УДК 021.8 + 025.1
ББК 78.34

О ДИСКРЕТНО-АВТОМАТНЫХ МОДЕЛЯХ КОНФОРМНОГО ПОВЕДЕНИЯ[1]

Семенов А.А.[2],
Кочемазов С.Е.[3]
(Институт динамики систем и теории управления СО РАН, Иркутск)

В работе для исследования феномена конформного поведения вводится дискретно-автоматная модель. Концептуально данная модель близка к дискретным моделям, используемым в компьютерной биологии для описания динамических процессов в генных сетях. Рассматриваемая модель, на наш взгляд, представляет практический интерес ввиду возможности использования для ее численного исследования символьных алгоритмов, хорошо зарекомендовавших себя в задачах верификации и криптоанализа.

Ключевые слова: модели конформного поведения, дискретно-автоматные модели, символьные алгоритмы, SAT.

Введение

В современных реалиях чрезвычайно актуальными являются вопросы, связанные с описанием динамики и прогнозированием коллективного поведения. Это обусловлено, в первую очередь, существенно возросшей в последние десятилетия скоростью обмена информацией как между отдельными людьми, так и между их группами, построенными в соответствии с различными принципами (группы по интересам, корпоративные группы и т.п.).

[1] *Работа выполнена при поддержке гранта РФФИ №11-07-00377а.*

[2] *Александр Анатольевич Семенов, кандидат технических наук, доцент (biclop.rambler@yandex.ru).*

[3] *Степан Евгеньевич Кочемазов, программист (veinamond@gmail.com).*

Наблюдающиеся кризисные ситуации в экономике и труднопредсказуемые социальные процессы демонстрируют слабую эффективность имеющихся подходов к управлению сложными мультиагентными системами и несовершенство методик, применяемых для прогнозирования их поведения.

Проблемам математического моделирования коллективного поведения посвящено множество работ. Кратко охарактеризуем наиболее значимые с нашей точки зрения.

Предлагаемые и изучаемые нами модели можно отнести к пороговым. По-видимому, первые пороговые модели, описывающие динамику коллективных процессов, были предложены в статье [33]. В дальнейшем пороговые модели коллективного поведения в различных аспектах изучались в большом числе работ: [1]-[4], [23]-[24] и др.

В ряде статей динамика коллективного поведения анализируется с теоретико-игровых позиций [1]-[4]. Так, в [4] описана общая теоретико-игровая модель конформного поведения, включающая пороговую составляющую и учитывающая различные виды взаимного влияния агентов.

В работе [12] для анализа коллективного поведения применяется вероятностный подход. Конкретно, для прогнозирования перехода каждого агента в некоторое состояние используются соотношения между априорной и апостериорной вероятностями такого перехода. При этом априорная вероятность выражает только личное мнение агента, а апостериорная – это вероятность принятия агентом решения после учета мнения его окружения.

В статьях [23], [24] в рамках пороговых моделей анализировались различные виды влияния одних агентов на других. Поведение агента при этом определяется информацией о связях и поведении его соседей. Фактически для описания взаимодействия агентов в рамках некоторой социальной группы в этих работах используются графы: агенты интерпретируются вершинами графа, а дуги или ребра графа определяют влияние одних агентов на других. Более точным термином для таких моделей является «сеть», поскольку вершины соответствующих графов могут быть

неравноправными, а ребра (дуги) нести различную смысловую нагрузку (как правило, в форме приписанных им чисел).

В последние годы появилось довольно много работ, в которых изучаются как статические, например, [27], [40], так и динамические [5], [27], [40] свойства социальных и информационных сетей. Книга [5] содержит целый ряд объединенных игровой составляющей подходов к моделированию динамики социальных сетей.

В рассматриваемых нами моделях конформного поведения, как и в «играх на сетях» [5], структура сети во времени никак не меняется. Однако описываемые и изучаемые далее модели не относятся к теоретико-игровым. В автоматных моделях функция, определяющая динамику сети, является дискретной функцией, для которой состояния равновесия и циклические режимы можно естественным образом интерпретировать на т.н. «графах состояний». Такого сорта «дискретно-автоматные» модели социальных групп имеют общие черты с моделями генных сетей, исследуемыми в информационной биологии [6]-[8], [10], [29], [30], [35].

Далее мы описываем простейшую дискретно-автоматную модель конформного поведения и изучаем в ее рамках ряд задач, имеющих, с нашей точки зрения, реальный практический смысл. Главная привлекательная черта дискретно-автоматных моделей состоит в возможности использовать для их численного исследования развитый аппарат символьных вычислений, показывающий хорошие результаты при решении трудных комбинаторных задач верификации и криптоанализа.

Приведем краткий план статьи. В первом разделе мы по аналогии с известными моделями, исследуемыми в информационной биологии, вводим простейшую дискретно-автоматную модель конформного поведения. Во втором разделе рассматриваются некоторые дополнительные детализации этой модели, имеющие самостоятельный комбинаторный смысл. В третьем разделе кратко описываются алгоритмы сведения комбинаторных задач, сформулированных в рамках рассматриваемой модели, к задачам решения булевых уравнений (в форме SAT-задач). Здесь же при-

ведена краткая информация по основным алгоритмам решения SAT-задач. В четвертом разделе приводятся результаты вычислительных экспериментов. В заключении кратко обсуждаются основные результаты статьи в сравнении с близкими по смыслу известными результатами, касающимися динамики коллективного поведения.

1. Дискретно-автоматная модель конформного поведения

Далее будем рассматривать дискретные функции следующего вида:

(1) $$f_G : \{0, \ldots, r\}^n \to \{0, \ldots, r\}^n, r \in N, n \in N,$$

задаваемые при помощи ориентированных графов (здесь через A^n обозначается множество всех слов длины n над конечным алфавитом A). Граф G, задающий функцию (1), имеет n вершин, называемых агентами (сам граф иногда называют сетью). Предполагается, что в каждый момент времени $t \in \{0, 1, \ldots\}$ произвольной вершине v_i с номером i, $i \in \{1, \ldots, n\}$, графа G приписано число $x_i(t) \in \{0, 1, \ldots, r\}$, называемое весом вершины v_i в момент времени t. Переходу от момента t к моменту $t+1$ соответствует синхронный пересчет весов всех вершин. Обычно правила пересчета 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 называется вектором состояния или состоянием данной системы. Переходы, совершаемые системой, похожи на переходы, совершаемые детерминированным конечным автоматом (ДКА) – различные векторы состояний можно рассматривать как аналоги состояний ДКА. Поэтому отображения вида (1) называются автоматными или дискретно-автоматными [6], [7].

Обозначим вектор состояния рассматриваемой системы в момент времени t через $w(t)$. Поскольку множество всех различных состояний описанной системы не превосходит $(r+1)^n$, то

для произвольного t_0, $t_0 \geqslant 0$, обязательно найдутся такие k, m, $0 \leqslant k < m$, что $w(t_0 + k) = w(t_0 + m)$. В этой ситуации говорим, что последовательность состояний $w(t_0 + k), \ldots, w(t_0 + m)$ образует цикл длины $m - k$. Цикл длины 1 называется стационарным состоянием или неподвижной точкой отображения (1).

Граф состояний дискретно-автоматной мультиагентной системы – это граф Γ_G на $(r + 1)^n$ вершинах. Каждой вершине соответствует уникальный вектор состояния. Вершины w, w' графа Γ_G соединены дугой (w, w') (направленной от w к w') тогда и только тогда (по определению), когда результатом применения отображения (1) к вектору состояния, соответствующему w, является вектор состояния, соответствующий w'. Ситуация (w, w), то есть петля, соответствует неподвижной точке отображения (1).

Как уже отмечалось выше, отображения вида (1) активно используются в информационной биологии для моделирования динамики генных сетей. По-видимому, первой дискретной моделью генной сети была модель С. Кауффмана [35], пример которой приведен на рисунке 1. Динамика поведения вершин в данной модели определяется булевыми функциями, которые задаются таблицами истинности. Конечно же, такой способ крайне неэффективен – для сети на 100 вершинах построить соответствующие таблицы истинности нельзя ни за какое разумное время. В [29] рассматривается модель Кауффмана, функции пересчета весов вершин в которой задаются булевыми формулами.

В работе [6] (см. также [10]) были введены дискретно-автоматные отображения вида (1), функции пересчета весов вершин в которых задаются при помощи следующих правил:

$$(2)\quad x_i(t+1) = \begin{cases} x_i(t) + 1, \text{если} \left(\sum\limits_{v_j \in V_i} x_j(t) = 0 \right) \text{и} \, (x_i < r) \\ x_i(t) - 1, \text{если} \left(\sum\limits_{v_j \in V_i} x_j(t) > 0 \right) \text{и} \, (x_i > 0) \\ x_i(t), \text{иначе} \end{cases}$$

Здесь и далее через V_i обозначено множество всех вершин графа G, дуги из которых входят в вершину v_i (т.о., можно сказать,

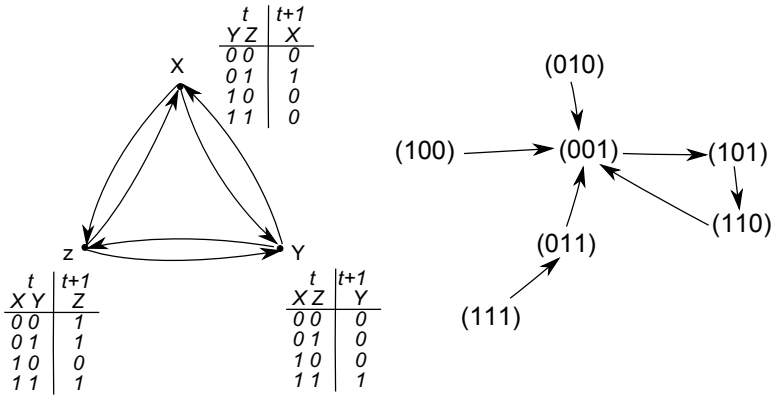

Рис. 1. Пример генной сети С. Кауффмана [35]. На рисунке изображен граф G сети с булевыми функциями пересчета весов, заданными таблицами истинности. Справа от графа сети изображен граф состояний Γ_G данной мультиагентной системы. Можно видеть, что соответствующее отображение не имеет неподвижных точек, но имеет цикл длины 3.

что вершины из V_i образуют «окружение», непосредственно влияющее на агента v_i). Для автоматных отображений с функциями весов вида (2) (т.н. «аддитивных автоматов»), граф G в которых является циркулянтом, в [6] были найдены необходимые и достаточные условия существования неподвижных точек.

В статье [7] для численного поиска неподвижных точек аддитивных автоматов (2), задаваемых случайными графами, был использован SAT-подход [22]. Удавалось находить неподвижные точки таких отображений, заданных случайными графами на 200 и более вершинах, с использованием обычного ПК. Более подробная информация о SAT-подходе будет приведена в разделе 3.

Далее мы вводим и исследуем простейшую дискретно-автоматную модель конформного поведения. В плане используемых понятий отправной точкой для нас стала работа [4]. При этом мы отмечаем, что рассматриваемая далее модель не учитывает теоретико-игровые аспекты, изученные в [4].

Определение 1. *Рассматриваем ориентированный граф G на n вершинах. Каждая вершина G интерпретирует агента, который может находиться в двух состояниях: состояние 0 – бездействие, и состояние 1 – действие. Произвольный агент v_i, $i \in \{1, \ldots, n\}$, называется θ_i-конформистом, если в момент времени $t+1$ он принимает решение «действовать», когда более $\lfloor \theta_i \cdot |V_i| \rfloor$ агентов из множества V_i в момент времени t приняли решение «действовать», в противном случае v_i принимает решение «бездействовать». Число $\theta_i \in [0, 1]$ называем далее порогом конформности агента (вершины) v_i. Вершина v_i имеет в момент времени t вес $x_i(t) = 1$, если v_i находится в состоянии «действовать», и $x_i(t) = 0$, если v_i находится в состоянии «бездействовать».*

Таким образом, в соответствии с определением 1, для любого агента v_i, $i \in \{1, \ldots, n\}$, динамику изменения его веса можно определить следующими соотношениями:

$$x_i(t+1) = \begin{cases} 1, & \sum\limits_{v_j \in V_i} x_j(t) > \lfloor \theta_i \cdot |V_i| \rfloor \\ 0, & \sum\limits_{v_j \in V_i} x_j(t) \leqslant \lfloor \theta_i \cdot |V_i| \rfloor \end{cases} \tag{3}$$

Данные соотношения задают дискретную функцию $f_G : \{0,1\}^n \to \{0,1\}^n$.

Описанная модель представляется вполне согласующейся с реальностью – например, при $\theta_i = \frac{1}{2}$ имеем агента, который принимает решение «действовать», только если большинство агентов, напрямую влияющих на него, принимают решение «действовать» (подобного рода ситуации весьма распространены на практике).

Далее мы исследуем возможность численного решения некоторых задач, естественным образом связанных с описанной моделью конформного поведения.

Определение 2. *Для рассматриваемой модели конформного поведения (3) одну итерацию синхронного пересчета весов всех вершин графа G далее называем контактом.*

Обозначим через $wt_\chi(w(t))$ вес Хэмминга вектора состояния $w(t)$ рассматриваемой системы в момент времени t.

Задача 1. *Для заданных порогов конформности всех агентов, заданных чисел $P, Q \in \{1, \ldots, n\}$, $Q > P$, и числа k выяснить, существует ли такое начальное состояние системы $w(0)$, что $wt_\chi(w(0)) \leqslant P$, а $wt_\chi(w(k)) \geqslant Q$.*

Грубо говоря, ищется начальное состояние, в котором решение «действовать» принимают не более P агентов, но после k контактов решение «действовать» принимают уже не менее Q агентов.

2. Дополнительные детализации рассматриваемой модели

Далее мы вводим в модель (3) дополнительные детализации, согласующиеся с реальными ситуациями и имеющие вполне конкретный комбинаторный смысл. А именно, будем полагать, что некоторые агенты всегда находятся в одном состоянии и ни при каких обстоятельствах это состояние не изменяют.

Определение 3. *Предположим, что в рамках рассматриваемой дискретно-автоматной модели (3) могут существовать агенты, которые в любом векторе состояния независимо от мнения их окружения принимают решение «действовать». Назовем таких агентов агитаторами. Также полагаем, что могут существовать агенты, всегда принимающие решение «бездействовать». Назовем их лоялистами. Полагаем, что остальные агенты являются конформистами с индивидуальными уровнями конформности и принимают решения в зависимости от окружающей обстановки. Этих агентов называем простыми агентами.*

Определение 4. *Рассматриваем произвольную мультиагентную систему в рамках модели конформного поведения (3), в которой имеются A агитаторов и L лоялистов, $A + L < n$. Если все простые агенты системы в начальном состоянии принимают решение «действовать», либо все они принимают решение «бездействовать», то назовем данное состояние системы начально-упорядоченным относительно действия (бездействия).*

Покажем, что справедливо следующее утверждение.

Утверждение 1. *Рассматриваем произвольную мультиагентную систему в рамках модели конформного поведения (3). Предположим, что в данной системе имеется A, $A \geqslant 0$, агитаторов, L, $L \geqslant 0$, лоялистов, и $n - A - L > 0$ простых агентов. Тогда для любого расположения агитаторов и лоялистов, любых уровней конформности простых агентов из любого начально-упорядоченного состояния описанная система достигнет стационарного состояния за не более чем $n - A - L$ контактов.*

Доказательство. Докажем справедливость утверждения для случая, когда система находится в начально-упорядоченном состоянии, в котором все простые агенты принимают решение «бездействовать». Предположим, что осуществляется один контакт между всеми агентами системы. Если ни один из простых агентов не изменил своего состояния, то имеем неподвижную точку (агитаторы и лоялисты не меняются по определению). Пусть на первом шаге некоторые простые агенты изменили свое состояние с 0 на 1. Пусть v – произвольный такой агент. Сказанное означает, что v изменил свое состояние с 0 на 1 только потому, что в его непосредственном окружении было достаточное (с позиции его уровня конформности) число агитаторов. Тогда этот агент уже не изменит свое состояние 1 ни на одной из последующих итераций (агитаторы не изменяют свое состояние). Если в результате второго контакта ни один простой агент не изменяет своего состояния, имеем неподвижную точку. Пусть v – произвольный агент, изменивший свое состояние на втором шаге. В силу сказанного выше, v изменил свое состояние с 0 на 1. Это произошло потому, что в непосредственном окружении v было достаточное (для его уровня конформности) число агитаторов и агентов, перешедших в состояние 1 после первого шага. Однако все эти агенты не изменяют своего состояния во всех последующих итерациях. Поэтому состояние 1 в дальнейшем не изменит и рассматриваемый агент. Далее по аналогии. Очевидно, что в любом случае не позднее, чем через $n - A - L$ итераций система перейдет в стационарное состояние. Случай с начальной упорядоченностью относительно действия разбирается аналогичным образом. Утверждение 1

доказано.

Данное утверждение представляется весьма полезным, поскольку его можно использовать для относительно эффективной оценки возможности достижения системой «критических» состояний (о соответствующей технике более подробно будет сказано в следующем разделе).

Определение 5. *Про мультиагентную систему в рамках модели (3) скажем, что она является* (A, L, α)*-критической относительно действия (бездействия), если существует такое расположение* A *агитаторов и* L *лоялистов, что, стартуя из начально-упорядоченного состояния относительно бездействия (действия), система через некоторое число контактов переходит в состояние, в котором* $> \alpha \cdot n$*,* $\alpha \in (0, 1)$*, агентов находятся в состоянии действия (бездействия).*

Понятие (A, L, α)-критической системы вполне согласуется с традиционными представлениями о «небезопасной социальной группе». Действительно, предположим, что система является (A, L, α)-критической относительно действия, например с $\alpha = \frac{9}{10}$. Это означает, что можно так расположить A агитаторов (относительно L лоялистов), что, стартуя из состояния, в котором все простые агенты бездействуют, система через некоторое число контактов придет в состояние, в котором $> 90\%$ агентов действуют. В соответствии с утверждением 1 для этого потребуется не более $n - A - L$ контактов. Заметим, что если система является (A, L, α)-критической, например относительно действия, то при соответствующем размещении агитаторов и лоялистов, стартуя из любого состояния (а не только из начально-упорядоченного), за не более чем $n - A - L$ контактов она достигнет состояния, в котором $> \alpha \cdot n$ агентов будут действовать. Это состояние, правда, может не быть стационарным. Если факт действия такого числа агентов воспринимается как негативное свойство рассматриваемой системы, то ее, конечно же, нельзя считать «безопасной».

В контексте введенных понятий естественным образом возникают следующие задачи исследования конформного поведения.

Задача 2. *Рассматривается мультиагентная система в*

рамках модели конформного поведения (3) с A агитаторами. Для состояния «бездействия» всех простых агентов требуется проверить, существует ли такое расположение агитаторов, при котором через некоторое число шагов ($\leqslant n - A$) более $\alpha \cdot n$ агентов будут находиться в состоянии действия (для различных уровней конформности агентов и различных α). Аналогичные задачи можно рассмотреть для лоялистов и для смешанных случаев.

Задача 3. *Предположим, что найдено некоторое размещение агитаторов, являющееся решением задачи 2. Рассматриваем систему с данным размещением агитаторов и полагаем, что L простых агентов становятся лоялистами. Требуется найти такое их расположение, чтобы начальное состояние данной системы, в котором все простые агенты бездействуют, было неподвижной точкой. Данная ситуация означает полное подавление лоялистами агитаторов.*

Задача 4. *Предположим, что найдено некоторое размещение A агитаторов, являющееся решением задачи 2. Рассматриваем систему с этим размещением агитаторов и полагаем, что L простых агентов становятся лоялистами. Требуется найти такое их расположение, что, стартуя из начального состояния, в котором все оставшиеся простые агенты действуют, через некоторое число контактов (не превосходящее $n - A - L$) система перейдет в состояние, в котором действуют $\leqslant \beta \cdot n$ агентов ($\beta < \alpha$). Данную ситуацию можно рассматривать как перевод лоялистами системы из «опасного» состояния в «более безопасное».*

3. Переход к SAT-задачам. Используемые алгоритмы решения SAT-задач

3.1. МЕТОДЫ СВЕДЕНИЯ КОМБИНАТОРНЫХ ЗАДАЧ К SAT

Известно довольно много примеров сведения различных комбинаторных задач к булевым уравнениям и, в конечном счете, к SAT-задачам. Напомним, что SAT-задачами (SAT – сокращение

от Satisfiability) [22] называются задачи поиска решений булевых уравнений вида КНФ=1, где КНФ – конъюнктивная нормальная форма [18]. Теоретическая возможность эффективных процедур сведения к SAT широкого класса комбинаторных проблем вытекает из теоремы С.А. Кука 1971г. [25]. Общие принципы построения сведений к SAT, главным образом, для систем различных ограничений (в рамках общей проблемы CSP – Constraint Satisfaction Problem), а также большое число ссылок можно найти в обзорной статье [42].

Мотивация перехода к SAT-задачам состоит в том, что алгоритмы их решения – это на сегодняшний день, пожалуй, наиболее эффективные эвристические комбинаторные алгоритмы, подтверждающие свою практическую применимость на задачах символьной верификации и даже на таком классе аргументированно трудных тестов, как задачи криптоанализа.

Для сведения к SAT задач поиска неподвижных точек и циклических режимов довольно сложных дискретно-автоматных отображений, моделирующих динамические процессы в генных сетях, в [8] был применен специальный программный комплекс TransAlg [13]. Данный программный комплекс работает с описаниями функций вида (1) в форме С-подобных программ. Результатом трансляции такой программы является не машинный код, а система булевых уравнений, которая в дальнейшем при помощи преобразований Цейтина [17] сводится к SAT-задаче.

Для сведения к SAT перечисленных выше задач в рамках модели конформного поведения (3) можно применить более адресный подход, использующий известные методы булевого кодирования целочисленных неравенств [31]. Кратко опишем здесь соответствующую технику.

Итак, рассматриваем модель вида (3). Предположим, что осуществляется k итераций синхронного пересчета весов вершин графа G (контактов). Состояние вершины v_i, $i \in \{1, \ldots, n\}$, после контакта с номером $t \in \{1, \ldots, k\}$ будем кодировать булевой переменной x_i^t; x_i^0 кодирует начальное состояние v_i.

Чтобы помечать, какие вершины являются агитаторами или

лоялистами, введем два дополнительных набора булевых переменных: $\{a_i\}_{i=1}^n$, $\{l_i\}_{i=1}^n$. Полагаем, что при $a_i = 1$, $l_i = 0$ вершина v_i является агитатором, при $a_i = 0$, $l_i = 1$ – лоялистом, а при $a_i = 0$, $l_i = 0$ – простым агентом; ситуация $a_i = 1$, $l_i = 1$ невозможна, и данный факт кодируется дизъюнктом $(\neg a_i \vee \neg l_i)$.

Пусть v_i – простой агент-конформист. Введем обозначение $\Theta_i = \lfloor \theta_i \cdot |V_i| \rfloor$. Переменная x_i^{t+1} принимает значение 1 тогда и только тогда, когда

$$\sum_{v_j \in V_i} x_j^t > \Theta_i. \tag{4}$$

Для булевого кодирования ограничений вида (4) можно использовать различные способы. Фактически мы должны эффективно записать условия истинности предиката $P_{\Theta_i}\left(x_{j_1}^t, \ldots, x_{j_{|V_i|}}^t\right)$, который истин на том и только том наборе значений переменных $x_{j_1}^t, \ldots, x_{j_{|V_i|}}^t$, на котором выполнено (4). Эти условия истинности могут быть записаны в виде системы булевых уравнений, от которой при помощи преобразований Цейтина [17] делается эффективный переход к одному уравнению вида КНФ=1, то есть к некоторой SAT-задаче.

Простейший способ построения такого рода системы уравнений – закодировать алгоритм подсчета числа единиц в векторе $x_{j_1}^t, \ldots, x_{j_{|V_i|}}^t$. Однако существуют более эффективные способы. В частности, для кодирования условий вида (4) можно использовать технику работы с т.н. «cardinality-ограничениями» (cardinality constraints [19], [20], [31],[37], [44]). Наиболее эффективные реализации этой техники, описанные в [19], [31], используют сортирующие сети. Основная идея данного подхода весьма проста и состоит в следующем. Мы можем отсортировать биты в произвольном булевом векторе $b_1, \ldots, b_m$ по возрастанию (полагаем, что старшие биты находятся слева), рассматривая их как натуральные числа из множества $\{0, 1\}$. Пусть $(s_1, \ldots, s_m)$ – результат сортировки. Очевидно, что $\sum\limits_{j=1}^{m} b_j > q$, $q \in \{0, 1, \ldots, m-1\}$, тогда и только тогда, когда $s_{q+1} = 1$. Остается выбрать алгоритм сортировки с наиболее компактной булевой кодировкой. Наилучшим

в этом смысле является алгоритм, использующий сортирующую сеть Батчера [11], [21].

При кодировании работы сортирующей сети со входом $(b_1, \ldots, b_m)$ и выходом $(s_1, \ldots, s_m)$ возникает $O\left(m \cdot log^2 m\right)$ дополнительных переменных и $O\left(m \cdot log^2 m\right)$ дизъюнктов. Легко понять, что в общем случае при кодировании с использованием сортирующих сетей k контактов в рассматриваемой модели конформного поведения и число переменных, и число дизъюнктов будут ограничены сверху величиной $O\left(k \cdot n^2 \cdot log^2 n\right)$. Учитывая, что основной интерес, в силу утверждения 1, представляют ситуации, когда $k \leqslant n - A - L$, можно считать SAT-подход вполне применимым для исследования моделей вида (3) с графами на сотнях вершин.

Для преобразования в SAT условий вида

$$\begin{aligned} &wt_\chi(w(0)) \leqslant P, \\ &wt_\chi(w(0)) \geqslant Q, \\ &wt_\chi(a_1, \ldots, a_n) \leqslant A, \\ &wt_\chi(l_1, \ldots, l_n) \leqslant L, \end{aligned}$$

также применима техника булевого кодирования cardinality-ограничений, использующая сортирующие сети.

3.2. АЛГОРИТМЫ РЕШЕНИЯ SAT-ЗАДАЧ

Здесь мы кратко останавливаемся на наиболее эффективных алгоритмах, используемых для решения SAT-задач. Практически в основе всех современных «промышленных» SAT-решателей, гарантирующих точное решение произвольной SAT-задачи, лежит алгоритм DPLL [26], а точнее его нехронологические версии, базирующиеся на идеях, впервые высказанных в работе [36].

Сам по себе алгоритм DPLL – это направленный обход дерева поиска с бэктрекингом и правилом распространения ограничений, названным впоследствии «Unit Propagation» [28]. Описанный в работе [36] алгоритм GRASP в дополнение к DPLL использует память для хранения информации о ходе поиска в форме булевых ограничений, называемых конфликтными дизъюнктами. Конфликтные дизъюнкты позволяют точно выявлять ответственные

за конфликт присвоения, что дает в ряде случаев возможность откатываться на более ранние уровни решения, чем уровень, который предшествует конфликтному уровню. Возможность такого рода «глубоких откатов» получила название «нехронологический бэктрекинг» или «бэкджампинг». В работе [46] был описан целый ряд конструкций, дополняющих общие идеи GRASP-а, что привело в итоге к возникновению нового поколения SAT-решателей, успешно применяемых к широкому классу комбинаторных задач.

Эффективность любого комбинаторного алгоритма должна подтверждаться его применимостью к решению аргументированно трудных тестов. Под аргументированно трудным мы понимаем тест, эквивалентный задаче, относительно которой есть четкая уверенность в ее высокой вычислительной сложности. Хорошим классом таких тестов являются задачи криптоанализа. В последние несколько лет растет интерес к применению SAT-подхода для решения этих задач: [14], [15], [38], [39], [43], [45].

Еще одну немаловажную положительную черту SAT-подхода составляют весьма естественные стратегии распараллеливания SAT-задач. Это позволяет использовать для их решения интенсивно развивающиеся в последние годы параллельные и распределенные вычислительные технологии и системы. В работах [14], [43] для криптоанализа алгоритма поточного шифрования A5/1 был использован распределенный SAT-решатель. Данный подход получил развитие в виде проекта добровольных распределенных вычислений SAT@home [9], [41]. Решению SAT-задачи в SAT@home предшествует стадия препроцессинга, на которой ищется декомпозиционное множество, используемое для распарарливания исходной задачи. При этом применяется специальная техника прогнозирования трудоемкости распределенного решения SAT-задач, базирующаяся на методе Монте-Карло [16].

Ряд примеров успешного применения SAT-подхода для поиска неподвижных точек и циклических режимов дискретно-автоматных отображений, используемых для моделирования динамики генный сетей, содержится в работах [7], [8].

4. Вычислительные эксперименты

Предложенная выше методика анализа дискретно-автоматных моделей конформного поведения была протестирована на случайным образом сгенерированных тестах. Конечно, такие тесты не связаны с реальными ситуациями коллективного поведения. Однако основной нашей целью на данном этапе была демонстрация принципиальной возможности решения соответствующих комбинаторных задач для размерностей, простой перебор для которых невозможен в принципе. На текущий момент мы никак не задействовали параллельные вычисления. Кроме этого, для рассматриваемых классов задач хорошие результаты могут давать различные неполные алгоритмы, которые также пока не применялись.

Во всех примерах рассматривались графы на $n = 100$ вершинах, которые генерировались по схеме, схожей с известной моделью Эрдеша-Реньи (G_{np}-графы, [27], [40]). А именно, в матрице смежности любая клетка, не находящаяся на главной диагонали, заполнялась единицей с вероятностью p и нулем с вероятностью $1 - p$; на главной диагонали везде ставились нули, что соответствовало отсутствию петель в графе. Отличие от модели Эрдеша-Реньи лишь в том, что генерируемые графы являются ориентированными, и их матрицы смежности не обязаны быть симметрическими. Конформность агентов также расставлялась случайным образом – для каждой вершины параметр θ_i генерировался как случайное число из отрезка $[0, 1]$.

Для каждой из задач 1-4 были сгенерированы по 10 тестов для значений параметра $p = 0,2$, $p = 0,3$, $p = 0,5$. Во всех вычислениях использовался SAT-решатель minisat2.2 [47], который запускался на одном ядре Core i7-3770k (16 Gb RAM, Ubuntu 12.04). При этом было выставлено ограничение по времени в 1200 секунд (20 минут), по достижении которого вычисление прерывалось с результатом «решение не найдено». В приведенных ниже таблицах содержатся средние по 10 тестам (для каждого значения p) данные.

В таблице 1 представлены результаты вычислительных экспериментов для задачи 1 с параметрами $n = 100$, $P = 40$, $Q = 80$ и $k = 10$. Отметим, что среди КНФ, SAT-задачи для которых были дорешены за отведенный лимит времени, были как выполнимые, так и невыполнимые. Таким образом, в этих случаях SAT-решатель либо находил соответствующее размещение агентов, являющееся решением задачи, либо доказывал отсутствие такого размещения.

Таблица 1. Результаты вычислительных экспериментов для задачи 1.

Вероятность дуги	Средний размер КНФ, Кб.	Решено тестов	Среднее время решения, с.
0,2	20049	10/10	144,57
0,3	35906	5/10	618,15
0,5	69047	7/10	487,40

В тестах по задаче 2 использовались следующие параметры: $n = 100$, $A = 20$, $\alpha = 0,8$, $k = 10$. Результаты тестов показаны в таблице 2.

Таблица 2. Результаты вычислительных экспериментов для задачи 2.

Вероятность дуги	Средний размер КНФ, Кб.	Решено тестов	Среднее время решения, с.
0,2	20471	10/10	4,08
0,3	34844	10/10	18,63
0,5	68980	10/10	97,19

Для каждого теста по задаче 2, используя найденное размещение агитаторов, строился соответствующий тест для задачи 3 с $L = 15$. Отметим, что неподвижные точки находились очень быстро. Соответствующие результаты представлены в таблице 3

Используя размещения агитаторов, найденные в тестах по задаче 2, строились тесты для задачи 4 с $L = 30$ и $\beta = 0,2$. Результаты этих экспериментов можно увидеть в таблице 4.

Таблица 3. Результаты вычислительных экспериментов для задачи 3.

Вероятность дуги	Средний размер КНФ, Кб.	Решено тестов	Среднее время решения, с.
0,2	2093	10/10	0,07
0,3	3373	10/10	0,20
0,5	6468	10/10	0,20

Таблица 4. Результаты вычислительных экспериментов для задачи 4.

Вероятность дуги	Средний размер КНФ, Кб.	Решено тестов	Среднее время решения, с.
0,2	20166	10/10	1,22
0,3	34540	10/10	2,74
0,5	68642	10/10	35,05

Особо отметим, что все КНФ, являющиеся тестами по задачам 2 - 4, оказались выполнимыми. То есть во всех этих задачах существовали размещения агитаторов и лоялистов с требуемыми свойствами.

5. Заключение

Статья посвящена исследованию коллективного поведения с позиций дискретно-автоматных моделей, схожих по своим свойствам с дискретными моделями генных сетей, изучаемыми в информационной биологии. Мотивация работы с данным классом моделей заключается в том, что для их численного исследования может быть применен аппарат символьных вычислений, демонстрирующий хорошие результаты на аргументированно трудных тестах из символьной верификации и криптоанализа.

В статье подробно рассмотрена простейшая дискретно-автоматная модель, в рамках которой естественным образом могут быть поставлены задачи описания поведения «небезопасных социальных групп». В частности, рассмотрена задача поиска на-

чальных состояний с относительно малым числом действующих агентов, из которых за небольшое число контактов возможны переходы в состояния с большинством действующих агентов. Далее в модель введены агенты, никогда не меняющие своего состояния (действия – агитаторы либо бездействия – лоялисты). Для этого варианта модели установлено, что для достижения ею стационарного состояния из состояния начальной упорядоченности требуется не более $n - A - L$ контактов (n – общее число агентов, A – число агитаторов, L – число лоялистов). Если достигнутое стационарное состояние таково, что более $\alpha \cdot n$, $\alpha \in (0, 1)$, агентов в нем, например действуют, притом что в начальном состоянии все простые агенты бездействовали, то соответствующая система называется (A, L, α)-критической. Системы данного типа при α, близких к 1, можно считать «небезопасными» в том плане, что для них можно указать размещение агитаторов, которое переведет почти всех бездействующих простых агентов в состояние действия. Однако в некоторых случаях ситуацию можно исправить, «перекупив» часть простых агентов и сделав их лоялистами – система из состояния действия всех простых агентов через некоторое число контактов переходит в состояние, в котором число действующих агентов не превосходит $\beta \cdot n$, где $\beta < \alpha$.

На наш взгляд, немаловажный аспект моделирования коллективного поведения составляет проблема построения таких моделей (безусловно, адекватных реальным ситуациям), для которых возможно эффективное их численное исследование. Причем важно не только наблюдать за их эволюцией во времени (что можно считать простой «прогонкой» модели), но и уметь находить конфигурации, соответствующие различным «экстремальным» формам поведения модели (например, искать как устойчивые, так и неустойчивые состояния, циклические режимы и т.п.). В ряде работ эти проблемы предлагается решать с теоретико-игровых позиций [1]-[5], [24] (поиск равновесий по Нэшу, состояний, оптимальных по Парето, и др.). Для рассматриваемых нами моделей данные задачи могут решаться без привлечения теории игр, а лишь отталкиваясь от весьма простых свойств дискретных функ-

ций.

В будущем интерес представляет развитие предложенного подхода в ряде естественных направлений. Например, не составит большого труда при трансляции в SAT учитывать различные виды нагрузки дуг графа G, которая может отображать, например «социальное давление» (как в [4]) либо какие-то другие факторы взаимного влияния агентов.

Сформулированное выше условие критичности для систем с агитаторами и лоялистами можно усилить, используя формализм дважды квантифицированных булевых формул (2-QBF) [22]. Скажем, если обозначить через S – размещение A агитаторов, а через T – размещение L лоялистов, то усиленное условие «критичности» системы может выглядеть следующим образом:

$$\exists S \forall T \Re(G, f_g, S, T, \alpha) = 1,$$

где $\Re$ – предикат, принимающий значение «истина», если система, будучи начально-упорядоченной с бездействием простых агентов, переходит в критическое (с параметром α) состояние относительно действия. Для проверки наличия этого свойства, в соответствии с утверждением 1, достаточно рассмотреть поведение системы при $\leqslant n - A - L$ контактах. Для численного исследования достижимости такого рода ситуаций могут быть использованы известные 2-QBF-решатели (например, [32], [34] и др.).

На данном этапе для решения сформулированных в работе задач использовались стандартные «промышленные» SAT-решатели (в частности, minisat2.2, [47]), основной областью применения которых является символьная верификация. Данные решатели всегда выдают точное решение SAT-задачи, то есть либо находят выполняющий КНФ набор, либо доказывают ее невыполнимость.

Отметим, что в рассмотренных в статье задачах в случае существования одного решения зачастую существует много других решений. Для таких задач вполне могут дать хорошие результаты различные неполные алгоритмы – например, те или иные стратегии локального поиска. В ближайшем будущем предполагается

использовать такие алгоритмы для работы с моделями конформного поведения, существенно превосходящими по размерности модели, рассмотренные в статье.

Авторы выражают глубокую благодарность Владимиру Валентиновичу Брееру за его ценные замечания, позволившие устранить целый ряд неточностей, имевшихся в первоначальной версии статьи.

Литература

1. БРЕЕР В.В. *Теоретико-игровая модель неанонимного порогового конформного поведения* //Управление большими системами. - 2010. - № 31. - С. 162–176.
2. БРЕЕР В.В., НОВИКОВ Д.А. *Пороговые модели взаимного страхования* //Математическая теория игр и ее приложения. - 2011. - Том 3. - № 4. - С. 3–22.
3. БРЕЕР В.В., НОВИКОВ Д.А. *Пороговая модель коррупционного поведения* //Системы управления и информационные технологии. - 2011. - № 3. - С. 73–75.
4. БРЕЕР В.В. *Теоретико-игровые модели конформного поведения* //Автоматика и телемеханика. - 2012. - № 10. - С. 111-126.
5. ГУБАНОВ Д.А., НОВИКОВ Д.А., ЧХАРТИШВИЛИ А.Г. *Социальные сети: модели информационного влияния, управления и противоборства.* //М.: Физматлит. 2010. – 228 с.
6. ГРИГОРЕНКО Е.Д., ЕВДОКИМОВ А.А., ЛИХОШВАЙ В.А., ЛОБАРЕВА И.А. *Неподвижные точки и циклы автоматных отображений, моделирующих функционирование генных сетей.* //Вестник Томского гос. ун-та. Приложение. - 2005. – № 14. - С. 206–212.
7. ЕВДОКИМОВ А.А., КОЧЕМАЗОВ С.Е., СЕМЕНОВ А.А. *Применение символьных вычислений к исследованию дискретных моделей некоторых классов генных сетей.* //Вычислительные технологии. 2011. - Т. 16. - № 1. С. 30–47.

8. 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.* //Материалы XVI Международной конференции «Проблемы теоретической кибернетики». Нижний Новгород. 2011. - С. 151–154.
9. ЗАИКИН О.С., СЕМЕНОВ А.А., ПОСЫПКИН М.А *Процедуры построения декомпозиционных множеств для распределенного решения SAT-задач в проекте добровольных вычислений SAT@HOME.* //Управление большими системами. - 2013. – Т. 43. - С. 138–156
10. *Системная компьютерная биология* //Под ред. Н.А. Колчанова, С.С. Гончарова, В.А. Лихошвая, В.А. Иванисенко. – Новосибирск: Изд-во СО РАН - 2008. - 767 с.
11. КОРМЕН Т., ЛЕЙЗЕРСОН Ч., РИВЕСТ Р. *Алгоритмы. Построение и анализ.* //М. МЦНМО. - 2002
12. КРАСНОЩЕКОВ П.С. *Простейшая математическая модель поведения. Психология конформизма* //Математическое моделирование. - 1998. - Т. 10. - № 7. – С. 76–92.
13. ОТПУЩЕННИКОВ И.В., СЕМЕНОВ А.А. *Технология трансляции комбинаторных проблем в булевы уравнения.* //Прикладная дискретная математика. - 2011. - № 1. - С. 96–115.
14. ПОСЫПКИН М.А., ЗАИКИН О.С., БЕСПАЛОВ Д.В., СЕМЕНОВ А.А *Решение задач криптоанализа поточных шифров в распределенных вычислительных средах.* //Труды ИСА РАН. - 2009. - № 46. - С. 119–137.
15. СЕМЕНОВ А.А., ЗАИКИН О.С., БЕСПАЛОВ Д.В., УШАКОВ А.А. *SAT-подход в криптоанализе некоторых систем поточного шифрования.* //Вычислительные технологии. - 2008. - Т. 13 - № 6. - С. 134–150.
16. СЕМЕНОВ А.А., ЗАИКИН О.С. *Алгоритмы построения декомпозиционных множеств для крупноблочного распараллеливания SAT-задач.* //Известия ИГУ. Серия: Математика. - 2012. – Т. 5. - № 4. - С. 79–94.

17. ЦЕЙТИН Г.С. *О сложности вывода в исчислении высказываний.* //Записки научных семинаров ЛОМИ АН СССР. - 1968, - т. 8. - С. 234–259.
18. ЯБЛОНСКИЙ С.В. *Введение в дискретную математику.* //М.: Наука. - 1986. - 384 с.
19. ASIN R., NIEUWENHUIS R., OLIVERAS A., RODRIGUEZ-CARBONELL E. *Cardinality Networks: a theoretical and empirical study.* //Constraints - 2011. - Vol. 16. № 2. - P. 195-221.
20. BAILLEUX O., BOUFKHAD Y. *Efficient CNF encoding of boolean cardinality constraints.* //LNCS. - 2003. – Vol. 2833. - P. 108-122.
21. BATCHER K.E. *Sorting Networks and their Applications.* // In Proc. of AFIPS. - 1968. – Vol. 32. - P. 307–314.
22. *Handbook of Satisfiability.* //edited by Biere A., Heule V., van Maaren H., Walsh T. - IOS Press - 2009. – 980 P.
23. BRAUN N. *Individual Thresholds and Social Diffusion* // Rationality and Society - 1995. – № 7.- P. 167–182.
24. CHWE M. *Structure and Strategy in Collective Action* // AJS - 1999. – Vol. 105. № 1. P.128–156.
25. COOK S.A *The complexity of theorem-proving procedures // Third annual ACM symposium on Theory of computing.* //Third annual ACM symposium on Theory of computing - 1971 - Ohio, USA. ACM. - 1971. - P. 151–159.
26. DAVIS M., LOGEMANN G., LOVELAND D. *A machine program for theorem proving* //Communication of the ACM. - 1962. – Vol. 5 - Issue 7. - P. 394 – 397.
27. DOROGOVTSEV S.N., GOLTSEV A.V., MENDES J.F.F. *Critical phenomena in complex networks* //Rev. Mod. Phys. - 2008. – Vol.80. - P. 1275–1335.
28. DOWLING W., GALLIER J. *Linear-time algorithms for testing the satisfiability of propositional Horn formulae.* //Journal of Logic Programming. - 1984. - №1(3). - P. 267–284.
29. DUBROVA E., TESLENKO M., MARTINELLI A. *Kauffman*

networks: analysis and applications. //Proc. Of ICCAD - 2005. - P. 479–484.

30. DUBROVA E., TESLENKO M. *A SAT-Based Algorithm for Finding Attractors in Synchronous Boolean Networks.* //IEEE/ACM Transactions on Computational Biology and Bioinformatics. 2011 – Vol. 8 - №5. - P. 1393–1399.
31. EEN N., SORENSSON N. *Translating Pseudo-Boolean Constraints into SAT.* //Journal on Satisfiability, Boolean Modeling and Computation. - 2006. - Vol. 2. - P. 1–26.
32. GIUNCHIGLIA E., MARIN P., NARIZZANO M *Reasoning with quantified boolean formulas. In Handbook of Satisfiability (editors: A. Biere, M.Heule, H. van Maaren, T. Walsh).* //2009. - IOS Press. - P. 761–780.
33. GRANOVETTER M. *Threshold Models of Collective Behavior* // AJS - 1978. - Vol. 83. № 6. - P. 1420-1443.
34. JANOTA M., MARQUES-SILVA J.P. *Abstraction-Based Algorithm for 2QBF* //LNCS. - 2011. – V. 6695. - P. 230–244.
35. KAUFFMAN S. A. *Metabolic stability and epigenesis in randomly constructed genetic nets.* //Theor. Biol. - 1969. - Vol. 22, - №. 3. - P. 437–467.
36. MARQUES-SILVA J.P., SAKALLAH K.A. *GRASP: A search algorithm for propositional satisfiability.* //IEEE Transactions on Computers. - 1999. - Vol. 48. - № 5. - P. 506–521.
37. MARQUES-SILVA J., LYNCE I. *Towards Robust CNF Encodings of Cardinality Constraints.* //LNCS. 2007. – Vol. 4741. - P. 483–497.
38. MCDONALD C., CHARNES C., PIEPRZYK J. *Attacking Bivium with MiniSat.* //Tech. Rep. - 2007/040 - ECRYPT Stream Cipher Project. 2007.
39. MIRONOV I., ZHANG L. *Applications of SAT Solvers to Cryptanalysis of Hash Functions.* //LNCS. - 2006. – Vol. 4121 - P. 102–115.
40. NEWMAN M.E.J. *The structure and function of Complex Networks* //SIAM Review. - 2003. - V. 45, - №2. - P. 167–256.
41. POSYPKIN M., SEMENOV A., ZAIKIN O. *Using BOINC*

desktop grid to solve large scale SAT problems. //Computer Science Journal. - 2012. - Vol. 13. - № 1. - P. 25–34.
42. PRESTWICH S. *CNF encodings. In Handbook of Satisfiability (editors: A.Biere, M.Heule, H. van Maaren, T. Walsh).* //2009. - IOS Press. - P. 75–97.
43. SEMENOV A., ZAIKIN O., BESPALOV D., POSYPKIN M. *Parallel logical cryptanalysis of the generator A5\1 in BNB-Grid system.* //LNCS. 2011. – V. 6873. - P. 473–483.
44. SINZ C. *Towards an Optimal CNF Encoding of Boolean Cardinality Constraints.* //LNCS. 2005. – Vol. 3709. - P. 827–831.
45. SOOS M., NOHL K., CASTELLUCCIA C. *Extending SAT Solvers to Cryptographic Problems.* //LNCS. - 2009. - Vol. 5584. - P. 244–257.
46. ZHANG L., MADIGAN C.F., MOSKEWICZ M.H., MALIK S. *Efficient conflict driven learning in a boolean satisfiability solver.* //In Proc. of ICCAD. - 2001. - P. 279–285.
47. http://www.minisat.se/.

ANALYSIS OF SOME DISCRETE-AUTOMATON MODELS OF THRESHOLD BEHAVIOUR

Alexander Semenov, Institute of System Dynamics and Control Theory SB RAS, Irkutsk, Cand.Sc., assistant professor (biclop.rambler@yandex.ru).
Stepan Kochemazov, Institute of System Dynamics and Control Theory SB RAS, Irkutsk, researcher (veinamond@gmail.com).

Abstract: In this paper to analyze the phenomenon of collective behaviour we introduce a discrete automaton model. This model is based on the ideas used in modern computational biology to describe dynamical processes in gene networks. In our oppinion the model introduced is interesting from the practical point of view because it is possible to use modern symbolic algorithms (used for example in cryptanalysis and verification) for its numerical study.

Keywords: models of collective behaviour, discrete-automaton models, symbolic algorithms, SAT.

Статья представлена к публикации
членом редакционной коллегии ...

Поступила в редакцию ...
Дата опубликования ...