

УДК 519.7

## ИССЛЕДОВАНИЕ ДИСКРЕТНО-АВТОМАТНЫХ МОДЕЛЕЙ ГЕННЫХ СЕТЕЙ НЕРЕГУЛЯРНОЙ СТРУКТУРЫ МЕТОДАМИ СИМВОЛЬНЫХ ВЫЧИСЛЕНИЙ \*)

А. А. Евдокимов, С. Е. Кочемазов, И. В. Отпущенников,  
А. А. Семёнов

**Аннотация.** Вводятся и изучаются дискретно-автоматные модели генных сетей с весовыми функциями вершин, учитывающими различные формы регуляторного взаимодействия агентов. Исследовано дискретное отображение, описывающее функционирование фрагмента генной сети бактерии *E. coli*. Для этого отображения с помощью SAT-подхода находятся его неподвижные точки (стационарные состояния). Также исследованы отображения, задаваемые случайными графами сети, которые генерировались в соответствии с известными моделями Гилберта — Эрдеша — Реньи и Уоттса — Строгатца. Для этих отображений найдены неподвижные точки и циклы длины 2 и 3. Настоящую работу можно рассматривать как обзор полученных нами результатов по дискретным моделям генных сетей и численным методам анализа их функционирования.

**Ключевые слова:** дискретная функция, дискретно-автоматная модель генных сетей, SAT.

### Введение

Генные сети [12] — активно изучаемые в последние годы модели биологических систем. Точнее говоря, они служат основой для моделирования процессов, протекающих в клетке: поддержание в организме устой-

---

\*) Исследование выполнено при частичной финансовой поддержке междисциплинарного интеграционного проекта СО РАН № 80 «Дифференциально-разностные и интегродифференциальные уравнения. Приложения к задачам естествознания», гранта Президента РФ СП–3667.2013.5, Российского фонда фундаментальных исследований (проекты 11–01–00997, 14–07–00403 и 14–07–31172мол\_а) и Совета по грантам Президента Российской Федерации для государственной поддержки ведущих научных школ (НШ–5007.2014.9).

чивых стационарных состояний, определение циклических режимов, контроль необратимых процессов развития и роста. Как математическая модель генная сеть представляет собой ориентированный граф, вершины которого интерпретируют биологических агентов, а дуги интерпретируют связи между агентами. Каждой вершине сети сопоставляется специальная функция, называемая весовой. Поведение весовой функции описывает динамику соответствующего агента — как правило, рассматривается динамика концентрации некоторого вещества, продуцируемого или поглощаемого агентом. Для произвольной вершины сети соответствующая ей весовая функция зависит от вершин, с которыми рассматриваемая вершина соединена дугами. Способы задания весовых функций могут быть различны. Известно, что в моделях генных сетей с непрерывными весовыми функциями изменения значений этих функций во времени могут описываться системами обыкновенных дифференциальных уравнений [2, 8, 12]. С другой стороны, исследуются и модели с дискретными функциями. Отметим, что в одной из первых работ, где рассматривались математические модели генных сетей — статье [21], динамика сети описывалась именно дискретными функциями.

Граф сети и весовые функции вершин, синхронно пересчитываемые в моменты времени  $t \in \{1, 2, \dots\}$ , определяют вектор-функцию, свойства которой отражают динамику сети в целом. Назовём её дискретно-автоматным отображением. В [1, 3, 6, 7, 9] для таких отображений и сетей регулярной структуры, заданных циркулянтными графами, получены условия возникновения и вид неподвижных точек и циклов функционирования.

В [5] рассматривались задачи поиска неподвижных точек дискретно-автоматных отображений, заданных графами случайной структуры, с весовыми функциями в вершинах сети из работы [1]. Для численного решения этих задач был применён SAT-подход [14].

В [4] предложена дискретная модель генной сети, функции в вершинах которой имеют существенно более сложную природу в сравнении с исследованными ранее.

В настоящей работе подробно описывается вычислительный аппарат, использованный в [4], и развит наш подход в направлении исследования сетей случайной структуры, сгенерированных в соответствии с известными моделями порождения случайных графов [16, 19, 20, 23, 25, 27].

В разд. 1 мы даём основные определения и приводим поясняющие примеры. Здесь же вводится обобщённая модель генной сети, учитывающая различные варианты взаимного регуляторного воздействия агентов.

В разд. 2 описывается сведение задач поиска неподвижных точек и циклов дискретных отображений к SAT-задачам. Также приводятся результаты использования SAT-подхода для поиска неподвижных точек отображения, заданного граф-структурой бактерии *E. coli*, изображённой ниже на рис. 3. В разд. 3 с использованием весовых функций из [4], исследуются задачи поиска неподвижных точек и циклов отображений, задаваемых сетями, которые сгенерированы в соответствии с двумя моделями порождения случайных графов.

### 1. Дискретные модели генных сетей

Рассматриваемые далее функции — дискретные функции вида

$$f_G : \{0, \dots, r\}^n \rightarrow \{0, \dots, r\}^n, \quad r \in N, \quad n \in N, \quad (1)$$

которые задаются при помощи ориентированных графов (через  $A^n$  обозначается множество всех слов длины  $n$  над конечным алфавитом  $A$ ). Граф  $G$ , задающий функцию (1), имеет  $n$  вершин, которые будем называть *агентами*. Сам граф  $G$  называем *сетью*.

Произвольной вершине  $v_i$ ,  $i \in \{1, \dots, n\}$ , графа  $G$  и каждому значению параметра  $t \in \{0, 1, \dots\}$  сопоставим число  $x_i(t) \in \{0, 1, \dots, r\}$ , называемое *весом вершины  $v_i$  в момент времени  $t$* . Переходу от момента  $t$  к  $t + 1$  соответствует синхронный пересчёт весов всех вершин. Везде далее правила пересчёта не зависят от конкретного значения  $t$  и целиком определяются структурой графа  $G$ .

Все описанные объекты в совокупности определяют дискретную динамическую мультиагентную систему. Набор весов всех вершин графа  $G$  в произвольный момент времени  $t$  называется *вектором состояния* или *состоянием данной системы*. Переходы, совершаемые системой, аналогичны переходам, которые совершает автономный детерминированный конечный автомат без выхода, а различные векторы состояний можно рассматривать как состояния такого автомата. Поэтому отображения вида (1) иногда называются *автоматными* или *дискретно-автоматными* [1, 5, 6].

Обозначим вектор состояния рассматриваемой системы в момент времени  $t$  через  $w(t)$ . Поскольку мощность множества всех различных состояний описанной системы не превосходит  $(r + 1)^n$ , для произвольного  $t_0 \geq 0$  обязательно найдутся  $k, m$ ,  $0 \leq k < m$ , такие, что  $w(t_0 + k) = w(t_0 + m)$ . В этом случае говорим, что последовательность состояний  $w(t_0 + k), \dots, w(t_0 + m)$  образует цикл длины  $m - k$ . Цикл длины 1 называется *стационарным состоянием* или *неподвижной точкой отображения* (1).

Граф состояний  $\Gamma_G$  дискретно-автоматной системы — граф на  $(r+1)^n$  вершинах, определяемый следующим образом. Каждой вершине соответствует вектор состояния системы. Вершины  $w, w'$  графа  $\Gamma_G$  соединены дугой  $(w, w')$ , направленной от  $w$  к  $w'$ , тогда и только тогда, когда результатом применения отображения (1) к вектору состояния  $w$  является вектор состояния  $w'$ . Ситуация  $(w, w)$ , т. е. петля, соответствует неподвижной точке отображения (1).

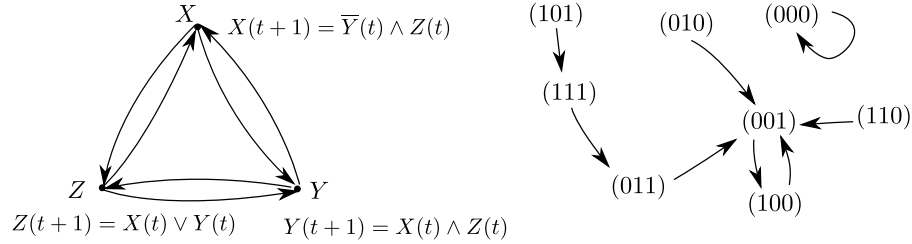


Рис. 1. Пример генной сети с весовыми функциями, заданными формулами. Справа изображён граф состояний  $\Gamma_G$

Как отмечалось выше, отображения вида (1) можно использовать для описания динамических процессов, происходящих в генных сетях. Соответствующие модели обычно называют дискретными, подчёркивая их отличие от моделей, в которых весовые функции непрерывны. Первой дискретной моделью генной сети была модель, предложенная в [21]. Весовые функции в данной модели — булевы функции, задаваемые таблицами истинности. Такой способ задания требует, вообще говоря, большого объёма данных. В [17] рассмотрены сети Кауффмана, весовые функции в которых задавались булевыми формулами. На рис. 1 приведён пример такой сети.

В [1] (см. также [12]) введены дискретно-автоматные отображения вида (1) с пороговыми весовыми функциями, заданными следующим образом:

$$x_i(t+1) = \begin{cases} x_i(t) + 1, & \text{если } \sum_{v_j \in V_i} x_j(t) = 0 \text{ и } x_i < r, \\ x_i(t) - 1, & \text{если } \sum_{v_j \in V_i} x_j(t) > 0 \text{ и } x_i > 0, \\ x_i(t) & \text{в противном случае.} \end{cases} \quad (2)$$

Здесь и далее через  $V_i$  обозначено множество всех вершин графа  $G$ , дуги из которых входят в вершину  $v_i$ . Параметр  $r$  в данных моделях называется *значностью*. На рис. 2 приведён пример функционирования

сети из 5 вершин с весовыми функциями (2) и показана ситуация входа в неподвижную точку (30300).

Для автоматных отображений с весовыми функциями вида (2) (так называемых «аддитивных автоматов»), в которых граф  $G$  является циркулянтном, в [1] найдены необходимые и достаточные условия существования неподвижных точек. В [5] на основе результатов [1] показано, что задача поиска неподвижных точек аддитивных автоматов, задаваемых сетями произвольной структуры, эффективно сводится к задаче поиска решений весьма простых систем булевых уравнений. Для решения последних в [5] применён SAT-подход.

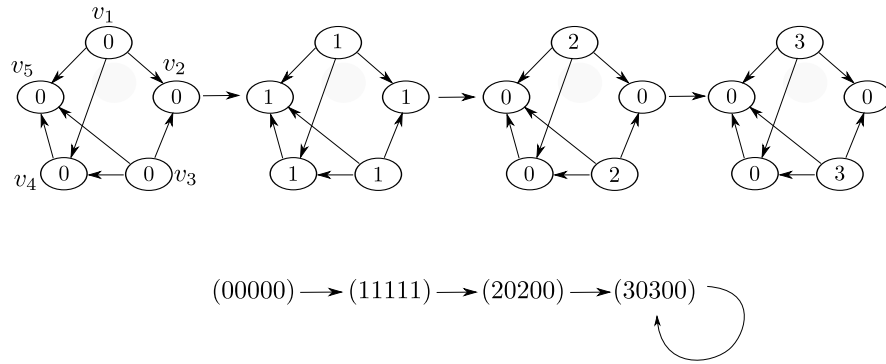


Рис. 2. Пример функционирования генной сети на 5 вершинах с весовыми функциями вида (2), значность  $r = 3$ . Слово (30300) является неподвижной точкой

Генные сети, возникающие на практике, могут содержать сотни и даже тысячи вершин с весовыми функциями довольно сложной природы. Далее рассматриваем один из таких примеров — фрагмент генной сети бактерии кишечной палочки *E. coli*, который предоставлен сотрудниками ИЦиГ СО РАН (рис. 3). Для этой сети строим дискретные весовые функции, которые учитывают исходные особенности функционирования сети. Эти функции близки к функциям (2), но имеют более разветвлённую логику задания. Приведённая далее модель схематично описана в [4].

На рис. 3 вершины графа интерпретируют различные белки-агенты, входящие в генную сеть *E. coli*. Раскраска вершины соответствует типу регуляторной деятельности белка, которая направлена на него самого (саморегуляция). При этом белая раскраска означает, что вершина никак сама себя не регулирует, серая раскраска означает, что вершина сама себя активирует, тёмная раскраска означает, что вершина сама себя ингибирует (подавляет), штриховая раскраска означает, что в зависимости

от концентрации белка вершина может себя как активировать, так и подавлять (смешанная саморегуляция).

Отметим, что раскраску произвольной вершины  $v_i$  можно рассматривать как раскраску петли  $(v_i, v_i)$ . Раскраска дуг интерпретируется по аналогии с раскраской вершин: серая дуга — активация, тёмная дуга — подавление, нет дуги — нет регуляции, пунктирная дуга — либо активация, либо подавление в зависимости от концентрации белка (смешанная регуляция).

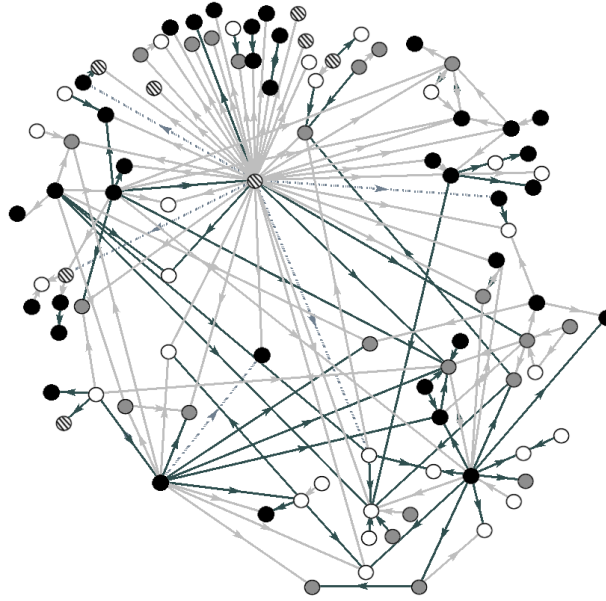


Рис. 3. Фрагмент генной сети бактерии *E. coli*

Сказанное означает, что все перечисленные случаи взаимовлияния можно учесть, задавая граф генной сети *E. coli* матрицей смежности  $A_G$  с элементами из  $\{0, 1, 2, 3\}$ . Элементы  $a_{ij}$  отражают форму воздействия вершины  $v_i$  на вершину  $v_j$ :

$$a_{ij} = \begin{cases} 0, & \text{если нет дуги/петли } (v_i \text{ не регулирует } v_j), \\ 1, & \text{если серая дуга/петля } (v_i \text{ активирует } v_j), \\ 2, & \text{если тёмная дуга/петля } (v_i \text{ подавляет } v_j), \\ 3, & \text{если пунктирная дуга/петля — смешанная форма регуляции; при этом } v_i \text{ может активировать или подавлять } v_j. \end{cases}$$

Далее строим весовые функции вершин сети *E. coli*, учитывающие все перечисленные выше факторы.

Как и выше, через  $V_i$  обозначается множество вершин, дуги из которых входят в  $v_i$ . При этом учитываются петли, т. е.  $v_i \in V_i$ , если  $v_i$  каким-либо образом себя регулирует. Через  $x_i(t)$  обозначен вес вершины  $v_i$  в момент времени  $t$ , параметр  $r$  определяет значность. Всем вершинам  $v_i$  и  $v_j \in V_i$  сопоставим вспомогательную функцию  $\delta_{ij}$ , заданную следующим образом:

$$\delta_{ij} = \begin{cases} +1, & (a_{ji} = 1) \wedge (x_j(t) > 0) \vee (a_{ji} = 3) \wedge (x_j(t) \geq \lfloor r/2 \rfloor), \\ -1, & (a_{ji} = 2) \wedge (x_j(t) > 0) \vee (a_{ji} = 3) \wedge (0 < x_j(t) < \lfloor r/2 \rfloor), \\ 0, & a_{ji} = 0 \vee x_j(t) = 0. \end{cases}$$

Смысл данной функции в том, что при пересчёте веса вершины  $v_i$  в каждый такт работы системы воздействие вершины  $v_j \in V_i$  учитывается в соответствии со следующими альтернативами:

- (i) как +1, если дуга  $(v_j, v_i)$  серая (т. е.  $v_j$  активирует  $v_i$ ) и при этом текущий вес вершины  $v_j$  больше нуля, либо если дуга  $(v_j, v_i)$  пунктирная, но текущая концентрация вершины  $v_j$  не меньше чем  $\lfloor r/2 \rfloor$ ;
- (ii) как -1, если дуга  $(v_j, v_i)$  тёмная ( $v_j$  подавляет  $v_i$ ) и текущий вес вершины  $v_j$  больше нуля, либо если дуга  $(v_j, v_i)$  пунктирная и текущая концентрация вершины  $v_j$  больше нуля, но меньше чем  $\lfloor r/2 \rfloor$ ;
- (iii) нет изменения, если отсутствует дуга  $(v_j, v_i)$ , либо если вес  $v_j$  равен нулю.

Весовая функция вершины задаётся следующим образом:

$$x_i(t+1) = \begin{cases} x_i(t) + 1, & \sum_{v_j \in V_i} \delta_{ij}(t) > 0 \text{ и } x_i(t) < r, \\ x_i(t) - 1, & \sum_{v_j \in V_i} \delta_{ij}(t) < 0 \text{ и } x_i(t) > 0, \\ x_i(t) & \text{в противном случае.} \end{cases} \quad (3)$$

## 2. Процедуры сведения задач поиска неподвижных точек и циклов автоматных отображений к задаче о булевой выполнимости (SAT)

Задача о булевой выполнимости (SAT) — поиск решений булевых уравнений вида  $\text{КНФ}=1$ , где КНФ — конъюнктивная нормальная форма. Эта задача является исторически первой NP-трудной задачей (NP-полнота соответствующей задачи распознавания есть следствие теоремы Кука 1971 г.). Несмотря на это, во многих частных случаях SAT весьма успешно решается при помощи разнообразных эвристических комбинаторных алгоритмов. Это свойство SAT привело к развитию вычислительных методов, основной областью применения которых является символьная ве-

рификация различных аппаратных и программных систем [14]. Алгоритмические идеи, используемые в современных SAT-решателях, настолько успешны, что позволяют справляться с такими аргументированно трудными тестами как задачи криптоанализа [22, 26]. Применение параллельных вычислительных технологий к решению SAT позволяет осуществлять криптоанализ реально используемых систем шифрования [11, 24].

Ещё одно позитивное свойство SAT состоит в том, что к ней можно эффективно (в общем случае за полиномиальное время) сводить комбинаторные задачи из очень широкого класса. Для практического достижения этих целей необходимо использование специализированных программных средств, которые реализуют на современных вычислительных архитектурах идеи булева кодирования алгоритмов, высказанные в [15] в отношении машины Тьюринга. Для этого подходят различные системы описания аппаратуры (например, VHDL или Verilog), применяемые при проектировании цифровых схем. Однако их использование сопряжено с целым рядом трудностей, обусловленных схемотехнической спецификой и, как следствие, избыточностью получаемого булева кода. Для сведения к SAT задач поиска неподвижных точек и циклов дискретных отображений, задаваемых генными сетями, мы использовали специализированный программный комплекс TransAlg, подробно описанный в [10].

На вход TransAlg получает описание рассматриваемой дискретной функции в виде программы, написанной на специальном C-подобном языке. Результатом компиляции этой программы является не машинный код, а система булевых уравнений, которая затем при помощи преобразований Цейтина [13] переводится в SAT-задачу. На рис. 4 приведён пример программы в системе TransAlg, которая описывает один шаг функционирования фрагмента генной сети с весовыми функциями вида (3).

Итогом трансляции подобных программ являются КНФ, в которых выделены множества переменных, кодирующих вход и выход рассматриваемого отображения вида (1). Условия на неподвижные точки и циклы записываются в виде дополнительных булевых ограничений, кодирующих равенства соответствующих входов и выходов.

После построения КНФ, кодирующей задачу поиска неподвижной точки или цикла, к полученной SAT-задаче можно применить любой из имеющихся в свободном доступе SAT-решателей. В вычислительных экспериментах использовались решатели Minisat 2.2 ([www.minisat.se](http://www.minisat.se)) и cryptominisat ([www.msoos.org/cryptominisat2](http://www.msoos.org/cryptominisat2)).

Первые примеры применения SAT-решателей к поиску неподвижных



точек и циклов дискретных отображений, задаваемых генными сетями, содержатся в [5, 18]. Как сказано выше, в [5] решены задачи поиска неподвижных точек для аддитивных автоматов, описанных в [1] и заданных сетями случайной структуры. Однако использованный в [5] подход не позволял сводить к SAT задачи поиска неподвижных точек и циклов более сложных дискретных автоматов, в том числе и отображений с весовыми функциями вида (3).

```

__in bit v[3][3];
__out bit ve[3][3];
int r = 7;
void F0 (int step){
    int hp0=(v[1]>=r/2);
    int hm0=(v[2]>0)+(v[1]>0)*(v[1]<r/2);
    if (hp0>hm0) & (v[0]<r){ve[0]=v[0]+1;}
    else if (hp0<hm0){ve[0]=v[0]-1;}
    else ve[0]=v[0];
}
void F1 (int step){
    int hp1 = (v[0]>0);
    int hm1 = (v[1]>0)+(v[2]>0);
    if ((hp1>hm1)&(v[1]<r)){ve[1]=v[1]+1;}
    else if ((hp1<hm1)&(v[1]>0)){ve[1]=v[1]-1;}
    else {ve[1]=v[1];}
}
void F2 (int step){
    int hp2=(v[0]>0)+(v[2]>=r/2);
    int hm2=(v[2]>0)*(v[2]<r/2);
    if ((hp2 > hm2)&(v[2]<r)){ve[2]=v[2]+1;}
    else if ((hp2 < hm2)&(v[2]>0)){ve[2]=v[2]-1;}
    else {ve[2]=v[2];}
}
void main(){
    F0(i);F1(i);F2(i);
}

```

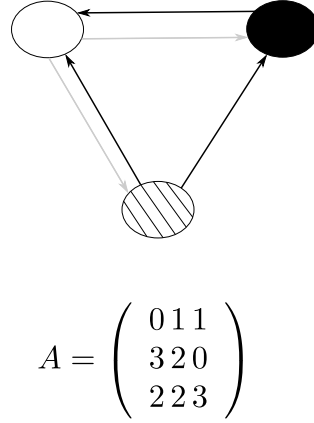


Рис. 4. Фрагмент сети из трёх вершин с весовыми функциями вида (3).

Слева приведён код TransAlg-программы, описывающей один шаг функционирования данной сети

Мы применили описанную в данном пункте технику булева кодирования алгоритмов к задаче поиска неподвижных точек отображения, заданного фрагментом генной сети *E. coli* (см. рис. 3) с весовыми функциями вида (3) и параметрами  $n = 96$ ,  $r = 31$ . КНФ, построенная при помощи системы TransAlg, состояла из 40,5 тысяч дизъюнктов над множе-

ством из 9017 переменных (объём в DIMACS-формате примерно 800 Кб). Соответствующие SAT-задачи решались на обычном персональном компьютере SAT-решателем minisat 2.2. При этом на поиск каждой новой неподвижной точки требовалось всего несколько секунд. Информация о каждой найденной неподвижной точке добавлялась затем в КНФ в виде запрещающего ограничения-дизъюнкта. В табл. 1 приведены два примера найденных неподвижных точек.

Т а б л и ц а 1

Примеры неподвижных точек дискретного отображения,  
заданного графом *E. coli*

(16,31,0,31,31,0,31,0,0,0,17,0,31,0,0,0,31,0,0,31,31,0,31,31,0,0,31,31,0,0,31,0,0,0,17,16,16,0,16,0,31,0,0,31,0,16,16,0,0,0,16,0,31,0,16,0,31,0,0,0,16,16,0,31,16,0,16,31,16,0,1,0,0,1,0,0,0,6,16,0,0,0,16,0,0,31,31,31,0,0,1,31,31,2,0)
(16,31,0,31,31,0,31,0,0,0,16,0,31,0,0,0,31,0,0,31,31,0,0,31,16,0,31,31,0,0,31,0,0,0,16,16,16,0,16,0,31,0,0,31,0,16,16,0,0,0,16,0,31,0,16,0,31,0,0,0,16,16,0,31,16,0,16,31,16,0,0,0,0,1,17,0,0,8,16,0,16,0,0,0,0,31,31,31,0,0,4,31,31,10,0)

### 3. Поиск циклов отображений, задаваемых двумя моделями порождения случайных графов

Удивительным представляется тот факт, что для отображения, определённого выше для *E. coli*, не обнаруживаются циклы малой длины. Это наблюдение стимулировало нас к рассмотрению сетей, близких *E. coli* функционально, но основанных на известных моделях порождения случайных графов. Далее мы рассматриваем две такие модели:  $G_{np}$ -модель и модель Уоттса — Строгатца.

Изначально  $G_{np}$ -модель или модель Гилберта [20] (см. также [19, 25]) позволяет порождать случайные неориентированные графы без петель. В соответствии с данной моделью случайный граф на  $n$  вершинах может быть сгенерирован независимым заполнением ячеек верхнего треугольника  $(n \times n)$ -матрицы смежности за исключением главной диагонали, элементами из  $\{0, 1\}$  с вероятностями  $1 - p$  и  $p$  соответственно (вероятность  $p$  постоянна и не зависит от номера ячейки матрицы). Данная модель естественным образом обобщается на ориентированные графы с петлями. Раскраска вершин и дуг также строится случайным образом. В рассматриваемых далее тестовых задачах на первом этапе генерировался обобщённый  $G_{np}$ -граф. Затем случайным образом наносилась раскраска вершин и дуг. Для этой цели последовательно просматривались

ячейки матрицы смежности, заполненные единицами, и для каждой такой ячейки раскраска соответствующей вершины или дуги выбиралась равновероятно из множества альтернатив  $\{1, 2, 3\}$ : 1 — серая, 2 — тёмная, 3 — пунктирная (по аналогии с сетью *E. coli*). На рис. 5 приведены примеры таких графов на 10 вершинах для значений вероятности  $p = 0,2$  и  $p = 0,5$ .

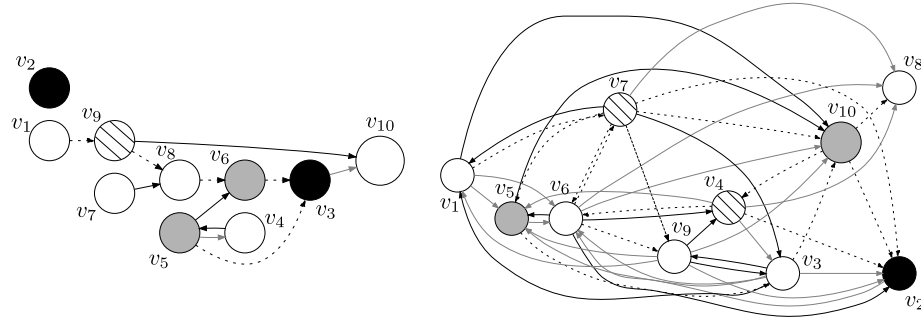


Рис. 5. Примеры сетей, сгенерированных в соответствии с  $G_{np}$ -моделью, слева вероятность дуги  $p = 0,2$ , справа —  $p = 0,5$ .

Все вычислительные эксперименты проводились на следующей платформе: одно ядро Intel Core i7-3770K, 4 Gb RAM, ОС Ubuntu 12.04, решатели Minisat 2.2, Cryptominisat 2.9.4. Для каждого набора значений параметров генерировалось 10 тестов, лимит времени решения на один тест — 2400 секунд.

На первом этапе мы рассмотрели дискретные отображения, заданные графами на 50 вершинах, сгенерированными в соответствии с обобщённой  $G_{np}$ -моделью. Весовые функции вершин имели вид (3) с параметром  $r = 31$  (значность). В большинстве решённых задач найдены циклы соответствующей длины или показано их отсутствие (см. табл. 2). Так, для графов с вероятностью дуги  $p = 0, 1$  в 3 тестах из 10 было показано, что циклов длины 3 не существует.

Ещё одной популярной моделью случайных графов является модель Уоттса — Строгатца [27]. Интерес к ней в последние годы вызван интенсивными исследованиями динамики роста глобальных информационных и социальных сетей.

В [27] отмечается, что при построении данной модели преследовалась цель учесть как детерминированные, так и случайные процессы, происходящие в реальных информационных сетях. Исходя из этого, случайный граф в соответствии с моделью [27] строится в два этапа. На первом строится регулярная структура — кольцо на  $n$  вершинах. Каждая

вершина соединена рёбрами с  $k$  «ближайшими соседями». Здесь можно использовать различные подходы. Обычно от фиксированной вершины  $v_i$ ,  $i \in \{1, \dots, n\}$ , кольцо обходится в двух направлениях. В каждом из этих направлений выбирается по  $k/2$  ближайших соседей, т. е. последовательно идущих вершин кольца, если  $k$  чётно, либо по  $\lfloor k/2 \rfloor$  и  $\lceil k/2 \rceil$  ближайших соседей в случае нечётного  $k$ . Все эти вершины соединяются рёбрами с  $v_i$ . Второй этап построения случайного графа в соответствии с данной моделью — это этап переключения рёбер. Рассматривается произвольное ребро  $(v_i, v_j)$ ,  $i < j$  (петли в исходной модели не допускаются). С вероятностью  $\beta$  принимается решение о переключении данного ребра (с вероятностью  $1 - \beta$  ребро не переключается). Рассматривается множество альтернатив выбора числа  $k \in \{1, \dots, n\}$  таких, что  $k \notin \{i, j\}$  и в текущем графе нет ребра  $(v_i, v_k)$ . Из данного множества альтернатив выбирается  $k$  в соответствии с равномерным распределением. Переключение ребра  $(v_i, v_j)$  состоит в его замене ребром  $(v_i, v_k)$ .

Т а б л и ц а 2

Результаты вычислений для функций, задаваемых случайными  $G_{np}$ -графами на 50 вершинах

Вероятность дуги	Длина цикла	Средний размер КНФ, Кб	Решено тестов	Среднее время решения, с
0,1	1	290	10/10	0,02
	2	558	10/10	0,79
	3	867	5/10	1604,95
0,2	1	471	10/10	0,73
	2	956	7/10	1104,39
	3	1522	0/10	2400

По аналогии с моделью Гилберта мы обобщили модель Уоттса — Строгатца на ориентированные графы с петлями следующим образом: на первом этапе считаем, что все дуги, соединяющие вершину  $v_i$  с её  $k$  соседями, исходят из вершины  $v_i$ . Далее делаются переключения по всем дугам в соответствии со стандартной моделью Уоттса — Строгатца. После этого по аналогии с *E. coli* случайным образом раскрашиваются вершины (0 — вершина себя не регулирует, 1 — активирует, 2 — подавляет, 3 — смешанная саморегуляция), а также дуги (случайный выбор альтернатив раскраски из  $\{1, 2, 3\}$ ). На рис. 6 приведены примеры графов, сгенерированных таким образом при  $n = 10$ ,  $k = 4$  и значениях вероятности  $\beta = 0, 2$  и  $\beta = 0, 5$ .

Были проведены вычисления по нахождению неподвижных точек и циклов для отображений, заданных сетями Уоттса — Строгатца с весо-

выми функциями вершин вида (3) и  $k = 10$ . Отметим, что во всех тестах, решённых за отведённый лимит времени, были найдены неподвижные точки и циклы длин 2 и 3 для сетей на 100 вершинах.

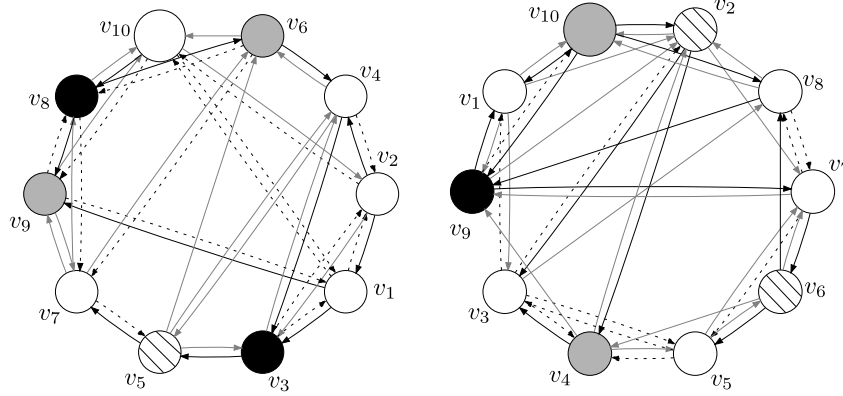


Рис. 6. Примеры сетей, сгенерированных в соответствии с моделью Уоттса — Строгатца, слева вероятность  $\beta = 0,2$ , справа —  $\beta = 0,5$

### Заключение

В статье рассмотрены дискретно-автоматные модели генных сетей, учитывающие такие виды регуляторного взаимодействия агентов — вершин сети, как активация, ингибирование и саморегуляция. Рассмотрены задачи поиска неподвижных точек и циклических режимов для автоматных отображений, задаваемых графами нерегулярной структуры. В частности, найдены неподвижные точки отображения, заданного известным фрагментом генной сети бактерии *E. coli* с дискретными весовыми функциями, учитывающими регуляторную деятельность белков, представленных в данной сети.

Введённые для модели *E. coli* весовые функции затем используются в сетях, сгенерированных в соответствии с известными моделями порождения случайных графов. В частности, рассмотрены обобщения  $G_{np}$ -модели и модели Уоттса — Строгатца. Для определяемых этими графами дискретных отображений найдены неподвижные точки и циклы длин 2 и 3.

Наши исследования по применению SAT-подхода к исследованию динамических свойств дискретно-автоматных отображений, заданных графами с определёнными на них дискретными функциями, показывают работоспособность данного подхода в отношении сетей с десятками и сотнями (в случае простых весовых функций) вершин. При этом зачастую

такие задачи удаётся решать даже без использования распределённых вычислений.

Вместе с публикацией [3] настоящую статью можно рассматривать как обзор полученных нами результатов по дискретным моделям генных сетей и численным методам анализа их функционирования.

## ЛИТЕРАТУРА

1. Григоренко Е. Д., Евдокимов А. А., Лихошвай В. А., Лобарева И. А. Неподвижные точки и циклы автоматных отображений, моделирующих функционирование генных сетей // Вестн. Томск. гос. ун-та. Прил. — 2005. — № 14. — С. 206–212.
2. Демиденко Г. В., Колчанов Н. А., Лихошвай В. А., Матушкин Ю. Г., Фадеев С. И. Математическое моделирование регуляторных контуров генных сетей // Журн. вычисл. математики и мат. физики. — 2004. — Т. 44, № 12. — С. 2276–2295.  
Demidenko G. V., Fadeev S. I., Likhoshvai V. A., Matushkin Yu. G., Kolchanov N. A. Mathematical simulation of regulatory circuits of gene networks // Comput. Math. Math. Phys. — 2004. — Vol. 44, N 12. — P. 2166–2183.
3. Евдокимов А. А. Дискретные модели генных сетей: анализ и сложность функционирования // Вычисл. технологии. — 2008. — Т. 13, № 3. — С. 31–37.
4. Евдокимов А. А., Кочемазов С. Е., Отпущенников И. В., Семенов А. А. Символьные алгоритмы решения булевых уравнений в применении к исследованию дискретных моделей генных сетей // Мат. XVI Междунар. конф. «Проблемы теоретической кибернетики» (Нижний Новгород, 2011 г.). — Нижний Новгород: Изд-во Нижегород. гос. ун-та, 2011. — С. 151–154.
5. Евдокимов А. А., Кочемазов С. Е., Семенов А. А. Применение символьных вычислений к исследованию дискретных моделей некоторых классов генных сетей // Вычисл. технологии. — 2011. — Т. 16, № 1. — С. 30–47.
6. Евдокимов А. А., Лихошвай В. А., Комаров А. В. О восстановлении структуры дискретных моделей функционирования генных сетей // Вестн. Томск. гос. ун-та. Прил. — 2005. — № 14. — С. 213–217.
7. Евдокимов А. А., Пережогин А. Л. Дискретные динамические системы циркулянтного типа с линейными функциями в вершинах сети // Дискрет. анализ и исслед. операций. — 2011. — Т. 18, № 3. — С. 39–48.  
Evdokimov A. A., Perezhogin A. L. Discrete dynamical systems of a circulant type with linear functions at the vertices of a network // J. Appl. Industr. Math. — 2012. — Vol. 6, N 2. — P. 160–166.
8. Колчанов Н. А., Латыпов А. Ф., Лихошвай В. А., Матушкин Ю. Г., Никуличев Ю. В., Ратушный А. В. Задачи оптимального

- управления в динамике генных сетей и методы их решений // Изв. РАН. Теория и системы управления. — 2004. — № 5. — С. 36–45.
- Kolchanov N. A., Latypov A. F., Likhoshvai V. A., Matushkin Yu. G., Nikulichev Yu. V., Ratushnyi A. V.** Optimal control problems in gene network dynamics and the methods for solving them // J. Comput. Syst. Sci. Int. — 2004. — Vol. 43, N 6. — P. 862–872.
9. **Кутумова Е. О.** Циклы функционирования дискретной модели регуляторного контура генной сети с пороговыми функциями // Дискрет. анализ и исслед. операций. — 2011. — Т. 18, № 3. — С. 65–75.
10. **Отпущенников И. В., Семенов А. А.** Технология трансляции комбинаторных проблем в булевы уравнения // Прикл. дискрет. математика. — 2011. — № 1. — С. 96–115.
11. **Посыпкин М. А., Заикин О. С., Беспалов Д. В., Семенов А. А.** Решение задач криптоанализа поточных шифров в распределённых вычислительных средах // Тр. ИСА РАН. — 2009. — № 46. — С. 119–137.
12. **Системная компьютерная биология.** — Новосибирск: Изд-во СО РАН, 2008. — 768 с.
13. **Цейтин Г. С.** О сложности вывода в исчислении высказываний // Зап. научн. семинаров ЛОМИ АН СССР. — 1968. — Т. 8. — С. 234–259.
14. **Biere A., Heule V., van Maaren H., Walsh T.** Handbook of satisfiability. — Amsterdam: IOS Press, 2009. — 980 p.
15. **Cook S. A.** The complexity of theorem-proving procedures // 3rd Ann. ACM Symp. Theory of Computing (Shaker Heights, Ohio, USA, 1971). — New York: ACM, 1971. — P. 151–159.
16. **Dorogovtsev S. N., Goltsev A. V., Mendes J. F. F.** Critical phenomena in complex networks // Rev. Mod. Physics. — 2008. — Vol. 80. — P. 1275–1335.
17. **Dubrova E., Teslenko M.** A SAT-based algorithm for finding attractors in synchronous Boolean networks // IEEE/ACM Trans. Comput. Biology Bioinform. — 2011. — Vol. 8, N 5. — P. 1393–1399.
18. **Dubrova E., Teslenko M., Martinelli A.** Kauffman networks: analysis and applications // Proc. ICCAD (San Jose, USA, 2005). — Washington, DC: IEEE Comput. Soc., 2005. — P. 479–484.
19. **Erdos P., Renyi A.** On random graphs // Publ. Math. — 1959. — Vol. 6. — P. 290–297.
20. **Gilbert E. N.** Random graphs // Ann. Math. Stat. — 1959. — Vol. 30, N 4. — P. 1141–1144.
21. **Kauffman S. A.** Metabolic stability and epigenesis in randomly constructed genetic nets // J. Theor. Biol. — 1969. — Vol. 22, N 3. — P. 437–467.
22. **Mironov I., Zhang L.** Applications of SAT solvers to cryptanalysis of Hash functions // Lect. Notes Comp. Sci. — 2006. — Vol. 4121. — P. 102–115.
23. **Newman M. E. J.** The structure and function of complex networks // SIAM Review. — 2003. — Vol. 45, N 2. — P. 167–256.

24. **Semenov A., Zaikin O., Bernalov D., Posypkin M.** Parallel logical cryptanalysis of the generator A5/1 in BNB-grid system // Lect. Notes Comp. Sci. — 2011. — Vol. 6873. — P. 473–483.
25. **Solomonoff R., Rapoport A.** Connectivity of random nets // Bull. Math. Biophys. — 1951. — Vol. 13. — P. 440–442.
26. **Soos M., Nohl K., Castelluccia C.** Extending SAT solvers to cryptographic problems // Lect. Notes Comp. Sci. — 2009. — Vol. 5584. — P. 244–257.
27. **Watts D. J., Strogatz S. H.** Collective dynamics of small-world networks // Nature. — 1998. — Vol. 393. — P. 440–442.

*Евдокимов Александр Андреевич,*

e-mail: evdok@math.nsc.ru

*Кочемазов Степан Евгеньевич,*

e-mail: veinamond@gmail.com

*Отпущенников Илья Владимирович,*

e-mail: otilya@yandex.ru

*Семёнов Александр Анатольевич,*

e-mail: biclop.rambler@yandex.ru

Статья поступила

2 сентября 2013 г.