Синтез эмерджентных сетей разбиений

Изучается специальный класс (m^2, m) разбиений (триномиальное семейство сетей) с заданной метрикой в пространстве разбиений. Анализируется топология неориентированных графов этого семейства, в которых вершинами являются разбиения, а наличие ребра определяется расстоянием между вершинами (разбиениями). Ряд топологических параметров таких графов однозначно определяется числами треугольника Паскаля. Численное моделирование для m= 3,4,…,9 и предложенный метод синтеза дают основание считать, что для любого m рассмотренное семейство организовано как объединение трех классов совершенных графов: клик, хордовых графов и двудольных графов. Экспериментально изучено поведение топологических характеристик графов триномиальных семейств при произвольном удалении клик графа и отмечен эмерджентный характер устойчивости топологических характеристик рассмотренных семейств графов.

Введение.

Сложными сетями принято называть обширное поле исследований различных моделей сетей – в первую очередь биологических, а также компьютерных, информационных, транспортных, энергетических. Одна из причин всплеска интереса к подобным моделям – доступность огромных баз данных, в которых можно найти *row data* обо всех мыслимых областях деятельности.

В такой обширной области неизбежно и довольно быстро возникает необходимость в поиске общих закономерностей, справедливых для изучаемого объекта, независимо от частных деталей таких как принципы функционирования элементов сети и/или структуры сети. Уже достаточно давно замечено[1], что архитектура различных сложных сетей имеет между собой много общего: модульную структуру, наличие повторяющихся топологических структур (motif), «близких» законов вероятностных распределений параметров сети, эффект «тесного мира» (*small world*) и масштабируемости (*scale-free*). Более того, отмечалось удивительное сходство некоторых топологических параметров случайных сетей с реальными сложными сетями. Следует, однако, учесть [2], что между искусственно созданными системами (даже с таким числом элементов, как в современных чипах или в Интернете) и биологическими системами существует кардинальное отличие: для биологических систем все функции, которыми они обладают, подчинены двум главным целям: выживанию и репродукции. Потому в русле таких исследований как генетические сети [4] и

DNA-computing [3] важно изучение методов синтеза сетей с упором на реализацию механизмов наследования (передачи наследственной информации), а также эффекты эмердженси[1] (emergence) [5] известные в эволюционной биологии.

Анализ поведения сложных сетей в общем случае не позволяет конструктивно расшифровать их структуру. Известно, например, [6], что, если приходится иметь дело с «абсолютным черным ящиком», т.е. автоматом, для которого неизвестна оценка числа состояний, задача расшифровки (восстановления автоматной диаграммы) является алгоритмически неразрешимой даже для «почти всех» таких автоматов. Но и в том случае, когда известна верхняя оценка числа состояний автомата, трудоемкость задачи его расшифровки оказывается в общем случае не меньше по порядку числа его состояний[2].

Особенности функционирования сложного объекта могут быть связаны с конкретными условиями процесса его создания. Простейший пример – выбор топологии интегральных схем. Зависимость между целями и структурой наглядно видна на примере биологических, в частности генетических сетей. Две главные цели живого – выживание и репродукция – диктуют такие базовые особенности конструкции и функционирования генетических сетей как стабильность и вариабельность.

В работах последнего десятилетия, особенно тех, которые посвящены анализу генетических сетей, рассматриваются различные методы, позволяющие на основании экспериментальных данных строить формальные моделей. Обычно, за основу берутся статические и динамические Байесовские сети, булевы сети, авторегрессионные модели, которые представляют в виде направленных и ненаправленных графов с вершинами, соответствующими активным элементам (генам, молекулам), а ребра – возможностям взаимодействия между этими элементами. В некоторых случаях, когда математическая модель представляет собой граф, появляется возможность оценивать топологическую «близость» модели и объекта: поведение диаметра графа, распределения средних длин кратчайших путей, размерности клик и распределения их числа, величин кластерного коэффициента и индекса плотности матрицы связности. Исследование топологии графов сложных сетей часто играет роль *benchmarking analysis*, зачастую позволяя приходить к нетривиальным гипотезам.

[1] Emergence – в теории систем, физике и ряде других областей этим термином принято обозначать такие ситуации, при которых «целое становится больше, чем просто сумма частей», точнее говоря, появление таких свойств, которые как будто бы появляются ниоткуда (трудно найти их причину, анализируя отдельные элементы целого)

[2] Эти факты справедливы лишь для «почти всех» сетей. Для расшифровки (исправного) оперативного ЗУ достаточно порядка логарифма числа его состояний.

Цель настоящей работы – изучение формальной модели класса сложных сетей, элементами которых являются разбиения и клики разбиений. Такая модель, с нашей точки зрения, может помочь пониманию того, что построение структуры (сети), обладающей свойствами «наследования», робастности и вариабельности можно осуществить на базе чрезвычайно простого (с позиций сложности по Колмогорову) исходного «элемента». Более того, такая сеть явно демонстрирует эффект emergence при анализе поведения в случае удаления составляющих её элементов.

Большинство моделей сложных сетей опирается на неявное, но кажущееся очевидным, предположение: установление связей каждого элемента сети с соседями определяется некими *внешними* обстоятельствами. Таким внешним фактором в социальных сетях является топология уже существующей сети (новая вершина присоединяется к самым популярным вершинам), в технических сетях – представление создателя этой сети о том, как элементы сети *должны быть* связаны друг с другом. В тоже время установление связей в биологических сетях есть внутреннее (*intrinsic*) качество элементов этих сетей, что наблюдается в экспериментах по экспрессии генов [ ] и, возможно, именно это качество может быть причиной эффектов эмердженси.

Предлагаемый в работе элемент это формальная модели некоторой «универсальной клетки»[3] основными свойствами которой являются внутренняя способность образовывать связи с другими «близкими» по свойствам клетками. Формальная модель этой «универсальной клетки» достаточно проста, что уменьшает трудоемкость числового моделирования, но, одновременно, позволяет на феноменологическом уровне рассмотреть вопросы роста и наследования. Наш подход к изучению сложных сетей нацелен на изучение некоторых общих закономерностей робастного поведения сетей специального класса в процессе эволюции и, прежде всего, обеспечения «выживания» при повреждениях структуры. Как и в большинстве реальных биологических объектов такое поведение основано на возможностях изменения топологии сети в процессе передачи ресурсов (сигналов), когда имевшиеся ранее пути доставки оказываются недоступными.

Дальнейшее изложение построено следующим образом.

В разделе 1 «Модель p-клетки» предложена модель элемента сети в виде разбиений из заданного класса и предложена метрика в пространстве разбиений. Раздел 2 «Принцип наследования для p-клеток» посвящен механизму передаче свойств от поколения к поколению. Раздел 3 «Семейства разбиений $F(\lambda_0(m))$ их топология» содержит итеративный

---

[3] На биологическом языке клетки способные делиться и создавать *любые* клетки организма данного вида называются *totipotency* (наиболее известный случай – споры).

способ синтеза триномиальных семейств разбиений, топологические свойства которых изучаются в дальнейшем. В разделе 4 «Робастность семейств $F(\lambda_0(m))$» приведены результаты численных экспериментов, которые демонстрируют эмерджентные свойства сетей клик: сохранение таких топологических параметров, как диаметр и кластерный коэффициент при удалении клик сети за счет перестройки топологии. «Заключение» посвящено обсуждению дальнейших исследований.

1. Модель р-клетки

Модель сети, рассмотренная в работе, основана на триномиальном семействе разбиений (некоторые сведения о разбиениях см. [7],[8],[9]). Элементы этой модели – разбиения из заданного класса – везде далее будем называть р-клетки

Определение 1. Пусть (n, m) – разбиение целого числа n (ресурс разбиения) не более чем на m целых частей – последовательность неотрицательных целых чисел $a_1 \geq a_2 \geq \cdot \cdot \cdot \geq a_m$, такая, что $n = a_1 + a_2 + \cdot \cdot \cdot + a_m$. Множество всех (n, m) - разбиений не более чем на m частей на обозначим $P(n, m)$. Пусть $\boldsymbol{n}$ множество значений ресурса и $\boldsymbol{m}$ множество значений числа частей. Тогда через $P(\boldsymbol{n}, \boldsymbol{m})$ будет обозначаться множество разбиений с различными значениями n, m, где $n \in \boldsymbol{n}$, $m \in \boldsymbol{m}$.

Определение 2. Определим расстояние ρ между парой разбиений α, $\boldsymbol{\beta} \in P(n, m)$ и $\boldsymbol{\alpha} \neq \boldsymbol{\beta}$ как

$$\rho(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \max |a_i - \beta_i|, \qquad (1)$$

где максимум берется по всем i ($i = 1,\dots, m$).

Если $\boldsymbol{\alpha}$ и $\boldsymbol{\beta}$ имеют различное число частей будем предполагать, что значение всех недостающих частей в меньшем по числу частей разбиении равно 0.

Для всех $\boldsymbol{\alpha} \neq \boldsymbol{\beta}$ справедливы неравенства,

$$1 \leq \rho(\boldsymbol{\alpha}, \boldsymbol{\beta}) \leq n(1 - 1/m)$$

(верхняя граница справедлива при $\boldsymbol{\alpha} = \{0, 0, \dots, n\}$ и $\boldsymbol{\beta} = \{ n/m, n/m, \dots, n/m\}$)

Везде дальше любое разбиение $\boldsymbol{\alpha} \in P(n, m)$ будем называть *р-клеткой*, если расстояние в $P(n,m)$ заданы в соответствии с Определением 2.

Определение 3. Для разбиения $\boldsymbol{\alpha} = (\alpha_1 , \dots, \alpha_m)$ просвет (*gap*) определяется как $k = \alpha_j - \alpha_{j+1}$ ($j =1,2,\dots, m-1$) [ ]. Пусть для некоторого разбиения $\boldsymbol{\alpha_0} = (\alpha_1 , \dots, \alpha_m) \in P(n,m)$ выполняется следующее: для любого α_j , α_{j+1} ($1\leq j \leq m-1$) величина просвета $k \geq 2$ одинакова . Такие разбиения будем называть *разбиениями с равномерным просветом* (РП(k)).

В дальнейшем в основном будет изучаться класс разбиений, в котором $n(m) = m^2 +tm$, где $t = 0,1,2,\dots$, который будет обозначаться как $P_t(n(m))$.

Множества разбиений $P_0(n(m)) \supset P(n, m)$ и $P_1(n(m)) \supset P(n, m)$, где $m \geq 3$, содержат одно РП(2) для каждого m. Это последовательность нечетных чисел $\boldsymbol{\lambda_0}(m) = 2m - 1, 2m-3, \ldots, 3, 1$ для $P_0(m)$ и последовательность четных чисел $\boldsymbol{\lambda_1}(m) = 2m, 2m-2, \ldots 4, 2$ для $P_1(m)$.

Определение 4. Для любого $\boldsymbol{\alpha} \in P(n, m)$ назовем *средним просветом*

$$g^*(\boldsymbol{\alpha}) = 1/(m-1) \sum_{i=1} (\alpha_i - \alpha_{i+1})$$

Для $g^*(\boldsymbol{\alpha})$ имеют место оценки

$$0 \leq g^*(\boldsymbol{\alpha}(m)) \leq n/(m-1)$$

2. Принцип наследования для р-клеток

Укажем вначале вид структуры «генотипа», т.е. первого поколения, и сформулируем закономерности перехода от поколения к поколению.

Определение 5. Назовем Р-клетку $\boldsymbol{\alpha} \in P(n, m)$ главой (*head*) множества разбиений (семьи) F_α, если F_α содержит все разбиения $\boldsymbol{\beta} \in P(n, m)$ такие, что $\rho(\boldsymbol{\alpha}, \boldsymbol{\beta}) = 1$. Семью F_α будем представлять графом $G(F_\alpha) = (V(P(n,m)),E)$, где V – множество вершин (р-элементов семьи F_α) и E – множество ненаправленных ребер $e(v_i, v_j)$, если $\rho(v_i, v_j) = 1$. Номер поколения h любой семьи F_α равен m-3, т.е. нулевым поколение будет семья из множества разбиений P(3,9). В качестве примера на рис. 1 приведена матрица связности для графа $G(F_\alpha) = (V(P(16,4)),E)$ и $\rho(v_i, v_j) = 1$.

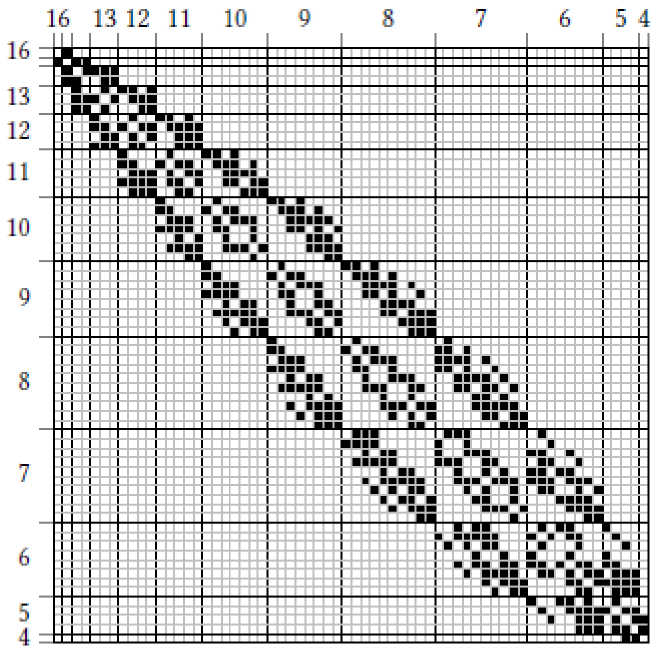

Рис.1 Матрица связности для графа $G(F_\alpha) = (V(P(16,4)),E)$ и $\rho(v_i, v_j) = 1$.

Рассмотрим эволюцию семейств F_α, т.е. процесс роста семьи при увеличении параметра *m*, особо отмечая самоповторяемость в процессе эволюции.

<u>Построение нулевого поколения.</u>

Главой семьи нулевого поколения является p-клетка $\lambda_0(3)$, которую будем называть *зародышевой.* Множество разбиений в этой семье обозначим как $F(\lambda_0(3))$. Будем рассматривать три непересекающихся подмножеств этой семьи, т.е. $F(\lambda_0(3)) = F_3^l \cup F_3^c \cup F_3^r$, где подмножество F_3^c содержит все разбиения у которых $a_1 = \lambda_1$, подмножество F_3^l разбиения у которых $a_1 = \lambda_1 - 1$ и F_3^r – разбиения у которых $a_1 = \lambda_1 + 1$

Пример 1. На рис. 2 показан граф $G(F_\alpha) = (V(P(9,3)),E)$, который содержит 12 разбиений: (9; 8,1; 7,2; 7,1,1; 6,3,0; 6,2,1; 5,4; 5,3,1; 5,2,2; 4,4,1; 4,3,2; 3,3,3). Граф семьи $F(\lambda_0(3))$ = (6,3,0; 6,2,1; 5,4,0; 5,3,1; 5,2,2; 4,4,1; 4,3,2) показан на рис.2, а на рис.3 его матрица связности. Кластерный коэффициент для графа семьи равен 0,465, а диаметр – 6. Красным обозначены вершины семьи $F(\lambda_0(3))$ подмножество вершин $V(P(9,3))$, которое представляет собой гексагон (группа симметрии 6n). «Зародыш» – разбиение [5,3,1] является ядром графа семьи $F(\lambda_0(3))$, который состоит из 6 клик размерности 3. Кластерный коэффициент графа равен 0,545, а диаметр равен 2. Матрица связности для этого графа показана на рис.2. Семья $F(\lambda_0(3))$ является подмножеством класса P(9,3),. Таким образом, семья $F(\lambda_0(3))$ является симметричным, плотным подмножеством класса разбиений P(9,3) из которого она происходит.

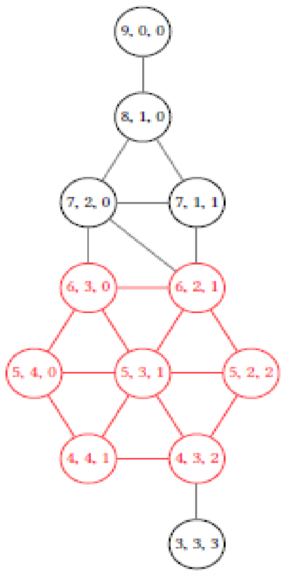

Рис.2. Граф $G(F_\alpha) = (V(P(9,3)),E)$,красные вершины принадлежат графу семьи $F(\lambda_0(3))$

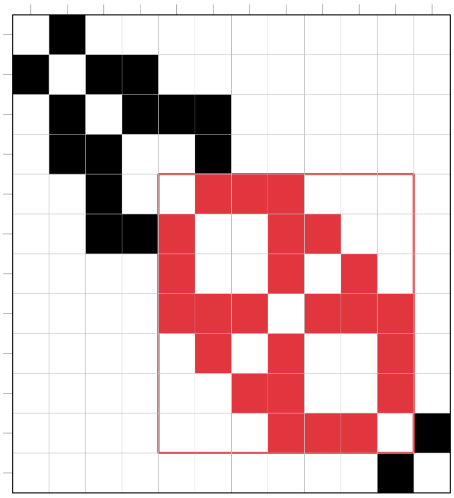

Рис. 3. Матрица связности графа $G(F_\alpha) = (V(P(9,3)),E)$, где красным обозначено наличие вершин семьи $F(\lambda_0(3))$

Замечание. Если в качестве зародыша будет выбрана клетка $\lambda_1(3)$, то семья $F(\lambda^1(3))$ будет состоять из разбиений: (6,4,2; 7,4,1;7,3,2; 6,5,1; 6,3,3; 5,5,2; 5,4,3) а графы семей $F(\lambda_0(3))$ и $F(\lambda_1(3))$ будут изоморфны.

Каждое поколение семей построено с помощью следующей итеративной процедуры.

Шаг 0. Построение нулевого поколения семьи $F(\lambda_0(3))$

Шаг 1. Построение семьи $F(\lambda_0(m+1)) = F_{m+1}{}^{l} \cup F_{m+1}{}^{c} \cup F_{m+1}{}^{r}$

1.1. Множество $F_{m+1}{}^{c}$

К каждому разбиению из $F_m{}^{c}$ в качестве новой первой части добавляем число 2m +1, а все остальные части разбиений остаются без изменений.

1.2. Множества $F_{m+1}{}^{l}$, $F_{m+1}{}^{r}$

Для каждого разбиения $\boldsymbol{\alpha} = (a_1, a_2, \cdot\cdot\cdot, a_{m+1}) \in F_{m+1}{}^{r}$ $a_1 = 2m+2$ и $\rho(\boldsymbol{\alpha}, \underline{\boldsymbol{\lambda}_0(m+1)}) = 1$

Для каждого разбиения $\boldsymbol{\alpha} = (a_1, a_2, \cdot\cdot\cdot, a_{m+1}) \in F_{m+1}{}^{l}$ $a_1 = 2m$ и $\rho(\boldsymbol{\alpha}, \underline{\boldsymbol{\lambda}_0(m+1)}) = 1$

Пример 2. Рассмотрим процедуру, приведенную выше, на примере перехода от семьи $F(\lambda_0(3))$ к семье $F(\lambda_0(4))$. В таблицах 1а и 1б показаны семьи $F(\lambda_0(3))$ и $F(\lambda_0(4))$, соответственно, а на рис.4 матрица связности графа $G(F_\alpha) = (V(P(16,4)),E$

	531	
630	*540*	*441*
621	*522*	*432*

Табл. 1а

	7531	
6640	7630	8620
6631	7621	8530
6622	7540	8521
6532	7522	8440
6541	7432	8431
6442	7441	8422

Табл.1 б.

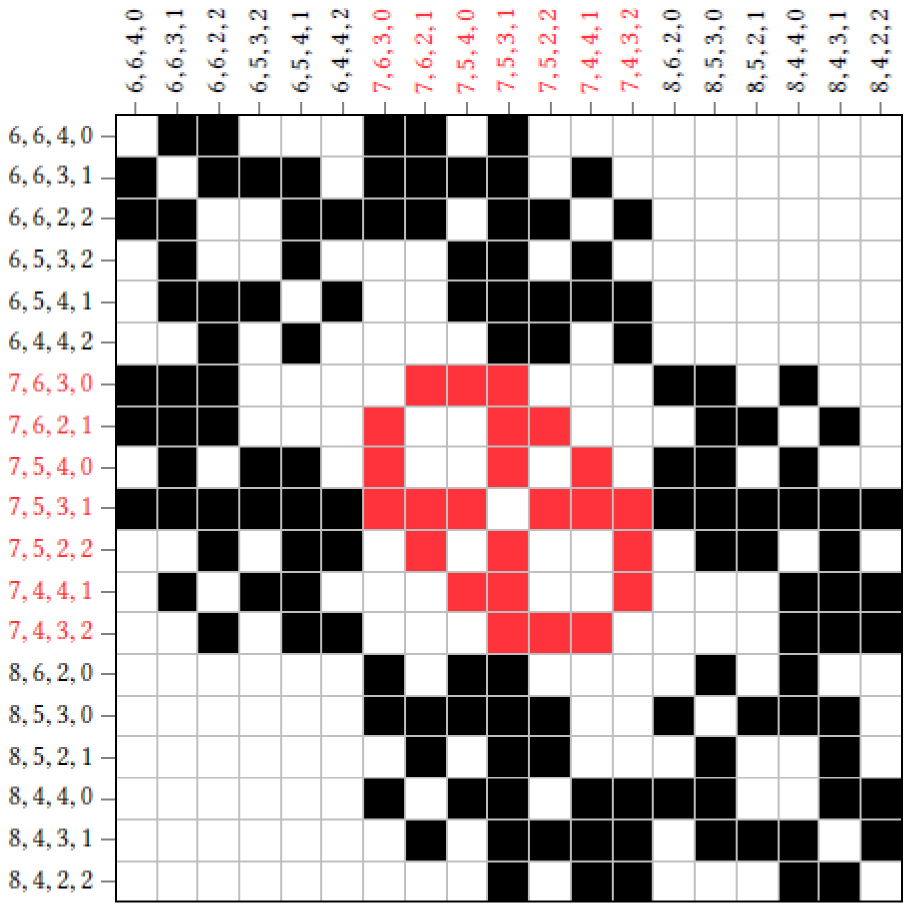

Рис.4. Матрица связности графа $G(F_\alpha) = (V(P(16,4)),E)$,где красным обозначены вершины семьи $F(\lambda_0(3))$

Центральная часть матрицы рис. 4 в точности повторяет матрицу связности графа семьи $F(\lambda_0(3))$. Итеративный характер процедуры наследования позволяет ожидать достаточно высокой предсказуемости топологических параметров семейств разбиений. В частности, «эффект матрешки», который наблюдается при сравнении матриц связности нулевого и первого поколения, повторяется для последующих поколений.

3. Семейства разбиений $F(\lambda_0(m))$ их топология.

Реализация итеративной процедуры построения семейств разбиений $F(\lambda_0(m))$ требует на каждом шаге $m \geq 3$ поиска всех разбиений, которые находятся на расстоянии 1 от $\lambda_0(m)$. Хотя проверка выполнения этого условия крайне проста, количество разбиенийпретендентов растет экспоненциально с ростом *n*. Выбор метрики (1) в пространстве разбиений дает возможность получить точные значения мощности множеств F_m^l, F_m^c, F_m^r для любых значений m. Заметим, что при заданных значениях *n* и *m* два разбиения α , β являются соседними – находятся на расстоянии 1 друг от друга – тогда и только тогда, когда разность $\alpha - \beta = (\alpha_1 - \beta_1, \ldots, \alpha_m - \beta_m)$ является набором из *m* символов (1,0,-1) сумма которых равна 0. Число таких соседних разбиений равно количеству перестановок символов (1,0,-1) при условии что сумма равна 0. Это в точности равно значениям триномиальных коэффициентов [10],[11] т.е. коэффициентов триномиального треугольника [4] (см. рис. 5)

				1				
			1	1	1			
		1	2	3	2	1		
	1	3	6	7	6	3	1	
1	4	10	16	19	16	10	4	1

Рис. 5. Первые 5 строк триномиального треугольника

Последовательность центральных триномиальных коэффициентов (k = 0) совпадает с последовательностью $|F_m^c|$ и для m = 3, 4, 5, 6, 7, 8, 9, 10…задается, соответственно, последовательностями 7, 19, 51, 141, 393, 1107, 3139, 8953,… (последовательность A002426 в OEIS [9]) и для $|F_{m+1}^l| = |F_{m+1}^r|$ последовательностью 6, 16, 45, 126, 357, 1016, 2907, 8350,…(последовательность A005717 in the OEIS). Известна также асимптотика для центрального триномиального коэффициента $a(m) \sim d\cdot 3^m / (m)^{1/2}$, где константа $d = \mathrm{sqrt}(3/\pi)/2$ Принцип построения семейств разбиений $F(\lambda_0(m))$ позволяет точно указать мощности множеств семейств F_m^l, F_m^c, F_m^r. Анализ топологии сетей, получившихся в результате реализации этого принципа, можно осуществить стандартным способом: изучая граф сети $G(V(F(\lambda_0(m)), E_1)$. В работе проведено численное моделирование [12] топологических параметров семейств $F(\lambda_0(m))$ для m= 3,…, 9, которое позволило установить значения этих параметров, в том числе: распределение степеней вершин, распределение клик, кластерный коэффициент, диаметр, плотность графа (отношение числа ребер к числу вершин), J-расстояния между кликами и ряд других[5]. Анализ этих результатов показал, что семейства разбиений $F(\lambda_0(m))$ обладают практически *полностью прогнозируемой топологической структурой.* В частности, любая семья состоит из набора клик (полных подграфов),

---

[4] Следуя обозначениям [10] триномиальный коэффициент $\binom{m}{k}_2$ где $m \geq 0$ и $-m \leq k \leq m$, задается коэффициентом при x^{m+k} в разложении $(1 + x + x^2)^m$.

[5] Поскольку числа вершин и их степени в изучаемых семействах растут экспоненциально, моделирование при значениях m > 9 требует вычислительных ресурсов недоступных для авторов

причем число клик и их размерности однозначно определяются параметром m и соответствуют числам треугольника Паскаля. В Таблице 2 показаны результаты численного моделирования, относящиеся к числам и размерностям клик. Из этой таблицы следует, что (по крайней мере, для m = 3,…,9) размерность клик (число p-клеток) задается первыми $\lceil m/2 \rceil$ значениями таблицы Паскаля, а число этих клик из столбца C(m,i) равно C(m,i) + C(m,m-*i*), если *i* нечетное и C(m,*i*), если *i* четно.

В Таблице 3 показаны результаты, относящиеся к распределению степеней вершин изучаемых семейств. Это дискретное распределение количество значений в котором равно $\lfloor m/2 \rfloor + 1$, а частоты встречаемости вершин с конкретной степенью примерно одинаковы для всех значений m из Таблицы 3 (см. рис. 4). Причины такого единообразного поведения, конечно, в итеративном характере построения сети при увеличении m.
Необходимо также отметить степенное с ростом m падение индекса плотности $I(G)=2|E|/(|V|)(|V-1|)$ графа $G(V(F(\lambda_0(m)), E_1)$, т.е. при увеличении размеров семейство $F(\lambda_0(m))$ становится все более разреженным, несмотря на увеличение количества и размерности клик, входящих в это семейство. Однако разреженность графа семейства не сказывается на величине его диаметра, который остается равен 2 (или 3 при исключении p-клетки главы семьи) при любом из рассмотренных m.

m	|V(m)|	КЛИКИ	C(m,1)	C(m,2)	C(m,3)	C(m,4)	C(m,5)	C(m,6)	C(m,7)	C(m,8)	Средняя частота встречаемости элементов
3	7	**размер**	**3**								**2,57**
		число	**3**	**3**							
4	19	**размер**	**4**	**6**							**3,57**
		число	**4**	**6**	**4**						
5	51	**размер**	**5**	**10**	**10**						**4,9**
		число	**5**	**10**	**10**	**5**					
6	141	**размер**	**6**	**15**	**20**						**6,53**
		число	**6**	**15**	**20**	**15**	**6**				
7	393	**размер**	**7**	**21**	**35**						**8,72**
		число	**7**	**21**	**35**	**35**	**21**	**7**			
8	1107	**размер**	**8**	**28**	**56**	**70**					**11.6**
		число	**8**	**28**	**56**	**70**	**56**	**28**	**8**		
9	3139	**размер**	**9**	**36**	**84**	**126**					**15,48**
		число	**9**	**36**	**84**	**126**	**126**	**84**	**36**	**9**	

Табл.2. Числа клик различных размерностей (для размерностей клик коричневым цветом показаны натуральные числа, красным – триангулярные числа, зеленым – тетраэдральные и синим – пентатопные числа). В последнем столбце под средней частотой встречаемости элементов следует понимать отношение суммы произведений числа клик на их размер к числу вершин |V| для каждого m.Причина того, что эта величина больше 1 состоит в том, что пересечение множества элементов в кликах ненулевое.

m	|V(m)|	Степени вершин	i(1)	i(2)	i(3)	i(4)	i(5)	I(G)=2|E|/(|V|)(|V-1|)	Cl.coef .
3	7	степень	3	6				0,57	0,545
		#	6	1					
4	19	степень	5	9	18				0,595
		#			1				
5	51	степень	13	25	50			0,36	0,564
		#			1				
6	141	степень	19	35	69	140		0,27	0,522
		#	20	90	30	1			
7	393	степень	49	95	191	392		0,22	0,471
		#	140	210	42	1			
8	1107	степень	69	132	261	533	1106	0,17	0,422
		#					1		
9	3139	степень	181	357	752	1499	3138	0,13	0,375
		#					1		

Табл.3 Распределение степеней вершин (красным выделено значение степени для главы семейства), значение индекса плотности и величина кластерного коэффициента

3.1. Граф клик семейств $F(\lambda_0(m))$

Из Таблицы 2 видно, что доля клик различного размера чрезвычайно велика в структуре графа $G(V(F(\lambda_0(m)), E_1)$. Перейдем от модели графа $G(V(F(\lambda_0(m)), E_1)$ к более обобщенной графовой модели семейства $F(\lambda_0(m))$, в которой для каждого $m \geq 4$ будем рассматривать графы клик $GQ\ (V_m, E(j))$, где вершинами графа будет множество клик Q графа $G(V(F(\lambda_0(m)), E_1)$ и для каждой пары клик $q_k, q_l \in Q$ ребро между кликами $e(q_k\, q_l) \in E(j)$ имеет место только когда выполняются условия для расстояний между этими кликами. В качестве такого расстояния при моделировании вычислялось J-расстояние d между кликами Q_i, Q_j как (см., например, [13])

$$d(Q_i, Q_j) = 1 - |Q_i \cap Q_j| / |Q_i \cup Q_j| \qquad (2)$$

Таким образом, для графа $GQ(V_m, E(j))$, где $|V_m| = \sum_{j\,(m)} n_j$ – сумма клик всех размерностей из Табл.2, а $e(q_k\, q_l) \in E(j)$, если $1 > d(Qi, Qj) > 0$.

Кроме того, для всех m будем рассматривать все графы клик без разбиения, которое является ядром и имеет максимальную степень $|V_m| - 1$, что на единицу уменьшит размер каждой клики.

4. Эмерджентная робастность семейств $F(\lambda_0(m))$

Экспериментальное изучение поведения графа клик $GQ\ (V_m, E(j))$ при произвольном удалении клик из множества вершин графа $GQ(V_m, E(j))$ приводит к ситуации, которую можно назвать эмерджентной робастностью.

Рассмотрим предварительно модель сети в виде связного неориентированного графа $G(V,E)$, где каждый элемент $\alpha \in V$ может находиться в двух состояниях: «свободен»

(1) и «занят» (0). Предположим, что путь в сети может быть проведен только через свободные элементы. Пусть в начальный момент времени все элементы сети свободны и диаметр сети равен D_0. В момент $t > 0$ с вероятностью $\varepsilon > 0$ какой-то элемент $\boldsymbol{\beta \in V}$ может перейти в состояние 0, в котором и остается неопределенное время. Очевидно, для графа $G_t(V/\beta, E)$ справедливо $D_t \geq D_0$. Будем говорить, что исходный граф сети деградирует, по мере того как его диаметр увеличивается. Максимальный уровень деградации соответствует потери связности. (Естественно, при вычислении диаметра деградирующего графа занятые элементам не рассматривают ни в качестве промежуточных, ни в качестве финальных). Тривиальным примером *минимально деградирующей* сети является сеть, которая представляет собой полный граф (клику). При любом числе занятых элементов, которое меньше размера клики, ее диаметр остается равным 1. Известна и цена такой устойчивости связности: плотность полного графа равна 1, т.е. степень каждой из N вершин равна N-1. При синтезе реальных больших сетей, содержащих десятки тысяч или более элементов, физически нереально обеспечить единичную плотность сети. Потому представляет интерес синтез сетей, которые могут достаточно долго оставаться связными при увеличении числа занятых элементов, но при этом максимальная степень элементов остается заметно меньше общего числа элементов сети.

Численное моделирование поведения графов триномиальных сетей разбиений при удалении клик из множества Q вершин графа клик демонстрирует эмерджентную робастность топологии графа клик, прежде всего, сохранение величины диаметра и кластерного коэффициента, хотя для всех промоделированных случаев (вообще говоря, для произвольных m), максимальная степень элементов графа более чем в два раза ниже величины $|V_m| - 1$ (см. Таблицу 3), а средняя величина степени вершин падает с ростом m и при m=9 составляет всего 14% от максимальной. Тем не менее граф клик остается связным при значительном (до 50%) удалении числа клик. Более того, диаметр графов с удаленными кликами может достаточно долго не увеличиваться, также, как и кластерный коэффициент, т.е. рассматриваемые семейства совершенно не желают деградировать.

Подробное описание полученных результатов требует отдельной публикации. Здесь же в качестве примеров будут приведем некоторые результаты для семейства $F(\lambda_0(5))$.

Пример 3.

На рисунках показаны графы клик семейства $F(\lambda_0(5))$, где у каждой вершины стоит номер клики и ее размер, в Таблице 4 приведена последовательность удаляемых клик и вид полученных результатов.

Шаг	Удаленные клики	Клики графа (размер-количе-ство)	D	№ рис
0	0 (исходный граф)	4-10, 9-20	3	6
1	C4-1 – C4-4	3-2, 4-4, 5-8, 7-4, 2-9	3	7
2	C4-1 – C4-4, C9-1 – C9-4	2-3 , 5- 4, 2-5	2	8
3	C4-1 – C4-4,C9-1 – C9-4, C4-5	6-5	2	9
4	C4-1 – C4-4,C9-1 – C9-4, C4-5, C9-5	3-3	2	
5	C4-1 – C4-4,C9-1 – C9-4, C4-5, C9-5 ,C9-6	3-3	2	
6	C4-1 – C4-4,C9-1 – C9-4, C4-5, C9-5 ,C9-6,C9-7		∝	

Таблица 4. Параметры эксперимента по удалению кик в семействе $F(\lambda_0(5))$.

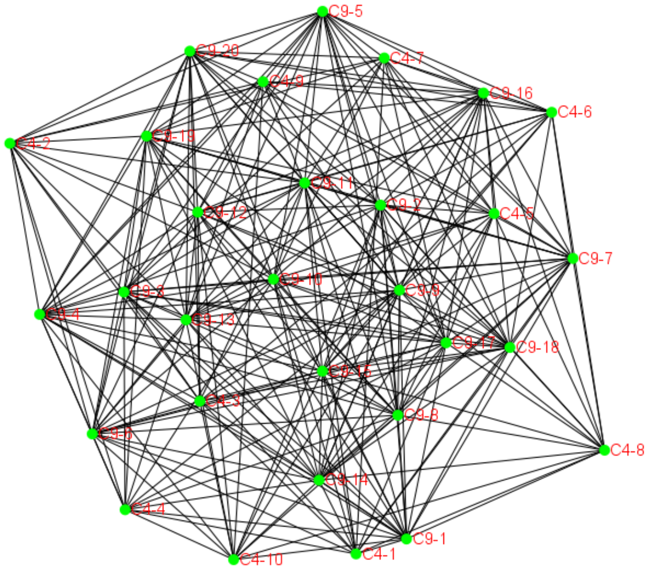

Рис.6

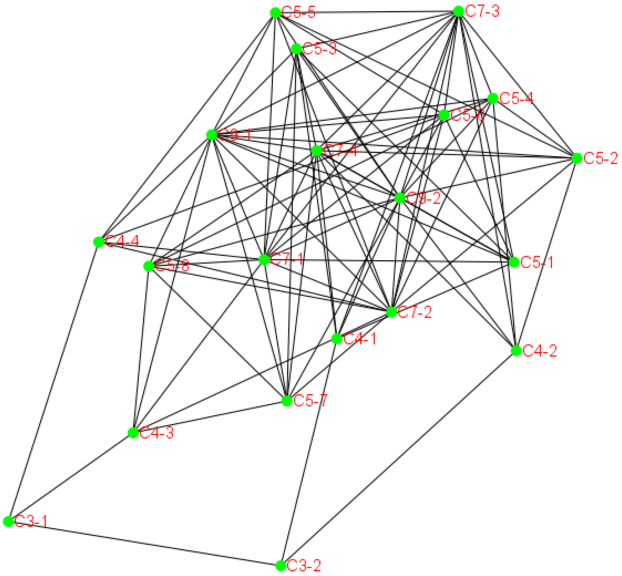

Рис.7.

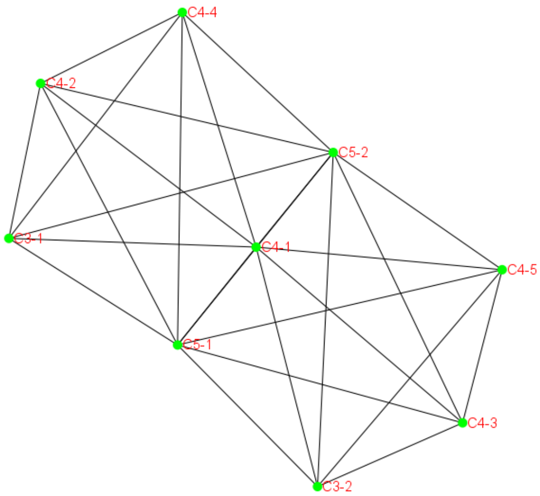

Рис. 8

Комментарий. После первого удаления четырех клик размера 4 граф становится намного «стройней» – вместо 30 исходных клик осталось 27, но зато разнообразие их размеров выросло в 3 раза. Если удалить кроме четырех клик размера 4 еще четыре клики размера 9 граф превращается в совершенный граф с центральной вершиной степени 7 (рис.8). Наконец, если к 8 удаленным кликам добавить последнюю клику размера 4 в исходном графе (C4-5), то граф превращается в одну клику размера 6 на «двух ногах» с диаметром 2, т.е. удаление клик даже уменьшило диаметр (граф как-бы сжимается, не теряя связности).
И даже после того как в список удаленных будет внесена еще одна клика C9-5, граф остается связным с диаметром 2: в нем три клики размера 3 каждая. Удаление клики C9-6, оставляет граф по-прежнему связным с диаметром 2 и только удаление еще одной клики

С9-7 приводит к потере связности. Таким образом, в этом эксперименте связность графа клик была потеряна после удаления 12 из 30 клик 40% клик.

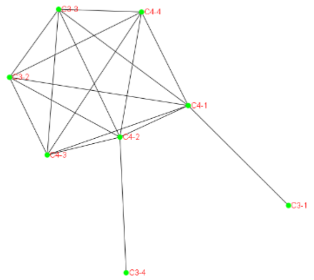

Рис.9

Рис.11

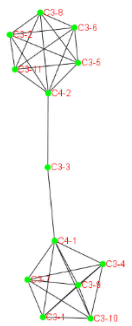

Рис. 10

Конечно, вид графов будет меняться в зависимости от порядка удаления клик – см. например, рис.9 и 10. Но что остается неизменным – явное «нежелание» графа распадаться на несвязанные части. А если все-таки приходится, то каждая часть «старается» остаться максимально связной (иметь минимальный диаметр), как на рис.11.

Заключение

Первоначально основной задачей, которую авторы ставили перед собой, было изучение очень простой исходной структуры сети, в которой возможно – также по очень простым правилам – наблюдать эволюционные изменения в структурах сетей и сравнивать их с активно исследованными в литературе. Наиболее интересным нам представлялся детерминированный самоподобный рост рассматриваемых структур. Однако, по мере расширения арсенала программных возможностей моделирования [13], наше внимание все больше занимали эффекты, связанные с поведением изучаемых сетей при удалении частей их структур. Стандартный научный подход к решению практических задач состоит в формулировке некоторого «показателя качества» решаемой проблемы и поиске методов оптимизации этого показателя при заданных ограничениях. Так, например, в задачах синтеза схем из ненадежных элементов (с ненулевой вероятностью неверного функционирования), обычно ищут возможности минимизации вероятности неверного ответа схемы, вводя *дополнительные* элементы. Возможно, это наилучший способ в условиях крайне ограниченного времени, которым располагают люди.

Природа не имеет временных ограничений и не стремится что-либо «оптимизировать». «Элементы», которыми она оперирует, стремятся «выжить», оптимизируя какие-то свои физические параметры, как «выживает» физический объект, который движется по градиенту в силовом поле. Биологические объекты, стремясь «выжить», в качестве одного из возможных решений нашли способ создавать семьи из себе подобных. Если и когда найденное решение оказывается приемлемым, оно будет повторяться, изменяясь со временем, в частности потому, что генотипу всегда «выгодно» с точки зрения выживания вида, реализовать множество фенотипов.

Наши эксперименты с моделью сети демонстрируют, что достаточно исходной «малости» - в виде задания некоторого метрического пространства и простейших правил объединения элементов этого пространства в «семьи» - для реализации поведения, которое производит впечатление чрезвычайно целесообразного. Можно предположить, что формальное объяснение эмерджентной природы робастности в рассмотренной модели следует искать в анализе вида матриц смежности для графов клик, наподобие того как в [5] при изучении активности биологических нейронных сетей было предложено анализировать матрицы весов нейронов для объяснения пространственно-временной активности групп нейронов.

Дальнейшие исследования в направлении изучения эмерджентной природы робастности позволят по-новому взглянуть как на проблемы отказоустойчивого проектирования и самовосстановления сложных сетей, так и на проблему эволюции сложных структур в целом.

Литература

1.M.E.J. Newman, The structure and function of complex networks, SIAM Rev.45 (2) (2003) 167–256

2. Leland H. Hartwell, John J. Hopfield, Stanislas Leibler and Andrew W. Murray From molecular to modular cell biology Nature · January 2000 DOI: 10.1038/35011540 · Source: PubMed

3. Richard F. Betzel1 and Danielle S. Bassett1 Multi-scale brain networks arXiv:1608.08828v2 [q-bio.NC] 4 Nov 2016

4. Pinheiro, A. V.; Han, D.; Shih, W. M.; Yan, H. (December 2011). "Challenges and opportunities for structural DNA nanotechnology". Nature Nanotechnology. **6** (12): 763–772.. doi:10.1038/nnano.2011.187.

5. Michael T. Schaub, Yazan N. Billeh, Costas A. Anastassiou, Christof Koch, and Mauricio Barahona Emergence of slow-switching assemblies in structured neuronal networks arXiv:1502.05656v3 [q-bio.NC] 20 Jul 2015

5.Трахтенброт Б.А., Барздинь Я.М. Конечные автоматы (Поведение и синтез). Наука, М., 1970

7. D. E. Knuth, The Art of Computer Programming, Volume 4A: Combinatorial Algorithms, Part 1. Addison-Wesley Professional, 2011.

8. Bocharov, Pavel, and Alexander Goryashko. "Evolutionary dynamics of partition games." In " Stability and Control Processes" in Memory of VI Zubov (SCP), 2015 International Conference, pp. 225-228. IEEE 2015. DOI: 10.1109/SCP.2015.7342108

9. P. Bocharov, “Partition Games Research Toolbox,” 2015. https://github.com/pbo/partition-games

10. George E. Andrews. 1971. Number Theory.

11. [2] George E. Andrews. 1990. Euler’s “Exemplum memorabile inductionis fallacis”and q-trinomial coefficients. J. Am. Math. Soc. 3, 3 (1990), 653–669. https://doi.org/10.2307/1990932

12. L. Samokhine Trinomial Family Research Toolbox, https://github.com/samokhine/gory 2017

13. Sven Kosub A note on the triangle inequality for the Jaccard arXiv:1612.02696v1 [cs.DM] 8 Dec 2016