

УДК 519.2

Случайные графы и грамматики на графах

© 1998 г. В. А. Малышев

Дается определение и изучаются простейшие свойства случайных грамматик на графах (многомерное обобщение случайной грамматики). Одним из базовых примеров является эволюция случайных графов, близкая по духу к теории ветвящихся процессов. Изучается асимптотический рост числа связанных компонент и независимых циклов, локальные свойства, фазовая диаграмма, шкалы топологии.

1. Введение

Грамматика, понятие, возникшее в лингвистике, стала одним из центральных понятий информатики. В [6] были введены случайные грамматики. В данной статье рассматривается их многомерное обобщение, случайные грамматики на графах. Их важность определяется тем, что этот класс случайных процессов содержит в себе многие известные конкретные процессы, а также выделяет новые связи информатики с современной математической физикой. В первой части статьи даются основные определения и свойства случайных грамматик на графах. Во второй части подробно изучается один из основных примеров, случайная эволюция графов. Постоянно прослеживается связь с классической теорией случайных графов [9, 10, 11], одной из центральных областей дискретной математики.

2. Случайные грамматики на графах

2.1. Грамматики на графах

Графы G , которые здесь рассматриваются предполагаются конечными и не направленными; $V = V(G)$ — множество вершин графа, $L = L(G)$ — множество ребер.

Спиновый граф (или спинграф, цветной граф, маркированный граф, спиновая система и т.д.) есть пара $\alpha = (G, s)$, где $s = s(\cdot)$ — функция $s: V \rightarrow S$, где S — алфавит или множество значений спина. Изоморфизм спиновых графов есть изоморфизм графов, уважающий спины. Пустой спиновый граф есть пустой граф без функции s . Случайная динамика спинграфов состоит в случайной последовательности их преобразований. Определим точно, какие преобразования рассматриваются.

9. Bush K. A. Orthogonal arrays of index unity, *Ann. Math. Stat.* (1952) **23**, 426–434.
10. Dénes J., Keedwell A. D. *Latin Squares and Their Applications*. Academic Press, New York, 1974.
11. Dénes J., Keedwell A. D. *Latin Squares. New Developments in the Theory and Applications*. North-Holland, Amsterdam, 1991.
12. Heise W., Quattrocchi P. *Informations- und Codierungstheorie*. Springer, Berlin, 1995.
13. Hill R. *A First Course in Coding Theory*. Clarendon Press, Oxford, 1994.
14. Koscielny C. Constructing a better cyclic code than cyclic Reed-Solomon code. *IEEE Trans. Inf. Theory* (1995) **41**, №4, 1191–1194.
15. Kurakin V. L., Kuzmin A. S., Mikhalev A. V., Nechaev A. A. Linear recurrences over rings and modules, *J. Math. Sci.* (1995) **76**, №6, 2793–2915.
16. van Lint J. H. *Introduction to Coding Theory*. Springer, Berlin, 1982.
17. Osborn J. M. New loops from old geometries. *Amer. Math. Monthly* (1961) **68**, 103–107.
18. Chein O., Pflugfelder H. O., Smith J. D. H. *Quasigroups and Loops: Theory and Applications*. Sigma Series in Pure Mathematics, v. 8, Heldermann, Berlin, 1990.
19. Zhu Lie. A short disproof of Euler's conjecture concerning orthogonal Latin squares. *Ars Combinatoria* (1982) **14**, 47–55.
20. Zierler N. Linear recurring sequences. *J. Soc. Ind. Appl. Math.* (1969) **7**, №1, 31–48.

Статья поступила 10.03.1998.

Определение 1. Замена $Sub = (\Gamma; \varphi: V_0 \rightarrow V(\Gamma'))$ определяется заданием двух малых спинграфов Γ и Γ' , подмножества $V_0 \subset V = V(\Gamma)$ и отображения $\varphi: V_0 \rightarrow V' = V(\Gamma')$; при этом как Γ , так и Γ' могут быть пустыми.

Определение 2. Преобразование $T = T(Sub)$ спинграфа $\alpha = (G, s)$, соответствующее данной замене Sub , определяется следующим образом. Фиксируем изоморфизм $\psi: \Gamma \rightarrow \Gamma_1$ на спиновый подграф Γ_1 спинграфа α . Рассмотрим несвязное объединение α с Γ' , удалим все ребра Γ_1 , удалим все вершины $\psi(V) \setminus \psi(V_0)$ вместе со всеми инцидентными им ребрами, отождествим каждую $\psi(v) \in \psi(V_0)$ с $v' = \varphi(v) \in \Gamma'$. Функция s на $V(G) \setminus V(\Gamma_1)$ наследуется из α , а на $V(\Gamma')$ из Γ' . Полученный спинграф обозначим $\alpha(Sub, \psi)$.

Примерами замен и преобразований могут быть удаление ребра, присоединение ребра в заданной вершине с другой новой вершиной и т.д.

Определение 3. Грамматика на графах определяется (начальным) спинграфом $\alpha_0 = (G_0, s_0)$ и конечным множеством замен $Sub_i, i = 1, \dots, r$. Грамматика на графах называется локальной, если все Γ , соответствующие всевозможным Sub_i , связны. Язык $L(\alpha_0, \{Sub_i\})$ определяется как множество всех спинграфов, которые могут быть получены из α_0 применением всевозможных замен в произвольном порядке. Точнее, $\alpha_0 \in L(\alpha_0, \{Sub_i\})$ и если $\alpha \in L(\alpha_0, \{Sub_i\})$, то $T\alpha \in L(\alpha_0, \{Sub_i\})$ для любого $T = T(Sub_i)$.

Грамматики на графах недавно сделали популярным разделом информатики (см. [2], а также труды четырех конференций по графовым грамматикам, труды первой из которых опубликованы в [3]). Заметим однако, что операции над графами в связи с теорией алгоритмов появились ранее (см. [8]). Важно понимать, что возникло (и это естественно) множество вариаций на тему грамматик на графах. Все они используют разные обобщения понятия спинграфа (см., например, [2, 3]). Примерами могут быть сети Петри с двудольными графами, гиперграфы с гиперребрами вместо ребер. Выбор конкретного определения зависит от локальной исторической традиции, личного вкуса и любимых применений. Однако, часто достаточно принятого здесь простейшего определения, так как более общие определения сводятся к принятому здесь переопределению вершин и расширением множества S .

2.2. Марковский процесс

Введем общий класс счетных цепей Маркова с непрерывным временем, случайные грамматики на графах. Состояниями такой цепи являются спиновые графы с заданным множеством спинов, то есть пары (G, s) . Процесс полностью задается начальным состоянием и интенсивностями $c(\alpha, \beta)$ преобразований из состояния $\alpha = (G, s)$ в состояние $\beta = (G_1, s_1)$.

Фиксируем грамматику на графах. Интенсивность скачка положительна тогда и только тогда, когда скачок определяется преобразованием, соответствующим замене из данной грамматики. Предполагается наличие аналога свойства трансляционной инвариантности: интенсивности $c(\alpha, \beta)$ зависят только от замены, но не от места (то есть не от выбора ψ), где эта замена делается. Иначе говоря, $c(\alpha, \beta)$ принимают конечное множество различных значений $c_i = c(Sub_i), i = 1, \dots, r$.

Стоит отметить, что введенный класс случайных процессов включает многие известные ранее процессы. Дадим несколько примеров (подробнее см. [7]).

- Ветвящиеся процессы. Для процесса Гальтона-Ватсона граф состоит из изолированных вершин, ребер нет, вершины соответствуют частицам, спин представляет тип частицы.
- Процессы с локальным взаимодействием на решетке. Граф фиксирован и не меняется со временем, только спины подвержены локальным скачкам.
- Теория очередей (см. [5]).
- Случайные фракталы (см. [6]).

2.3. Существование

Мы дадим условия отсутствия взрыва (то есть условия конечности числа скачков на любом конечном интервале времени) для введенной выше цепи Маркова с непрерывным временем. В отличие от ее одномерного аналога (см. [6]) можно легко привести примеры грамматик на графах, когда за конечное время произойдет бесконечное число скачков.

Типичный пример таков. Пусть нет спинов, а среди подстановок есть такая, что граф Γ' состоит из трех ребер, инцидентных одной общей вершине, и Γ — его подграф с двумя ребрами. Если некоторая вершина уже имеет N инцидентных ребер, то интенсивность следующего скачка в ней будет $N^2/2$, что ведет к взрыву.

Однако можно выделить достаточно общий класс грамматик на графах, где эта неприятность не возникает.

Определение 4. Фиксируем грамматику на графах и обозначим $n(\alpha, \Gamma)$ число всевозможных вложений Γ в спинграф α , пусть $n(\alpha)$ — максимум $n(\alpha, \Gamma)$ по всевозможным Γ , входящим в замены Sub с $c(Sub) > 0$. Функцию $f(\alpha)$ назовем пробной функцией, если она является верхней гранью для $n(\alpha)$, то есть $n(\alpha) \leq f(\alpha)$, и кроме того для любого преобразования T

$$f(T\alpha) - f(\alpha) < c,$$

где c — постоянная.

Лемма 1. Если существует пробная функция, то соответствующая цепь Маркова не допускает взрыва.

Доказательство. Из определений следует, что $f(\alpha(t))$ и $n(\alpha(t))$ растут не быстрее, чем линейно с ростом числа скачков. Отсутствие взрыва тогда очевидно, так как при этом условии последовательность случайных величин $n(\alpha(t))$ доминируется числом частиц в процессе чистого рождения на Z_+ , выходящем из $N = n(\alpha(0))$ и имеющем интенсивности перехода $p(i \rightarrow i + d) = Ci$, где C — максимальная интенсивность замены для нашего процесса. Но такой процесс рождения не взрывается.

Замечание 1. Тогда минимальный процесс описывается просто: каждый образ $\psi(\Gamma)$ ждет независимо от остальных соответствующее экспоненциальное время и, когда первый из них дождется, немедленно делает свою подстановку, после чего все начинается сначала.

Предложение 1. Пусть существует $d > 0$ такое, что локальная грамматика на графах обладает следующим свойством: для любого $\alpha \in L(\alpha_0, \{Sub_i\})$ вершина не может иметь более d инцидентных ребер. Тогда цепь Маркова не взрывается.

Доказательство. Для доказательства достаточно взять $f(\alpha) = c_1|V(G)|$, $\alpha = (G, s)$, где c_1 — постоянная. Заметим, что число различных d -окрестностей ограничено константой $C(d)$, и поэтому число возможных морфизмов ψ ограничено величиной $c_2|V(G)|$, где c_2 — постоянная. Тогда для всех $\alpha = (G, s)$

$$\sum_{\beta: \beta \neq \alpha} c(\alpha, \beta) < f(\alpha).$$

Следующее условие менее ограничительно и охватывает много интересных примеров.

Предложение 2. Пусть множество F замен данной грамматики на графах разбито на два непересекающихся подмножества:

$$F = F_1 \cup F_2, \quad F_1 \cap F_2 = \emptyset.$$

Пусть F_1 -преобразования имеют линейную пробную функцию $f(\alpha) = C|V(\alpha)|$. Предположим также, что для всех α, α_1 таких, что α_1 может быть получено из α одним F_2 -преобразованием, имеет место неравенство

$$\sum_{\beta} c(\alpha, \beta) \geq \sum_{\beta} c(\alpha_1, \beta).$$

Тогда цепь Маркова не взрывается.

3. Динамика графов на больших временах

3.1. Динамика среднего поля

Определения. Здесь грамматика на графах состоит из добавления и удаления ребер. Каждое ребро удаляется с интенсивностью μ . В каждой вершине v с интенсивностью λ_1 добавляется ребро, инцидентное данной вершине, причем другая вершина является новой. Добавляемые таким образом ребра мы иногда будем называть ребрами типа 1. Две только что определенные подстановки одни и те же, как для модели среднего поля, так и в рассматриваемой далее локальной модели, однако следующая замена различает эти две модели. Для обеих моделей в каждой вершине v с интенсивностью λ_2 добавляется ребро (называемое ребром типа 2), соединяющее v с другой уже существующей вершиной $v' \neq v$. При этом в модели среднего поля вершина v' выбирается равновероятно (с вероятностью $1/(V-1)$) среди существующих $V-1$ вершин, отличных от v . В локальной модели v' выбирается с равной вероятностью среди вершин из d -окрестности v для некоторого фиксированного d . Таким образом, в локальной модели графы могут иметь петли (v, v) .

Обозначим $G_t = G(t)$ граф в момент t , $V(t) = V(G(t))$. Обозначим $L_i(t) = L_i(G(t))$, $i = 1, 2$, множество ребер типа i в момент t и положим $L_1(t) + L_2(t) = L(t)$. Далее мы обозначаем одной буквой множество и число элементов в нем.

Пусть в момент ноль есть только одна изолированная вершина. В каждой вершине $v \in V(G_t)$ и $l \in L(G_t)$ независимо от остальных вершин происходят скачки со следующими инфинитезимальными переходными вероятностями

$$\begin{aligned} \mathbf{P}(G_{t+\Delta t} = G_t \cup (v, v_{\text{new}}) \mid G_t) &= \lambda_1 \Delta t + o(\Delta t), \\ \mathbf{P}(G_{t+\Delta t} = G_t \setminus l \mid G_t) &= \mu \Delta t + o(\Delta t), \end{aligned}$$

и для каждой $v, v' \in V(G_t)$

$$\mathbf{P}(G_{t+\Delta t} = G_t \cup (v, v') \mid G_t) = \frac{1}{V(t) - 1} \lambda_2 \Delta t + o(\Delta t)$$

в случае модели среднего поля и

$$\mathbf{P}(G_{t+\Delta t} = G_t \cup (v, v') \mid G_t) = \frac{1}{|O_d(v)|} \lambda_2 \Delta t + o(\Delta t)$$

для локальной модели.

Удобно рассматривать граф как одномерный топологический комплекс, чтобы изучать его топологические и метрические свойства (локальные и глобальные). Наибольший интерес это представляет в размерности, большей единицы. Размерность единица, однако, является хорошей моделью, и более того, любая графовая грамматика имеет одной из компонент обычный граф. Примером локальных свойств может быть структура окрестности вершины, примером глобальных — число независимых циклов $C(t)$ (число образующих фундаментальной группы). Напомним, что связный одномерный комплекс гомотопен пучку окружностей $S^1 \vee S^1 \vee \dots \vee S^1$.

Асимптотика роста. Процесс $V(t)$ образует цепь Маркова с вероятностями перехода

$$\mathbf{P}(V(t + \Delta t) = V(t) + 1 \mid V(t)) = \lambda_1 V(t) \Delta t + o(\Delta t).$$

Этот процесс известен в классической теории вероятностей как процесс Юла (см. [15]). Мы будем иногда пользоваться следующим обозначением: если при $n \rightarrow \infty$ случайные величины $A(t)$, $B(t)$, $B(t) \neq 0$, определены на одном вероятностном пространстве, то $A(t) \sim B(t)$ означает, что $A(t)/B(t) \rightarrow 1$ с вероятностью единица.

Лемма 2. При $t \rightarrow \infty$

$$V(t) \sim \xi \exp(\lambda_1 t),$$

где ξ — положительная случайная величина, $\mathbf{E}\xi = 1$. Более того,

$$L_i(t) \sim \frac{\lambda_i}{\lambda_1 + \mu} \xi \exp(\lambda_1 t), \quad i = 1, 2.$$

Доказательство. Утверждение следует из хорошо известных результатов в теории ветвящихся процессов. Эволюция пары $(V(t), L_1(t))$ не зависит от процесса добавления и удаления вершин типа 2 и ее, таким образом, можно рассматривать отдельно. Аналогично, эволюция $(V(t), L_2(t))$ не зависит от процесса удаления вершин типа 1. Математическое ожидание $\mathbf{E}L_i(t)$ можно найти из следующего уравнения для средних. Из равенства

$$\mathbf{E}L_i(t + \Delta t) = \mathbf{E}L_i(t)(1 - \mu \Delta t) + \lambda_i \Delta t \exp(\lambda_1 t)$$

следует, что

$$\frac{d\mathbf{E}L_i(t)}{dt} = -\mu \mathbf{E}L_i(t) + \lambda_i \exp(\lambda_1 t),$$

откуда находим, что

$$\mathbf{E}L_i(t) \sim c_i \exp(\lambda_1 t), \quad c_i = \frac{\lambda_i}{\lambda_1 + \mu}.$$

Замечание 2. Свойства случайной величины ξ можно легко извлечь из явного вида производящей функции

$$\begin{aligned} F(s, t) &= \sum_{k=1}^{\infty} s^k \mathbf{P}(V(t) = k) = s \exp(-\lambda_1 t) (1 - (1 - \exp(-\lambda_1 t))s)^{-1} \\ &= \exp(-\lambda_1 t) \sum_{j=1}^{\infty} s^j (1 - \exp(-\lambda_1 t))^{j-1} \end{aligned}$$

(см. [15, стр. 109]). В частности, полагая $s = a \exp(\lambda_1 t)$ можно видеть, что ξ имеет экспоненциальный хвост.

3.1.1. Локальные свойства. Древняя история

Локальная конечность. Фиксируем момент времени t и рассмотрим вершины, которые появились задолго до момента t . Заметим, что множество всех вершин можно естественным образом упорядочить: вершине припишем номер i , если это i -я появившаяся вершина. Пусть $l_i(t)$ — случайное число ребер инцидентных вершине i в момент t ; условимся полагать $l_i(t)$ равным нулю, если до момента t появилось менее i вершин.

Лемма 3. Если $\mu \neq 0$, то для каждой возрастающей функции $i = i(t) = o(\exp \lambda_1 t)$ и любого фиксированного $k = 0, 1, \dots$ существует предел распределений $l_i(t)$ при $t \rightarrow \infty$

$$\pi(k) = \lim_{t \rightarrow \infty} \mathbf{P}(l_i(t) = k), \quad k = 0, 1, \dots$$

Если фиксировать некоторую возрастающую функцию $\varphi(t) = o(\exp \lambda_1 t)$, то предел равномерен по всем $i(t) \leq \varphi(t)$.

Доказательство. Инфинитезимальные вероятности скачков при условии, что общее число вершин фиксировано, равны

$$\begin{aligned} \mathbf{P}(l_i(t + dt) = l - 1 \mid l_i(t) = l) &= \mu l, \\ \mathbf{P}(l_i(t + dt) = l + 1 \mid l_i(t) = l) &= \lambda_1 + \lambda_2 + \lambda_2(V - 1) \frac{1}{V - 1} = \lambda_1 + 2\lambda_2. \end{aligned}$$

Мы видим, что эти интенсивности в действительности не зависят от V и тогда $l_i(t)$ есть цепь Маркова. Эта цепь, одинаковая для всех i , очевидным образом эргодична и ее сходимость к стационарному распределению экспоненциальна. Обозначим эту цепь \mathcal{M} , пусть $\pi(k)$ — ее стационарные вероятности.

Топологический хаос. Справедлива следующая теорема.

Теорема 1. Пусть $p_{ij}(k, l; t)$ — вероятность того, что в момент времени t вершины $i = i(t)$ и $j = j(t)$, $i(t) < j(t) = o(\exp \lambda_1 t)$, имеют соответственно k и l инцидентных ребер. Тогда при $t \rightarrow \infty$

$$p_{ij}(k, l; t) \rightarrow \pi_k \pi_l.$$

Рассмотрим пару $(l_i(t), l_j(t))$. Это уже не цепь Маркова. Однако запишем снова условные переходные вероятности для этой пары при фиксированном числе вершин:

$$\begin{aligned} \mathbf{P}(l_i(t+dt) = l-1, l_j(t+dt) = k \mid l_i(t) = l, l_j(t) = k) &= \mu l, \\ \mathbf{P}(l_i(t+dt) = l, l_j(t+dt) = k-1 \mid l_i(t) = l, l_j(t) = k) &= \mu k, \\ \mathbf{P}(l_i(t+dt) = l+1, l_j(t+dt) = k \mid l_i(t) = l, l_j(t) = k) &= \lambda_1 + 2\lambda_2 - \frac{2}{V-1}, \\ \mathbf{P}(l_i(t+dt) = l+1, l_j(t+dt) = k+1 \mid l_i(t) = l, l_j(t) = k) &= \frac{2\lambda_2}{V-1}, \end{aligned}$$

и т. д. Так как $V(t) \rightarrow \infty$ п.н., процесс $(k(t), l(t))$ сходится к цепи Маркова $\mathcal{M} \times \mathcal{M}$, являющейся декартовым произведением двух цепей Маркова \mathcal{M} .

Локальная тривиальность топологии. Пусть $\mathbf{P}_i^t(O_d(i) = g)$ — вероятность того, что d -окрестность вершины i в момент t изоморфна графу g .

Лемма 4. При $t \rightarrow \infty$ вероятности $\mathbf{P}_i^t(O_d(i) = g)$, $i = i(t) = o(\exp(\lambda_1 t))$, стремятся к некоторому числу $p(d; g)$. В модели среднего поля они могут быть положительными, только если g является деревом.

Доказательство этого утверждения вполне аналогично предыдущим. Основное замечание состоит в том, что маловероятно провести ребро 2 внутри $O_d(i)$.

3.1.2. Новая история

Из экспоненциальной асимптотики следует, что $V(t-s)/V(t)$ стремится к нулю, если $s \rightarrow \infty$ и $t \rightarrow \infty$ с произвольной скоростью. Это означает, что большинство вершин достаточно молоды, то есть их дата рождения равна $t-s$ для некоторой константы s , где s не зависит от t .

Мы будем нумеровать молодые вершины иначе, чем мы это делали для старых вершин. Фиксируем s и пусть $v(s; t, \omega)$ вершина (определяемая п.н. единственным образом), чье время рождения самое близкое к $t-s$ среди остальных вершин. Пусть $l(\tau)$ — марковский процесс \mathcal{M} числа ребер, инцидентных фиксированной вершине, стартующий с $l(0) = 1$. Он делает скачки 1 с интенсивностью $\lambda_1 + 2\lambda_2$ и -1 с интенсивностью $\mu l(\tau)$.

Лемма 5. Пусть $\eta(t, v)$ — число ребер вершины v в момент t . Тогда при $t \rightarrow \infty$

$$\mathbf{P}(\eta(t, v(s; t, \omega)) = D) \rightarrow \mathbf{P}(l(s) = D),$$

$$\mathbf{P}(\eta(t, v(s_1; t, \omega)) = D_1, \eta(t, v(s_2; t, \omega)) = D_2) \rightarrow \mathbf{P}(l(s_1) = D) \mathbf{P}(l(s_2) = D), \quad s_1 \neq s_2.$$

Кроме того, вероятность $\mathbf{P}(O_d(v(s; t, \omega)) = g)$ имеет предел $P(s, g)$, который может быть отличен от нуля, только если g является деревом.

Доказательство. Возьмем две вершины, 1 и 2, появившиеся, скажем, в моменты $t-s_1 < t-s_2$ соответственно. Вероятность того, что вторая вершина произошла от первой, стремится к нулю при $t \rightarrow \infty$. Тогда вероятность

$$\mathbf{P}(l_1(t) = D_1, l_2(t) = D_2) \rightarrow \mathbf{P}(l(s_1) = D_1) \mathbf{P}(l(s_2) = D_2)$$

при $t \rightarrow \infty$, что доказывается так же, как и раньше.

Для произвольного D положим $V(D, t)$ равным числу вершин в момент t с D инцидентными ребрами. Средняя плотность числа вершин, появившихся в интервале $(t, t + dt)$, равна

$$\rho(t) = \frac{d \exp(\lambda_1 t)}{dt} = \lambda_1 \exp(\lambda_1 t).$$

Таким образом,

$$\frac{\mathbf{E}V(D, t)}{\exp(\lambda_1 t)} = \lambda_1 \int_0^t \exp(\lambda_1(s-t)) \mathbf{P}(l(t-s) = D) ds.$$

Отсюда видно, что $V(D, t)$ стремится экспоненциально быстро к нулю, если $D \rightarrow \infty$. Точная формулировка дается ниже. Пусть $V(>R, t)$ — множество вершин в момент t таких, что $l_v(t) > R$.

Лемма 6. *Существует функция $\varepsilon(R) > 0$ такая, что $\varepsilon(R) \rightarrow 0$ и*

$$\limsup_{t \rightarrow \infty} \mathbf{P}(|V(>R, t)| > \varepsilon(R)V(t)) \rightarrow 0$$

при $R \rightarrow \infty$.

3.1.3. Фазовая диаграмма

Пусть $V_D(t)$ и $V_{>D}(t)$ равны числу вершин, принадлежащих связным компонентам размера, равного D и большего D соответственно.

Теорема 2. *Если $\lambda_1 + 2\lambda_2 < \mu$, то равномерно по t имеет место экспоненциальная оценка*

$$\frac{\mathbf{E}V_{>D}(t)}{\exp(\lambda_1 t)} < a \exp(-bD).$$

Более того, существует функция $\varepsilon(D) > 0$ такая, что $\varepsilon(D) \rightarrow 0$ и

$$\limsup_{t \rightarrow \infty} \mathbf{P}(|V_{>D}(t)| > \varepsilon(D)V(t)) \rightarrow 0$$

при $D \rightarrow \infty$.

Кроме того,

$$\frac{C(t)}{\exp(\lambda_1 t)} \rightarrow 0.$$

Обозначим через $P_D(v, t)$ вероятность того, что данная вершина принадлежит связной компоненте размера D . Для оценки этой вероятности нам нужна следующая лемма.

Лемма 7. *Если $\lambda_1 + 2\lambda_2 < \mu$, то каждая вершина v имеет среднее число ребер, равное $E(t) = \mathbf{E}l_v(t) \leq 1$. Кроме того, если t_v — время появления вершины v , то для любого $\varepsilon > 0$ существует такое $\delta = \delta(\varepsilon) > 0$, что $E l_v(t) \leq 1 - \delta$ при $t - t_v \geq \varepsilon$.*

Доказательство. Обозначим $p_k(t)$ вероятность того, что $l_v(t) = k$ и пусть

$$E(t) = \sum_{k=1}^{\infty} k p_k(t).$$

Тогда

$$E(t + \Delta t) = E(t) + (\lambda_1 + 2\lambda_2)\Delta t p_0(t) + (\lambda_1 + 2\lambda_2)\Delta t(1 - p_0(t)) - \sum_{k=1}^{\infty} \mu k \Delta t p_k(t).$$

Отсюда следует, что

$$\frac{d\mathbf{E}l_v(t)}{dt} = \lambda_1 + 2\lambda_2 - \mu E(t).$$

Неподвижная точка этой динамической системы $E_f = (\lambda_1 + 2\lambda_2)/\mu$ и если стартовать из некоторой точки a , то мы останемся всегда внутри отрезка $[0, \max(a, E_f)]$.

Вернемся к доказательству теоремы. Для фиксированного d окрестность $O_d(v(s; t, \omega))$ вершины $v = v(s; t, \omega)$ является п.н. случайным деревом. Кроме того, распределение числа вершин, находящихся на расстоянии d от v совпадает с распределением числа частиц в некотором процессе Гальтона–Ватсона с дискретным временем и с одним типом частиц. Точное утверждение следующее. Для каждого d при $t \rightarrow \infty$ это число стремится к процессу Гальтона–Ватсона с дискретным временем $t = 0, 1, \dots, d$. Среднее число потомков у предельного процесса ограничено числом $\mathbf{E}l_v(t)$ (единственное несущественное отличие от стандартного процесса Гальтона–Ватсона состоит в том, что когда остается одна частица, она не может погибнуть, мы говорим в этом случае, что есть отражение в нуле). Последняя оценка доказывается так же, как и ранее. Заметим, что порядок пределов следующий: $\lim_{d \rightarrow \infty} \lim_{t \rightarrow \infty}$, причем первый предел (по времени) существует для всех значений параметров.

Тогда средний максимальный радиус дерева равен среднему времени до вырождения введенного процесса Гальтона–Ватсона. Этот процесс докритический при условиях леммы. Поэтому число вершин, лежащих на конечном расстоянии от v п.н. конечно, а это и есть связная компонента, в которой лежит v .

В двух разных вершинах эти процессы почти независимы. Поэтому нетрудно оценить дисперсию $V_D(t)$, откуда и будет следовать первое утверждение теоремы.

Мы знаем, что все d -окрестности являются деревьями с вероятностью, стремящейся к единице, откуда следует результат относительно малости числа циклов.

Перейдем теперь к другому крайнему случаю.

Теорема 3. Если $\lambda_2 > \mu/2$ и $\lambda_1 \ll \lambda_2$, то найдется такая постоянная $C > 0$, что вероятность того, что существует гигантская связная компонента размера, большего, чем $\exp(\lambda_1 t)$, стремится к единице. Более того, для числа циклов $C(t)$

$$\liminf_{t \rightarrow \infty} \frac{C(t)}{\exp(\lambda_1 t)} > 0.$$

Доказательство. Проще всего доказать это сведением к ситуации классической теории случайных графов. Пусть $V_1(t)$ — число вершин, появившихся до момента $t-1$. Рассмотрим множество $L_1(t, \omega)$ всех ребер типа 2, возникших в интервале времени $[t-1, t]$. Заметим сначала, что эти ребра можно упорядочить по времени их возникновения и получить последовательность l_1, l_2, \dots . Применим к этой последовательности следующий процесс прореживания: удалим из нее те l_i , для которых среди l_1, l_2, \dots, l_{i-1} уже было ребро между теми же вершинами, что и у l_i . Пусть $K(\omega)$ — число ребер, оставшихся после такого прореживания. Точнее, рассмотрим только

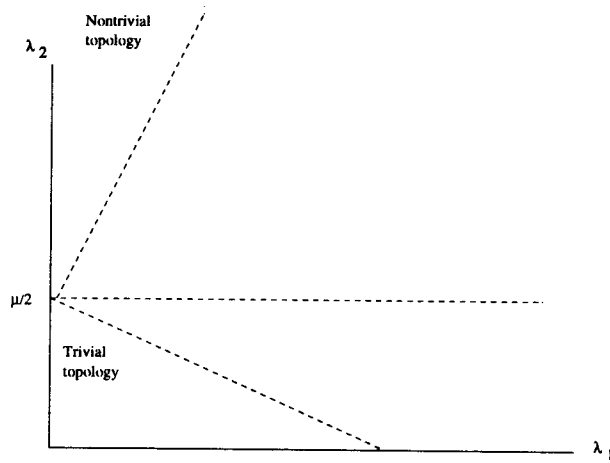


Рис. 1. Фазовая диаграмма

ребра, обе вершины которых принадлежат $V_1(t)$. Оставшееся множество ребер также упорядочено, причем условный процесс на прореженном множестве, описывающий прибавление новых ребер имеет следующее свойство: он совпадает с процессом равновероятного выбора ребра среди оставшихся ребер в $L(V_1(t))$. Рассмотрим условное распределение этих ребер при условии, что $L_1(t, \omega) = L_1$, $K(\omega) = K$: при этом распределении все такие $\binom{L_1}{K}$ возможностей имеют равную вероятность.

Оценим наиболее вероятные значения K и L_1 . Рассмотрим все вершины, появившиеся до момента $t - 1$. Из уравнения

$$\frac{d\mathbf{E}n_v(t)}{dt} = 2\lambda_2 - \mu E(t)$$

мы получим (используя первое условие теоремы), что в момент t для всех таких вершин $\mathbf{E}n_v(t) > 1$. Из второго условия теоремы следует, что большинство ребер соединяют вершины из $V_1(t)$. Точнее, для любого достаточно малого $\varepsilon > 0$ существует $\lambda_1(\varepsilon)$ такое, что при $\lambda_1 < \lambda_1(\varepsilon)$ и $t \rightarrow \infty$

$$\mathbf{P}((1 - \varepsilon)\mathbf{E}n_v(t) < K(t)/V_1(t) < (1 + \varepsilon)\mathbf{E}n_v(t)) \rightarrow 1.$$

Для окончания доказательства теперь достаточно применить известные результаты из классической теории случайных графов (см. [9]), так как с вероятностью, стремящейся к единице, мы находимся в области изменения параметров классической теории графов, когда существует гигантская компонента.

Некоторые улучшения. Оценки фазовой диаграммы, повидимому, могут быть существенно улучшены, и здесь делается несколько шагов в этом направлении.

Прежде всего, можно оценить точнее число $r_v(t) = \mathbf{E}L_1(t)$ из последней теоремы. Интуитивная картина такова. Ребра типов 1 и 2 получаются из разных случайных процессов. Поэтому полезно разделить их роль в образовании связанных компонент. Ребра типа 1 дают много связанных деревьев, остающихся конечными при $t \rightarrow \infty$. Это подтверждается следующей леммой.

Лемма 8. Вероятность того, что вершина $v(s; t, \omega)$ графа $G_1(t)$ содержится в связной компоненте размера D ограничена величиной $ca^{\log D}$ при некоторых $c > 0$ и $0 < a < 1$.

Доказательство. Допустим, что самая старая вершина v_1 в этой компоненте появилась в момент $t - s$. В частности, это означает, что процесс l_v никогда не достигал нуля за время s . Отсюда следует оценка сверху вида a^s при некотором $a < 1$. Другой источник оценок возможен в ситуации, когда много вершин появилось в течение короткого интервала времени, например, в интервале $[t - 1, t]$. Последнее дает самую худшую оценку, ее мы и используем.

Возникающие так связные компоненты можно рассматривать как эффективные вершины нового графа. Это фактор-граф $F(t)$ графа $G(t)$, получаемый отождествлением всех вершин каждой отдельной связной компоненты $G_1(t)$ и удалением всех ребер, ставших петлями в результате такой факторизации. Иначе говоря, удаляются все ребра типа 1 и все ребра типа 2, ставшие петлями. Число вершин в фактор-графе равно $m_1(t)$, и

$$\mathbf{E}m_1(t) \sim \frac{\mu}{\lambda_1 + \mu} \exp(\lambda_1 t)$$

ввиду формулы Эйлера $m = V - L_1$ и асимптотики для $V(t)$ и $L_1(t)$, полученной выше.

Оставшиеся ребра типа 2 можно рассматривать как ребра фактор-графа. Тогда число ребер на одну вершину становится равным

$$\frac{L_2(t)}{\mathbf{E}m_1(t)} = \frac{2\lambda_2}{\mu}.$$

Если бы ребра типа 2 были равномерно распределены между вершинами $F(t)$, то результат следовал бы из теории случайных графов, так как $2\lambda_2/\mu$ есть порог, устанавливаемый последней теорией. В этом случае разделяющей линией на фазовой диаграмме могла бы быть $\lambda_2 = \mu/2$. Мы не можем однако использовать здесь теорию случайных графов ввиду отсутствия равномерности.

Предполагая существование пределов, рассмотрим предельные параметры

$$L = \lim_{t \rightarrow \infty} \frac{\mathbf{E}L(t)}{\exp(\lambda_1 t)}, \quad V = \lim_{t \rightarrow \infty} \frac{\mathbf{E}V(t)}{\exp(\lambda_1 t)}, \quad C = \lim_{t \rightarrow \infty} \frac{\mathbf{E}C(t)}{\exp(\lambda_1 t)}, \quad m = \lim_{t \rightarrow \infty} \frac{\mathbf{E}m(t)}{\exp(\lambda_1 t)}.$$

Лемма 9. Если $\lambda_2 \geq \mu$, то $m > 0$ и $C > 0$.

Доказательство. Заметим, что

$$\frac{\lambda_1 + \lambda_2}{\lambda_1 + \mu} = \frac{\lambda_2}{\mu} \implies \lambda_2 = \mu$$

и

$$\frac{L}{V} = \frac{\lambda_1 + \lambda_2}{\lambda_1 + \mu}$$

монотонно возрастает на каждой вертикальной прямой $\lambda_1 = c$. По формуле Эйлера

$$V(t) - L(t) + C(t) = m(t).$$

Неравенство $m \leq C$ является необходимым и достаточным условием того, что $V(t) \leq L(t)$. Неравенство $m > 0$ справедливо, так как вероятность вырождения ветвящегося процесса всегда положительна.

Заметим, что это не доказывает существования гигантской компоненты.

3.2. Локальная динамика

Для локальной модели асимптотический рост $V(t)$, $L_1(t)$, $L_2(t)$ такой же, как и в модели среднего поля: действительно, вывод этих формул не использовал способа выбора второй вершины при проведении ребра типа 2.

Локальные свойства. Только при $d = 1$ процессы $l_v(t)$ устроены достаточно просто, как в случае среднего поля. Но случай $d = 1$ в этом смысле вырожден: если отождествлять всякий раз все ребра, соединяющие одну пару вершин, то получим дерево. При $t \rightarrow \infty$ вероятности $\mathbf{P}_i^t(O_d(i) = g)$ стремятся к пределу, обозначаемому $\mathbf{P}_i(O_d(i) = g)$. В локальном случае этот предел может быть положительным, даже если g не является деревом. Отсюда следует, что

$$\liminf_{t \rightarrow \infty} \frac{C(t)}{\exp(\lambda_1 t)} > 0$$

для всех параметров. При $d = 1$ это очевидно, мы опускаем доказательство для $d > 1$.

Связные компоненты. Число связных компонент стремится к бесконечности, так как уже отделившаяся компонента никогда не воссоединится с остальными ввиду условия локальности прибавления ребер типа 2. Любые две вершины, скажем i и j , с положительной вероятностью окажутся в разных компонентах. В локальной модели не должно быть гигантской компоненты (ни для каких параметров). Наша гипотеза состоит в том, что это имеет место и для всех значений $d < \infty$. Эта модель не должна иметь фазового перехода.

3.3. Шкалы топологии

Можно пытаться изучать более тонкую структуру связной компоненты в модели среднего поля, но мы вместо этого определим новую модель, в которой всегда есть только одна компонента. В этой новой модели (модели 2) переходы те же самые с одним единственным исключением: удаляя ребро типа 1, мы будем одновременно отождествлять обе его вершины. Здесь мы ограничимся формулировкой проблем и сделаем некоторые замечания по их поводу.

Ясно, что $m(t) \equiv 1$ и что $V(t) = L_1(t) + 1$ теперь является процессом рождения-гибели.

Лемма 10. *Цепь Маркова эргодична, если $\lambda_1 < \mu$. Она транзитивна, если $\lambda_1 > \mu$.*

Если рассматривать только вершины и ребра типа 1, то это процесс рождения-гибели, для которого классификация хорошо известна. Ребра же типа 2 не оказывают влияния на классификацию, так как интенсивность их гибели в каждой вершине растет с увеличением числа вершин.

Будем рассматривать только транзитный случай $\lambda_1 > \mu$, где для средних справедлива следующая цепочка:

$$\begin{aligned} \mathbf{E}L_1(t + \Delta t) &= \mathbf{E}L_1(t)(1 - \mu\Delta t + \lambda_1\Delta t) \implies \frac{d\mathbf{E}L_1(t)}{dt} = (\lambda_1 - \mu)\mathbf{E}L_1(t) \\ &\implies \mathbf{E}L_1(t) \sim \exp(\lambda_1 - \mu)t. \end{aligned}$$

Аналогично модели 1, находим, что

$$\mathbf{E}L_2(t) \sim c_2 \exp(\lambda_1 - \mu)t, c_2 = \frac{\lambda_2}{\lambda_1}.$$

Тогда вершины и ребра типа 1 образуют дерево и каждое добавленное ребро типа 2 увеличивает число независимых циклов на единицу. Поэтому

$$C(t) = L_2(t) \sim c_2 \exp(\lambda_1 - \mu)t,$$

как для локального случая модели 2, так и для модели 2 среднего поля.

Простота в поведении основных переменных позволяет обсудить две более тонкие характеристики, шкалы топологии и род графа. Первая связана с естественным вопросом: на какой шкале можно видеть нетривиальную топологию?

Но сначала надо знать, какие шкалы вообще возможны. Это вопрос о метрических свойствах графа: о росте его среднего диаметра $\mathbf{E}(D(t))$. В модели 2 диаметр графа совпадает с диаметром дерева из ребер типа 1. Гипотеза состоит в том, что этот рост имеет порядок ct , $c > 0$.

Рассмотрим множества $C(k, t)$, $k = 2, 3, \dots$, всех циклов длины k , не являющихся линейными комбинациями циклов из множеств $C(2, t), \dots, C(k-1, t)$. Пусть $c(k, t)$ — ранг группы (с коэффициентами в Z_2), порожденной $C(k, t)$, то есть число элементов ее базиса. Положим

$$F(k, t) = \frac{1}{C(t)} \sum_{i=2}^k c(i, t).$$

Это есть функция распределения и хотелось бы знать ее график, где на оси абсцисс откладываются скорости роста $k(t)$.

Для конкретности ограничимся шкалами ct^α , $0 \leq \alpha \leq 1$. Случай $\alpha = 0$ соответствует минимальной шкале. Дадим определение плотности топологии относительно такого выбора шкал:

$$T(\alpha) = \lim_{\varepsilon \rightarrow 0} \lim_{t \rightarrow \infty} \frac{1}{C(t)} \sum_{k=\varepsilon t^\alpha}^{\varepsilon^{-1}t^\alpha} c(k, t).$$

Лемма 11. Для локальной модели $T(0) = 1$ и, таким образом, для всех других шкал α плотность равна нулю.

Доказательство. По определению каждое новое ребро типа 2 связывает две вершины в d -окрестности. Это означает, что новый цикл появляется внутри этой окрестности. Он проходит через данное ребро и другой уже существовавший путь в этой окрестности.

В случае среднего поля кажется естественным, что $T(1) = 1$. Грубо говоря, это означает, что большинство независимых циклов имеет длину, линейную по t . Чтобы получить нетривиальную топологию на шкалах $0 < \alpha < 1$, полезны промежуточные модели: добавлять ребра типа 2, выбирая их с равной вероятностью в t^α -окрестности вершины (это соответствует моделям типа Ван дер Ваальса в статистической механике).

Род графа. Здесь мы коснемся двух других тонких свойств графа, имеющих отношение к физике. Первое касается вложений графа в двумерные поверхности.

Можно рассматривать граф G как топологическое пространство, то есть как одномерный комплекс. Тогда вложение графа в компактную ориентируемую поверхность R есть его непрерывное отображение φ в R , взаимно однозначное на $\varphi(R)$. Вложение называется клеточным, если $R \setminus \varphi(G)$ гомеоморфно несвязному объединению открытых дисков. Родом $\rho(G)$ графа G называется минимальный род поверхности R , в которую граф может быть вложен. Заметим, что род всегда конечен, так как для каждого нового ребра можно приклеить новую ручку к R . Приведем следующий результат.

Лемма 12. Если $\lambda_2/\lambda_1 > 2$, то с вероятностью, стремящейся к единице,

$$\rho(G_t) > \frac{1}{6}(\lambda_2/\lambda_1 - 2) \exp((\lambda_1 - \mu)t).$$

Доказательство. Условия леммы означают, что асимптотически $L(t) > 3V(t)$. Тогда утверждение леммы следует из известной оценки для рода в [17].

Вторая проблема состоит в попытке моделировать расширение Вселенной после Большого взрыва. Хотелось бы, чтобы две возникшие вершины либо схлопывались (отождествлялись), либо разлетались друг от друга. В нашей же модели они п.н. отождествляются. Действительно, пусть новая вершина v' появилась вместе с ребром типа 1 от вершины v . Расстояние в этот момент между v и v' равно единице. С вероятностью единица это ребро будет удалено и тогда эти две вершины будут отождествлены. Аналогично показывается, что любые две вершины будут отождествлены п.н. Однако число вершин растет быстрее, чем происходит их отождествление. Правильная формулировка эффекта разбегания могла бы быть следующей: существуют потомки (возрастающих поколений) любой пары вершин, разлетающиеся друг от друга с положительной вероятностью.

Список литературы

1. Rosenfeld A. *Picture Languages*. Academic Press, London, 1979.
2. Habel A. Hyperedge replacement: grammars and languages. *Lecture Notes in Computer Sci.* (1992) **643**.
3. Proceedings of the International Workshop. *Lecture Notes in Computer Sci.* (1979) **73**.
4. Palmer E. M., *Graphical Evolution*. Wiley, New York, 1985.
5. Малышев В. А. Взаимодействующие цепочки символов. *Успехи матем. наук* (1997) **52**, №2, 59–86.
6. Малышев В. А. Случайные грамматики. *Успехи матем. наук* (1998) **53**, №2, 107–134.

7. Malyshev V. A. Stochastic evolution via graph grammars. INRIA Preprint. March, 1998.
8. Колмогоров А. Н., Успенский В. А. К определению алгоритма. *Успехи матем. наук* (1958) **13**, №4, 3–28.
9. Bollobás B. *Random Graphs*. Academic Press, London, 1985.
10. Erdős P., Rényi A. On the evolution of random graphs. *Bull. Inst. Int. Statist. Tokyo* (1961) **38**, 343–347.
11. Колчин В. Ф. *Случайные отображения*. Наука, Москва, 1984.
12. Севастьянов Б. А. *Ветвящиеся процессы*. Наука, Москва, 1971.
13. Alon A., Spencer J. *The Probabilistic Method*. Wiley, New York, 1991.
14. Karonski M. Random Graphs. In: *Handbook of Combinatorics* (Graham R., Grottschel M., Lovasz L., Eds). MIT Press, Cambridge, 1995, pp. 351–380.
15. Athreya K. V., Ney P. E. *Branching Processes*. Springer, Berlin, 1972.
16. Gross J., Tucker T. *Topological Graph Theory*. Wiley, New York, 1987.
17. Harary F. *Graph Theory*. Addison–Wesley, Reading, Mass., 1972.

Статья поступила 03.03.1998.