности в сетевых моделях разного типа [адаптировано из Watts, Strogatz, 1998].

С.Строгатц и Д.Воттс предложили математическую модель сетей «малого мира», свойства которой представляют собой суперпозицию свойств регулярных, периодических решеток и пуассоновских случайных сетей (см. рис. 2) [Watts, Strogatz, 1998]. В такой модели, как и в модели П.Эрдеша, связи образуются случайным образом, но роль случайности математически ограничена повышенной вероятностью установления связей между соседями узла (узлами, связанными с ним). В результате мы получаем аналог идеи М.Грановеттера: сеть с большим количеством кластеров и небольшим количеством связей между кластерами, которые поддерживают, однако, общую связанность сети. Сколько бы узлов ни содержала сеть малого мира, расстояние между любыми двумя узлами в такой сети оказывается сравнительно коротким.

Сетевая модель «малого мира», разработанная С.Строгатцем и Д.Воттсом имела нерегулярную топологию. Она получила название «сложной» в противопоставление классическим регулярным графам. Узлы в сложной сети могут выполнять разные роли, быть более или менее важными для поддержания целостности сети.

В 1999 году ученый-физик из университета Нотр-Дам А.-Л.Барабаши изучал свойства реальных сетей с несколько иной точки зрения. Если С.Строгатц и Д.Воттс исходили из феномена «малого мира», то А.-Л.Барабаши решил исследовать закон распределения узлов реальных сетей по числу связей [Barabasi, 2002, p. 31–32]. Для многих сетей (World Wide Web, сети сотрудничества актеров Голливуда, метаболических сетей, лингвистических сетей, структуры авиационных сообщений в США и т.д.) вместо ожидавшегося вероятностного распределения узлов по числу связей q по закону

Пуассона полученное распределение приблизительно подчиняется свойственному всем критическим состояниям степенному закону (powerlaw) [Barabasi, 2002, p. 86–87; Mitchell, 2011, p. 247–257].

В отличие от нормального распределения, у степенного закона нет пика. Это постепенно снижающаяся кривая, подразумевающая, что большое количество маленьких событий сосуществует с маленьким количеством гигантских событий [Mitchell, 2011, p. 239–243]. Как пишет А.-Л.Барабаши, если бы распределение людей по росту подчинялось степенному закону, большинство людей были бы карликами, но при этом никто не удивился бы встретить на улице человека ростом в тридцать метров [Barabasi, 2002, p. 67]. В случайных сетях (модели П.Эрдеша и А.Реньи и модели Д.Воттса и С.Строгатца) распределение степеней у узлов довольно маленькое. У этих сетей есть характеризующий их масштаб. Он выражается через среднее количество связей (среднюю степень) и фиксированный пик степенного распределения. В сетях, подчиняющихся степенному закону (модель А.-Л.Барабаши), нет характерного узла, то есть нет характерного масштаба. В таких сетях небольшое число узлов имеет громадное число связей, а огромное число узлов содержит лишь несколько связей.

Такие сети получили название «безмасштабных» (scale free) (см. рис. 1 d). Это название не было придумано специально для этого типа сетей, а было взято из теории критических явлений, где флуктуации в критических состояниях также подчиняются степенному закону, а саму теорию безмасштабных сетей стали рассматривать как один из сценариев выхода сложных систем в критическое состояние.

А.-Л.Барабаши предложил простую и элегантную модель возникновения и эволюции безмасштабных сетей. Она показала, что для возникновения безмасштабности необходимо предпочтительное связывание: вероятность, с которой новый узел соединится с уже существующим узлом, зависит от числа связей, которым этот узел уже обладает [Barabasi, 2002, p. 86–87]. Степень предпочтительного связывания отражается в мере ассортативности: корреляции между степенью узла и степенями его соседей. Так, в социальных сетях эта корреляция положительна (популярные люди дружат с популярными людьми), а в биологических и технологических – отрицательна (ссылки на странице Википедии ведут на частные конечные сайты).

Как показали исследования последних лет, большинство сетей в живой природе являются сетями малого мира, множество из них также являются безмасштабными [Barabasi, 2012]. Анализ структуры этих сетей позволяет делать предположения об эволюционных процессах, приведших к формированию того или иного строения системы. Для исследования структуры сложных сетей в XXI веке была очерчена новая междисциплинарная отрасль науки – наука о сложных сетях (network science).

Аппарат создания и анализа структурно-сетевой модели

Исследователь прибегает к структурному сетевому моделированию для визуализации массива однородных данных, описывающих некий объект. Сетевая модель является графом с N узлов и L ребер (связей). Так, например, исследователь может визуализировать базу данных о свободных ассоциациях, где узлами являются слова, а ребра между узлами отражают связь «слово-стимул + слово-ответ». Если ребра между узлами содержат информацию о направлении связи («мышь» ассоциируется с «сыром», но «сыр» не ассоциируется с «мышью»), то сеть называется направленной. Ребра также могут иметь разный вес («мышь» гораздо чаще ассоциируется с «сыром», чем с «хлебом»). Если между двумя узлами i и j имеется связь, их называют соседями.

Путь в сети означает последовательность шагов между узлами, в которой шаг равняется одному ребру. Например, в семантической сети между словами «мышь» и «корова» нет прямой ассоциации, однако от одного узла до другого существует путь через цепь ассоциаций «мышь» – «сыр» – «молоко» – «корова». Такой путь будет равен трем шагам. Таким образом, длина пути между i и j равняется $n - 1$, где n – количество узлов от начального до конечного.

Визуализация данных в структурно-сетевой модели является первым шагом, но сила метода заключается в возможности извлечения важных знаний о системе через статистический анализ топологии сети. Представляется, что топология несет эволюционный отпечаток и функциональную

необходимость [Borge-Holthoefler, Arenas, 2010]. Подробный разбор доступных метрик можно найти, например, в [Carrington et al., 2015]. Рассмотрим лишь несколько метрик, часто используемых в исследованиях когнитивных моделей.

1. *Степень* (degree) узла i отражает количество его связей, или, другими словами, количество его соседей. В направленной сети k_o (out-degree) – количество исходящих связей, k_i (in-degree) – количество входящих связей. Степень является локальным признаком, характеризующим отдельный узел. Однако можно посчитать среднюю степень $\langle k \rangle$ – глобальную характеристику сети:

$$\langle k \rangle = \frac{\sum_{i=1}^N k_i}{N},$$

где N – количество узлов.

В фонологической сети, к примеру, степень слова может отражать количество слов, сходных с ним по звучанию [Vitevitch et al., 2011]. Средняя степень может говорить о том, насколько высоко сходство звучания для выборки слов.

2. *Короткие пути*. Выше мы обсудили понятие «путь». В некоторых случаях от одного узла к другому можно добраться большим или меньшим количеством шагов. Например, в ассоциативной сети слова «мышь» и «сыр» могут быть связаны напрямую, а также цепочкой «мышь» – «грызть» – «сыр». Понятие «короткий путь» (shortest path) обозначает минимальное количество шагов, которое необходимо совершить, чтобы добраться от i к j . Если между i и j существует лишь один путь, он и является «коротким путем».

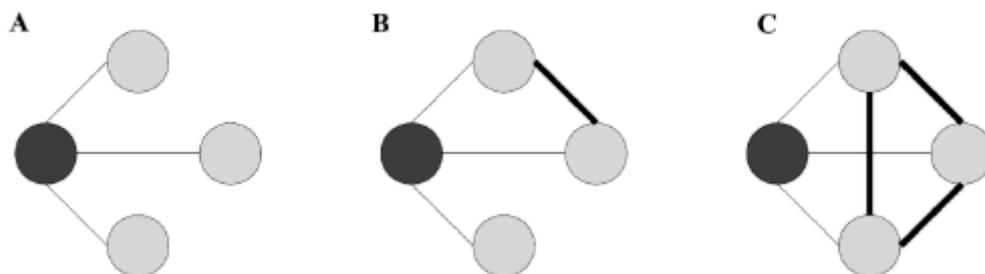


Рис. 3. Коэффициент кластеризации в ненаправленной сети.

Примечания. (а) – нет связей между соседями узла, $C = 0$; (б) – одна связь между соседями узла, $C = 1/3$; (с) – все соседи узла связаны друг с другом, $C = 3/3$.

3. *Коэффициент кластеризации* (clustering coefficient) отражает, сколько соседей узла также являются соседями друг друга (см. рис. 3). Например, в фонологической сети коэффициент кластеризации слова отражает, насколько слова, схожие с этим словом по звучанию, также сходны по звучанию друг с другом. Коэффициент кластеризации C_i узла i является пропорцией количества существующих связей (e_i) между его соседями (k_i) относительно количества всех возможных связей между его соседями:

$$C_i = \frac{2e_i}{k_i(k_i - 1)}.$$

Средний коэффициент кластеризации сети $\langle C \rangle$ является средним значением C_i для всех узлов в сети:

$$\langle C \rangle = \frac{1}{N} \sum_i c_i$$

4. *Меры центральности* (centrality measures) позволяют ранжировать узлы в сети по их относительной важности. Условно важность понимается как ответственность узла за связанность (проводимость) сети (ее участка), но конкретные трактовки варьируют. Примером меры центральности является оценка промежуточности (betweenness) узла. Эта мера была предложена Д.Фриманом на материале социологии [Freeman, 1977]. Ученый предположил, что узел в сети важен настолько, насколько он является промежуточной станцией на коротких путях других узлов друг к другу. Иными словами, для оценки промежуточности необходимо посчитать, сколько коротких путей проходит через данный узел. По словам самого Д.Фримана, эта мера отражает ответственность узла, какой «стресс» придется на него во время активации сети. Оценка промежуточности узла i рассчитывается по формуле:

$$C_B(i) = \sum_{j \neq i \neq k} \frac{\sigma_{jk}(i)}{\sigma_{jk}}$$

где σ_{jk} – количество кратчайших путей от j к k ; $\sigma_{jk}(i)$ – количество коротких путей от j к k , проходящих через узел i .

Родственным мерам центральности является понятие *коэффициента фрагментации* [Vitevitch, Goldstein, 2014]. Эта мера позволяет выявить ключевые узлы, удаление которых приводит к значительному удлинению коротких путей и разрыву сети. В экстремальном случае удаление ключевых узлов приводит к полной фрагментации сети на отдельные компоненты.

5. *Ассортативность* (degree-degree correlation). Мера ассортативности оценивает, есть ли корреляция между степенью узла и степенями его соседей. В ассортативных сетях (положительная корреляция) узлы с большим количеством связей (высокой степенью) стремятся быть связанными с узлами, у которых сопоставимо высокое количество связей, и наоборот. В дисассортативных сетях (отрицательная корреляция) узлы с высокой степенью стремятся быть связанными с узлами с низкой степенью (см. рис. 4). Корреляция рассчитывается через коэффициент корреляции Пирсона для степеней узлов на обоих концах связи.

Дисассортативность считается нежелательным топологическим свойством сети. Ассортативные сети более устойчивы к удалению узлов, однако они также отличаются более высокой транзитивностью. Например, в эпидемиологии из-за ассортативности социальных сетей заражение человека с большим количеством контактов с высокой вероятностью приводит к заражению людей со столь же большим количеством контактов [Keeling, 2005].

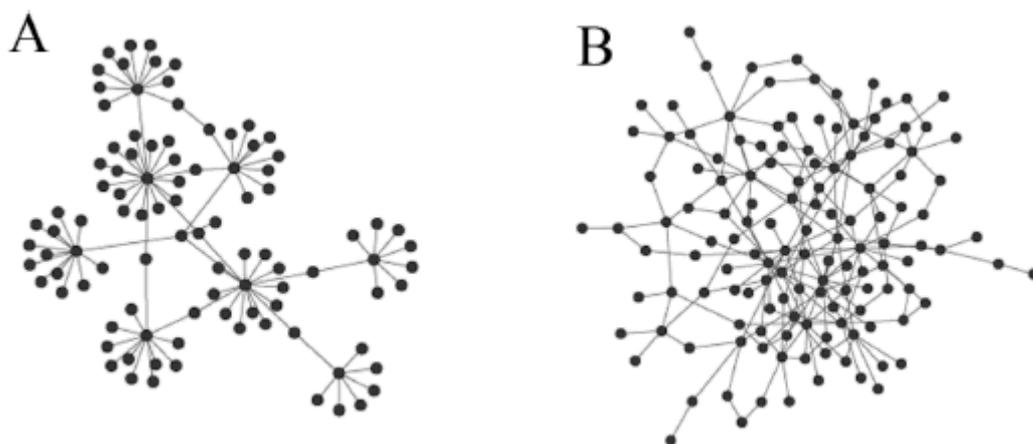


Рис. 4. Дисассортативная (А) и ассортативная (В) сети [НАО, Li, 2011].

6. *Оценка распределения степеней* позволяет исследователю понять, какой математической модели соответствует полученная им сеть. Для оценки, является ли полученная исследователем сеть сетью малого мира, производится сравнение ряда ее характеристик, таких как средний короткий путь и средний коэффициент кластеризации, с аналогичными характеристиками случайной сети, обладающей равным количеством узлов и связей. Сети малого мира характеризуются высоким (относительно случайной сети) средним коэффициентом кластеризации и сопоставимым средним коротким путем [Watts, Strogatz, 1998]. Распределение степеней в сетях малого мира может подчиняться экспоненциальному закону (*exponential degree distribution*) или степенному (*power law degree distribution*). В случае, если распределение степеней подчиняется степенному закону, сеть малого мира именуется безмасштабной сетью.

Структурные сетевые модели в когнитивных науках

В когнитивных науках количество работ, выполненных с использованием структурно- сетевого метода, растет в геометрической прогрессии (частичная библиография ссылок собрана в http://www.cs.upc.edu/~rferrericancholinguistic_and_cognitive_networks.html). Обзор всех направлений исследований не входит в задачи данной статьи. Мы изложим лишь принципы структурно-сетевого моделирования на примере сетевых моделей когнитивных процессов (операций обработки информации).

В русле этого направления структурный подход к сетевому моделированию тесно сближается с концептуальным подходом в ряде ключевых аспектов. Во-первых, крупнейшие достижения на данный момент сосредоточены в психолингвистике, как вначале произошло и с концептуально-сетевыми моделями. Во-вторых, наблюдается постепенная переориентация исследователей, использующих структурно-сетевой метод, с формальных на когнитивные аспекты языка. В-третьих, легко отметить преемственность в когнитивно-ориентированных структурных моделях основных понятий неассоцианизма, переосмысление их в контексте науки о сложных сетях. Некоторые термины вроде «распространения активации» и «порога извлечения» используются в неизменном виде, но несут в себе другое содержание. Структурный подход к сетевому моделированию сохраняет свой специфический взгляд на моделируемый объект. Центром внимания исследователя всегда является топология сети – визуализация данных, описывающих систему.

Структурно-сетевой метод органично приложим к языковым объектам: язык является сложной системой, состоящей из огромного количества лексических единиц. Эти единицы вступают в различные отношения (ассоциации) друг с другом, будь то отношения по смыслу, по составу (буквы, фонемы) или по структуре (синтаксис). Слова и их взаимосвязи образуют ассоциативные сети, которые эволюционируют в течение жизни языка, группы и отдельного человека. В лингвистическом контексте взаимодействие языковых единиц создает более высокие и сложные уровни лексической организации – предложение, текст и т.п. В когнитивном контексте взаимодействие психических репрезентаций лексических единиц порождает, будучи обусловлено психическими закономерностями,

мышление, чтение, восприятие и порождение речи и т.д.

Современная лингвистика содержит большие базы данных о различных аспектах языка, которые могут быть эффективно изучены структурно-сетевым методом. Перед началом моделирования исследователь решает, что будет принято им за узлы (буквы, слова, фонемы) и что будет определять наличие или отсутствие связи между узлами. В качестве массива данных, описывающих систему, принято использовать ряд источников. В первую очередь это словари и тезаурусы, содержащие формальные сведения о лексических единицах. Во-вторых, ученые проводят анализ текстов, выявляя, например, частотности совместной встречаемости. Для этого используются корпуса статей и сборники художественной литературы. Сеть, построенная на таких данных, отражает возможность одновременного появления слов в осмысленных предложениях конкретного языка. Наконец, данные могут быть собраны в экспериментах с живыми людьми.

Релевантные исследования, выполненные в русле психолингвистики, характеризуются разной степенью интереса к психической организации языкового знания. Как уже было упомянуто, на заре науки о сложных сетях наибольшее внимание уделялось формальным аспектам языка. Ученых интересовали следующие вопросы: что характеризует строение языка? Насколько универсальными являются эти характеристики? Каковы отпечатки эволюции в сложившейся структуре?

Попытки ответить на эти вопросы оказались продуктивны. Так, ученые выяснили, что большинство лингвистических сетей вне зависимости от языка являются сетями малого мира [Borge-Holthoef, Arenas, 2010]. Речь идет, к примеру, о базе WordNet (сеть гипонимов), корпусе ACE (сеть частотности совместной встречаемости), English Free Association Norms USF-FA (сеть свободных ассоциаций), тезаурусе Моби (сеть концептов), тезаурусе Рожета (сеть смысловых ассоциаций). Все они характеризуются маленькими короткими путями, высокой кластеризацией и ассортативностью. Часть из них (WordNet, ACE) также демонстрируют распределение степеней по степенному закону, то есть являются безмасштабными сетями. В обзоре 2010 года С.Строгатц и коллеги сравнили фонологические сети пяти языков мира: английского, испанского, китайского, баскского и гавайского [Arbesman et al., 2010]. Несмотря на частные отличия, все эти сети являются сетями малого мира и (кроме китайского) обладают распределением степеней, отвечающим усеченному степенному закону (truncated power law).

Понимание структурных особенностей сети позволяет исследователям выдвигать гипотезы об эволюционных механизмах, приведших к возникновению той или иной структуры, и делать компьютерные симуляции процессов роста сети [Strogatz, 2001]. Так, в работе Р.Феррер-и-Канчо и Р.Соле была построена сеть из 470 000 слов корпуса BNC (British National Corpus) [Ferrer-i-Cancho, 2003]. Распределение степеней в такой сети отвечает двум режимам степенного закона. Экспонента одного из них сходна с моделью А.-Л.Барабаси. Из-за наличия двух режимов авторы делают вывод, что лексикон разделен на группу ядерных слов и группу периферических слов. Ядерный лексикон содержит слова, известные всем представителям социальной группы носителей языка, а слова на периферии известны лишь их части. По мнению авторов, рост лексикона подчиняется предпочтительному связыванию, при котором сеть разрастается вокруг ядра, содержащего наиболее независимые от контекста слова.

В теориях, подобных теории Феррер-и-Канчо, можно заметить начало когнитивной трактовки лексических феноменов. Несмотря на свою формальность, описания процессов роста языка учитывают социальный контекст эволюции языка и индивидуальные условия обучения. Это смещение акцентов базируется на идее, что структурные особенности языка имеют глубокие психические причины и следствия. Ученые задают вопросы: как специфические (относительно систем другой природы) и универсальные (для языков мира) свойства языка обусловлены необходимостью обеспечивать познавательную работу человека? Как топологии языка и языковой памяти (ментального лексикона) обусловлены ограничениями когнитивной системы в переработке информации? Возможно ли описать принципы роста языковой памяти в течение жизни? Как структурные особенности языковой памяти (ментального лексикона) обуславливают распад и деформацию речи в патологии?

В современном структурно-сетевом подходе когнитивная система не сводится к сети. Сеть понимается как структура, на основе которой разворачивается когнитивная динамика. Когнитивная

система представляет собой единство структурного и динамического аспектов. Уточняя эту мысль, сформулируем три основных положения, принимаемые в когнитивных исследованиях на базе аппарата науки о сложных сетях. Эти положения принимаются вне зависимости от специфики изучаемого объекта [Baronchelli et al., 2013].

1. Когнитивные процессы рассматриваются как динамика, осуществляющаяся на субстрате сети (структурном аспекте системы).
2. Структура сети определяет характер динамики, осуществляющейся на ее основе.
3. Та или иная структура сети сложилась в результате эволюционного процесса и несет отпечаток функциональной необходимости.

Что представляет из себя структурный аспект когнитивной системы (первое положение)? В работах, выполненных в русле структурного подхода к сетевому моделированию когнитивной системы, для обозначения языковой памяти чаще всего используется понятие ментального лексикона [Vitevitch, 2008]. Ментальный лексикон является структурным аспектом когнитивной системы, а также прямым аналогом «ассоциативной структуры». Его понимают как структуру памяти человека, в которой в сетевой форме закодирована лексическая информация.

Конечно, попытки структурно- сетевого моделирования ментального лексикона угрожают частичным отрывом от ценностей и принципов, принимаемых психолингвистами в имплементации структурного подхода. Сила структурно- сетевого метода заключается в возможности детального анализа топологии реальной системы, представленной в образе сети. Сеть является визуализацией данных, описывающих систему. Чем точнее и полнее эти данные, тем корректнее визуализация и сильнее результат анализа. В таком контексте целесообразнее всего моделирование систем, данные о которых могут быть собраны непосредственно (результат МРТ, база данных о социальных контактах, входящие и исходящие пакеты информации, буквенное строение слов русского языка и т.д.). К сожалению, связи между психическими репрезентациями или языковая память подлежат лишь косвенному картированию, что накладывает на модель нежелательный отпечаток недостоверности и требует от ученого тщательного обоснования выбора для визуализации тех или иных баз данных.

Данные, которые ученый использует в моделировании, несут разную когнитивную нагрузку в зависимости от их источника. Так, можно предположить, что данные, получаемые из словарей, наиболее формальны и плохо отражают психическую реальность. Художественные тексты являются слепком письменной речи человека конкретного времени и культуры. В свою очередь данные, полученные в ассоциативных экспериментах, собраны путем организации и фиксации процесса навигации человека по его языковой памяти. Решение автора использовать те или иные данные зависит от стоящей перед ним проблемы и конкретной интерпретации связей между лексическими единицами. Общепринято представление, что единицей кодировки в ментальном лексиконе является слово. При этом организация знания подчиняется не только смысловым связям, но также орфографическим и фонологическим. На данном этапе сети, моделируемые различными авторами, чаще всего отражают один из уровней кодировки лексического знания. В этом контексте выбор баз данных, составленных на основе психологических экспериментов, не всегда является обязательным и даже необходимым.

Иллюстрацией этого утверждения может выступать, к примеру, фонологическая сеть английского языка, созданная в лаборатории М.Витевича в 2008 году [Vitevitch, 2008]. Узлами в такой сети являются транскрипции выборки из 20000 слов в Merriam-Webster Pocket Dictionary. Связь между словами устанавливается в случае их фонетического сходства. Определение фонетического сходства следует традиции, заложенной в работах П.Люче и Д.Писони. П.Люче и Д.Писони в серии психолингвистических экспериментов убедительно продемонстрировали, что шаг, который человек совершает в задаче на изменение услышанного слова, сводится к одной фонеме [Luce, Pisoni, 1998]. М.Витевич операционализировал этот принцип в алгоритме: связь устанавливается между двумя словами в фонологической сети, если одно слово образуется из другого путем замены, добавления или удаления одной фонемы. Таким образом, в приведенном примере визуализируемый массив данных формален (напрямую извлечен из словаря), однако сам принцип организации сети основан на принципе, полученном и проверенном в психолингвистических экспериментах. С определенной

долей условности такая мера позволяет принимать итоговую структуру сети как отражающую реальное строение уровня ментального лексикона, в котором лексические знания закодированы по фонетическому признаку.

Что представляет из себя динамический аспект когнитивной системы (первое положение)?

Используется понятие «навигация» в ментальном лексиконе, наиболее близкое к понятию «извлечение» в семантической памяти. Если структура ментального лексикона является прямым аналогом ассоциативной структуры, то «навигация» или «динамика» прямо описывается в терминах распространения активации, хотя механизмы распространения поставлены в подчеркнуто зависимое положение к структуре сети. Так, М.Витевич использует понятие активации для описания механизмов извлечения слов из фонологической сети / памяти [Vitevitch, Goldstein, 2014]. Его интерпретация использует представления о распространении активации А.Коллинза и Э.Лофтуса, а также идеи о пороге извлечения или реактивном потенциале, развитые еще в работах Ч.Кофера: при восприятии ключа активация возникает на определенном участке ассоциативной сети, содержащей схожие по составу слова, и распространяется от узла к узлу. Если на узел попало достаточно активации, он извлекается. Чем шире активация диффундирует по сети, тем меньше ее попадает на отдельные узлы, тем меньше вероятность их извлечения.

Как структура ментального лексикона влияет на динамику, осуществляющуюся на ее основе (второе положение)?

Именно в ответе на этот вопрос более всего раскрывается специфика структурного подхода относительно концептуального. Ученый, использующий структурно-сетевой подход, принимает решение о том, какой массив данных он будет визуализировать, при этом целью является не объяснительная модель, а именно наиболее точная визуализация реального строения системы. Когда сеть построена, исследователь использует аппарат статистического анализа, разработанный в науке о сложных сетях, для извлечения важных сведений о структуре реальной когнитивной системы. Для того чтобы понять, каким закономерностям подчиняется распространение активации по сети (динамика), автор выдвигает гипотезы о влиянии на него отдельных топологических свойств сети – и проверяет их в экспериментах с людьми (например, зависит ли сложность решения анаграмм от центральности исходных слов в орфографической сети).

Предположения о влиянии структуры на динамику также не являются свободными. Поскольку наука о сетях опирается на математические модели, постольку авторы руководствуются универсальными представлениями о преимуществах и ограничениях конкретной сетевой структуры. В итоге получается, что законы влияния топологии на динамику переносятся, например, из эпидемиологии в социологию, из социологии в психолингвистику.

Так, с когнитивистской точки зрения свойства малого мира сети необходимы для обеспечения эффективности когнитивной работы: высокая кластеризация собирает в единые кластеры сходные элементы информации; маленькие расстояния между единицами знания облегчают их поиск и извлечение; ассортативность обеспечивает устойчивость системы памяти к «забыванию» отдельных элементов и т.д. Длина коротких путей, к примеру, является важным условием эффективного извлечения информации: у детей с задержкой развития речи пути в семантической памяти длиннее нормы, а людей с болезнью Альцгеймера – короче [Borge-Holthoefner, Arenas, 2010]. Слова с высоким коэффициентом кластеризации в фонологической сети лучше запоминаются, распознаются и извлекаются, чем слова с низким коэффициентом кластеризации [Vitevitch et al., 2011].

Больше результатов когнитивных исследований, выполненных в русле структурного подхода к сетевому моделированию, можно найти, например, в обзорной статье А.Барончелли [Baronchelli et al., 2013].

Финансирование

Исследование выполнено при финансовой поддержке Российского научного фонда, проект 17-78-30035 «Психологические факторы экономической и социальной конкурентоспособности России».

Литература

- Anderson J.R. A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 1983, 22(3), 261–295.
- Arbesman S., Strogatz S.H., Vitevitch M.S. Comparative analysis of networks of phonologically similar words in English and Spanish. *Entropy*, 2010, Vol. 12, 327–337.
- Barabasi A.-L. *Linked: The new science of networks*. New York, NY: Plume, 2002.
- Barabasi A.-L. The network takeover. *Nature Physics*, 2012, Vol. 8, 14–16.
- Baronchelli A., Ferrer-i-Cancho R., Pastor-Satorras R., Chater N., Christiansen M.H. Networks in cognitive science. *Trends in Cognitive Science*, 2013, 17(7), 348–360.
- Borge-Holthoefer J., Arenas A. Semantic networks: structure and dynamics. *Entropy*, 2010, Vol. 12, 1264–1302.
- Carrington P.J., Scott J., Wasserman S. *Models and methods in social network analysis*. Cambridge: Cambridge University Press, 2015.
- Cofer Ch.N., Foley J.P. Mediated generalization and the interpretation of verbal behavior I Prolegomena. *Psychological Review*, 1942, 49(6), 513–540.
- Collins A.M., Loftus E.F. A spreading activation theory of semantic processing. *Psychological Review*, 1975(6), Vol. 82, 407–428.
- Ferrer-i-Cancho R., Sole R.V. Least effort and the origins of scaling in human language. *Proceedings of the National Academy of Sciences*, 2003, 100(3), 788–791.
- Freeman L. A set of measures of centrality based upon betweenness. *Sociometry*, 1977, 40(1), 35–41.
- Granovetter M.S. The strength of weak ties. *American Journal of Psychology*, 1973, 78(6), 1360–1380.
- Hao D., Li C. The Dichotomy in Degree Correlation of Biological Networks. *PLoS ONE*, 2011, 6(12), e28322. doi:10.1371/journal.pone.0028322
- Hinton G.E. Learning distributed representations of concepts. *Proceedings of 8th Annual conference of the Cognitive Science Society*, Amherst, 1986. pp. 84–119.
- Keeling M. The implications of network structure for epidemic dynamics. *Theoretical Population Biology*, 2005, 67(1), 1–8.
- Luce P.A., Pisoni D.B. Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 1998, 19(1), 1–36.
- Mitchell M. *Complexity: A guided tour*. New York, NY: Oxford University Press Inc, 2011.
- Milgram S. The Small World Problem. *Psychology Today*, 1967, 1(1), 61–67.
- Rumelhart D.E., McClelland J.L. On learning the past tenses of English words. In: *Parallel distributed processing*. Cambridge, Massachusetts: MIT Press, 1986. Vol. 2, pp. 110–146.
- Strogatz S.H. Exploring complex networks. *Nature*, 2001, Vol. 410, 268–276.
- Vitevitch M. What can graph theory tell us about Word Learning and Lexical Retrieval? *Journal of Speech, Language, and Hearing Research*, 2008, 51(2), 408–422.
- Vitevitch M.S., Ercal G., Adagarla B. Simulating retrieval from a highly clustered network: Implications for spoken word recognition. *Frontiers in Psychology*, 2011, Vol. 2, 369.

Vitevitch M.S., Goldstein R. Keywords in the mental lexicon. *Journal of Memory and Language*, 2014, Vol. 73, 131–147.

Watts D.J., Strogatz S.H. Collective dynamics of 'small-world' networks. *Nature*, 1998, Vol. 393, 409–410.

Поступила в редакцию 12 августа 2017 г. Дата публикации: 29 октября 2017 г.

[Сведения об авторе](#)

Морозова Ольга Андреевна. Аспирант, Институт психологии Российской академии наук, ул. Ярославская, д. 13, 129366 Москва, Россия.
E-mail: helgamoro@gmail.com

[Ссылка для цитирования](#)

Стиль psystudy.ru

Морозова О.А. Структурное сетевое моделирование в когнитивной науке. *Психологические исследования*, 2017, 10(55), 1. <http://psystudy.ru>

Стиль ГОСТ

Морозова О.А. Структурное сетевое моделирование в когнитивной науке // *Психологические исследования*. 2017. Т. 10, № 55. С. 1. URL: <http://psystudy.ru> (дата обращения: чч.мм.гггг).

[Описание соответствует ГОСТ Р 7.0.5-2008 "Библиографическая ссылка". Дата обращения в формате "число-месяц-год = чч.мм.гггг" – дата, когда читатель обращался к документу и он был доступен.]

Адрес статьи: <http://psystudy.ru/index.php/num/2017v10n55/1474-morozova55.html>

[К началу страницы >>](#)