

ՀՀ ԳԱԱ ԻՆՖՈՐՄԱՏԻԿԱՅԻ ԵՎ ԱՎՏՈՄԱՏԱՑՄԱՆ ՊՐՈԲԼԵՄՆԵՐԻ
ԻՆՍՏԻՏՈՒՏ

Հովհաննիսյան Մինաս Հենրիկի

Բազմամասշտաբային ճարտարապետությամբ բարդ ցանցային համակարգերի նպատակային նախագծումը բաշխված հաշվողական ռեսուրսների օգտագործմամբ

Ե13.04 – «Հաշվողական մեքենաների, համալիրների, համակարգերի և ցանցերի մաթեմատիկական և ծրագրային ապահովում» մասնագիտությամբ տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի հայցման ատենախոսության

ՍԵՂՄԱԳԻՐ

Երևան – 2017

ИНСТИТУТ ПРОБЛЕМ ИНФОРМАТИКИ И АВТОМАТИЗАЦИИ НАН РА

Оганесян Минас Генрикович

Целевое проектирование сетевых систем многомасштабной архитектуры с использованием распределенных вычислительных ресурсов

АВТОРЕФЕРАТ

Диссертации на соискание ученой степени кандидата технических наук по специальности

05.13.04 – “Математическое и программное обеспечение вычислительных машин, комплексов, систем и сетей”

Ереван – 2017

Ատենախոսության թեման հաստատվել է Երևանի Պետական Համալսարանում

Գիտական ղեկավար՝	Ֆիզ.-մաթ. գիտ. դոկտոր	Ս.Ա.Նիգիյան
Ընդդիմախոսներ՝	տեխ. գիտ. դոկտոր	Հ.Հ.Հարությունյան
	տեխ. գիտ. թեկնածու	Մ.Ղ.Գյուրջյան

Առաջատար կազմակերպություն՝ Ռուսաստանի Գիտությունների Ակադեմիայի
Ն.Ն. Սեմյոնովի անվ. քիմիական ֆիզիկայի ինստիտուտ

Պաշտպանությունը կայանալու է 2017թ. Հունիսի 8-ին, ժ.16:00-ին ՀՀ ԳԱԱ
Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտում գործող 037
«Ինֆորմատիկա» մասնագիտական խորհրդի նիստում հետևյալ հասցեով՝ Երևան,
0014, Պ. Սևակի 1:

Ատենախոսությանը կարելի է ծանոթանալ ՀՀ ԳԱԱ ԻԱՊԻ գրադարանում:
Սեղմագիրն առաքված է 2017թ. Մայիսի 8-ին:

Մասնագիտական խորհրդի գիտական

քարտուղար, Ֆիզ.-մաթ. գիտ. Դոկտոր



Հ. Գ. Սարուխանյան

Тема диссертации утверждена в Ереванском Государственном Университете

Научный руководитель:	доктор физ.-мат. наук	С. А. Нигилян
Официальные оппоненты:	доктор тех. наук	Г.А. Арутюнян
	кандидат тех. наук	М.К. Гюрджян

Ведущая организация:

Институт химической физики им. Н.Н. Семенова Российской академии наук.

Защита состоится 8-го июня 2017г. в 16:00 на заседании специализированного
совета 037 “Информатика” Института проблем информатики и автоматизации НАН
РА по адресу: 0014, г. Ереван, ул. П. Севака 1.

С диссертацией можно ознакомиться в библиотеке ИПИА НАН РА.

Автореферат разослан 8-го мая 2017г.

Ученый секретарь специализированного

совета, доктор физ.-мат. наук



А. Г. Саруханян

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Теория сетей является активно развивающейся областью науки. Множество объектов и явлений представляют в виде статических и динамических сетевых структур, что позволяет применять аппарат теории сетей для широкого спектра явлений и задач. Большой интерес вызывают критические явления в сетях, происходящие в результате различных процессов, ввиду наблюдаемых аналогий с различными физическими, химическими и биологическими явлениями. Важным случаем процесса, приводящего к подобному явлению – фазовому переходу – является процесс направленного развития сетей с топологическими ограничениями. Изучение упомянутых явлений невозможно осуществить в полной мере без их программной реализации и анализа экспериментальных результатов, полученных на основе компьютерных симуляций. Эти результаты необходимо получать для различных значений параметров рассматриваемых явлений; необходимо рассматривать ансамбли сетей, которые включают тысячи сетей с одинаковыми параметрами; кроме того, само явление может требовать проведения множества шагов, количество которых достигает десятков миллионов. Поэтому, программная реализация требует применения различных техник распределения и распараллеливания выполняемых вычислений. Множество существующих программных систем моделирования сетей не позволяют рассматривать динамические процессы на сетях, а предназначены для их статического анализа и визуализации. Системы, позволяющие моделировать динамические процессы в сетях, являются узконаправленными и предназначены для проведения исследований в специфичных областях, таких, как, например, сети протеиновых взаимодействий или компьютерные сети. Другие системы предоставляют общие подходы для моделирования подобных процессов, однако от пользователя требуется полное описание необходимых действий, что требует много времени и соответствующих знаний среды от пользователя для моделирования конкретного процесса и не обеспечивает гибкости решения относительно модификаций процесса. Кроме того, результаты экспериментов требуют соответствующей автоматической обработки с целью получения необходимых характеристик и динамик их изменения, что также необходимо реализовать пользователю. Следовательно, актуальной задачей является создание программной системы, которая позволит выполнять необходимое моделирование рассматриваемой динамики с возможностью легкой и быстрой настройки под конкретные характеристики и последующего автоматического анализа полученных экспериментальных данных.

Объект и предмет исследования. Основными объектами исследования данной работы являются:

1. Алгоритмы и программные методы моделирования динамики направленного развития случайных сетей.

2. Параллельные и распределенные технологии моделирования.
3. Динамика направленного развития сложных сетевых структур.
4. Критические и коллективные явления в сетевых структурах.

Основными предметами исследования представленной работы являются:

1. Применение параллельных и распределенных технологий для эффективного компьютерного моделирования сетей и процесса направленного развития в них.
2. Исследование поведения сетевых структур и их характеристик при направленном развитии.
3. Исследование фазового перехода при рассматриваемой динамике.

Цель и задачи исследования. Основной целью диссертационной работы является разработка алгоритмов и программных методов для эффективного компьютерного моделирования направленных процессов в сетях и их применение для экспериментального исследования процесса направленного развития сетей.

Исходя из указанной цели, в работе поставлены и решены следующие задачи:

1. Разработка алгоритмов и программных методов для моделирования процесса направленного развития сетей.
2. Разработка программных средств для моделирования процесса направленного развития сетей с возможностью получения строения и основных свойств рассматриваемых сетей.
3. Моделирование динамики развития сетей с использованием разработанной системы; получение и анализ экспериментальных данных.
4. Исследование процесса направленного развития сетевых структур и характеристик изучаемых и конструируемых сетей.

Научная новизна. Разработаны алгоритмы и программные методы моделирования направленных процессов в сетях. На их основе разработана программная система, позволившая получить экспериментальные результаты и выявить приведенные ниже новые явления, наблюдаемые при моделировании направленного развития случайных сетей Эрдёша-Реньи с топологическими ограничениями:

1. Процесс направленного развития с сохранением степеней вершин приводит к расщеплению сети на плотные (почти полные) компоненты со слабыми связями между компонентами. Данное явление называется структурообразованием.
2. Расщепление сопровождается образованием двух зон в спектральной плотности матриц смежности рассматриваемого ансамбля сетей. Основная зона в процессе развития сужается и меняет форму от Вигнеровской в начале к треугольной в конце рассматриваемого процесса. Вторая зона в

начале состоит из нескольких изолированных значений, а в конце процесса зона становится плотной и связанной.

3. Матрицы смежности сетей изменяют свою структуру от случайных в начале рассматриваемого процесса к блочно-диагональным в конце. Матрицы смежности финальных сетей содержат блоки примерно равных размеров, количество которых обратно пропорционально вероятности связей между вершинами сети.

Практическая ценность. Разработанная программная система может быть использована для разностороннего исследования различных направленных динамик в сетях. Экспериментальные результаты исследований, проведенных на случайных сетях Эрдёша-Реньи, могут быть использованы для описания различных физических и биологических явлений, таких, как явление почкования в мембранах липидов, некоторых явлений из теории струн и квантовой гравитации, а также для описания процессов в нейронных сетях мозга. Процесс направленного развития сетей с топологическими ограничениями можно применять как инструмент для разбиения случайных сетей на оптимальные части из почти полных подсетей.

На защиту выносятся следующие положения:

1. Разработаны алгоритмы и программные методы моделирования процесса направленного развития сетей. Разработана программная система, в которой реализованы соответствующие алгоритмы с применением программных методов. Система, кроме моделирования процесса направленного развития сетей, позволяет проводить разносторонний анализ сетей и получение динамик различных свойств. Система показала свою эффективность в ходе проведенных экспериментальных исследований и внедрена в Институте химической физики им. Н.Н. Семенова РАН.
2. Случайная сеть в результате процесса направленного развития с топологическими ограничениями расщепляется на множество почти полных кластеров (клик).
3. Количество кластеров, образующихся в результате процесса направленного развития случайных сетей Эрдёша-Реньи, обратно пропорционально вероятности связи.
4. Слабые связи между кластерами, образующимися в результате процесса направленного развития случайных сетей Эрдёша-Реньи, определяют масштабно-инвариантные свойства финальных сетей.

Апробация результатов работы. Основные результаты диссертации были представлены на семинаре кафедры программирования и информационных технологий ЕГУ, на общем семинаре факультета информатики и прикладной математики ЕГУ, на научной школе-семинаре "Технологии моделирования сложных физических и биологических систем" (Цмакаог, 2015), на международной конференции Computer Science and Information Technologies - 2015 (CSIT, Ереван,

2015) и на международном семинаре Workshop on Critical and collective effects in graphs and networks (CCEGN, Москва, 2016).

Публикации. Основные результаты диссертации опубликованы в 3 публикациях.

Структура и объем диссертации. Диссертация состоит из введения, трех глав и заключения общим объёмом 110 страниц, включая 94 наименования литературы. Работа включает 24 рисунка, 3 диаграммы и 3 таблицы.

СОДЕРЖАНИЕ РАБОТЫ

Во введении представлен общий обзор области исследований, актуальность рассматриваемых задач и приведены основные результаты диссертационной работы.

В главе 1 представлен обзор области исследований. Представлена теория сетей, основные рассматриваемые задачи и имеющиеся результаты. Представляется процесс направленного развития сетей, который является объектом исследования и обосновывается актуальность его исследования. Далее представлены вопросы программного моделирования процессов в сетях, доступные системы и библиотеки моделирования сетей и приводятся их основные недостатки в контексте программного моделирования рассматриваемой задачи. Далее формулируются объекты и предметы исследований, представляются поставленные цели и задачи исследований и описываются методология и методы исследований.

В главе 2 представлены основные теоретические и технологические результаты и идеи, лежащие в основе диссертационной работы. Глава состоит из трех разделов.

Раздел 2.1 описывает модель случайных графов Эрдёша-Реньи. Описываются основные свойства сетей, методы визуализации сетей, представлены способы обнаружения сообществ, а также спектральные свойства матриц.

Раздел 2.2 представляет идею семплирования и дает описание одного из его распространенных вариантов – алгоритма Метрополиса – Гастингса. Данная идея лежит в основе процесса направленного развития сетей, рассматриваемого в данной работе, ввиду того, что позволяет выполнить структурно-параметрическую идентификацию наилучшей статистической модели случайного процесса и является эффективным приемом для логичного смыслового увязывания статистических свойств выборки и цели моделирования.

Раздел 2.3 описывает модель процесса направленного развития сетей. Далее приведен краткий обзор этого процесса.

Рассматривается классическая модель случайных графов Эрдёша-Реньи – $G = (V, E)$, где $V(G) = \{x_1, \dots, x_N\}$ – непустое конечное множество вершин и $E(G)$ –

конечное множество ребер. Модель обладает симметричной матрицей смежности A с элементами a_{ij} , такими, что $a_{ij} = 1$, если $(x_i, x_j) \in E(G)$ и $a_{ij} = 0$ в противном случае.

Предположим задана некоторая сеть G . Обозначим через $T(G)$ количество циклов длины три в G . Далее, будем рассматривать процесс развития – *эволюцию* – сети в направлении роста $T(G)$. Данный процесс представляет из себя последовательность трансформаций начального графа $G - \{G = G_0, G_1, G_2, \dots, G_k\}, k \gg 1$. Каждый последующий граф получается из предыдущего с помощью некоторого переключения связей. Рассматриваемые в данной работе варианты переключения связей приведены ниже.

Значение $T(G)$ в процессе переключения связей может меняться. Обозначим через Δ_i изменение рассматриваемой величины в течении одного шага процесса – $\Delta_i = T(G_i) - T(G_{i-1}), 1 \leq i \leq k$. Для принятия или отклонения полученной сети применяется алгоритм Метрополиса-Гастингса, а именно – вероятность ω принятия переключения определяется следующим образом:

$$\omega = \begin{cases} 1, & \text{if } \Delta_i \geq 0 \\ e^{\mu \Delta_i}, & \text{if } \Delta_i < 0 \end{cases},$$

где μ – параметр модели: чем больше значение μ , тем меньше вероятность принятия сетей при отрицательном значении Δ_i . Очевидно, что все шаги, ведущие к росту величины $T(G)$, принимаются безусловно. Параметр модели μ имеет физический смысл – в термодинамике он соответствует понятию химического потенциала, используемого при описании состояния систем с переменным числом частиц и представляет собой энергию добавления одной частицы в систему без совершения работы.

Эволюция представляется *эволюционной траекторией*, т.е. изменением количества циклов длины три как зависимости от числа элементарных переключений (времени).

В ходе изучения динамики нас интересуют равновесные состояния и, поэтому, динамический процесс продолжается до наступления сходимости. Под условием сходимости подразумевается продолжение эволюции до наступления *стабилизации* – состояния системы, при котором количество циклов длины три не подвергается значительным изменениям на достаточно длинном участке переключений, $\Phi(G, \mu)$, и система колеблется (*флуктуирует*) около некоторого фиксированного значения. Естественно ассоциировать это состояние с равновесным состоянием.

Процесс переключения связей, основанный на определенной стохастике, принято также называть *рандомизацией*. Случай без топологических ограничений достаточно изучен. Случай с наложением топологических ограничений является новым направлением рассмотрения данной динамики – фазовый переход в системе при топологических ограничениях на мотивы был рассмотрен в 2015 году.

Сети, полученные в результате данного процесса, будем называть *финальными*.

Далее приводится описание двух рассматриваемых в диссертационной работе вариантов переключения связей в течении процесса.

Процесс переключения без сохранения степеней. На каждом шаге процесса случайным образом (имеется в виду равномерное распределение по всем связям сети) выбирается одна связь – (u_1, v_1) . Далее, из всех возможных пар несвязанных вершин случайным образом (равномерное распределение по всем указанным парам) выбирается одна пара – $\{u_2; v_2\}$. Связь (u_1, v_1) удаляется из сети, а пара $\{u_2; v_2\}$ – соединяется ребром. В результате данного переключения сохраняется общее количество связей в сети, но меняется топология. Далее приведен рисунок (рис. 1),

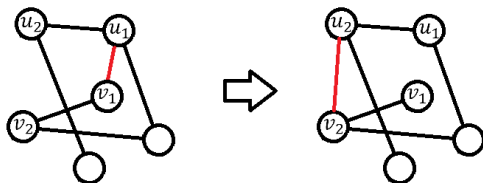


Рисунок 1. Переключение связей с сохранением степеней узлов

показывающий одно такое переключение.

Процесс переключения с сохранением степеней. В данном варианте на каждом шаге случайным образом выбирается два ребра $\{(u_1, v_1); (u_2, v_2)\}$ при соблюдении следующего условия: в сети отсутствуют ребра (u_1, u_2) и (v_1, v_2) . Далее два выбранных ребра удаляются из сети и добавляются ребра (u_1, u_2) и (v_1, v_2) . В

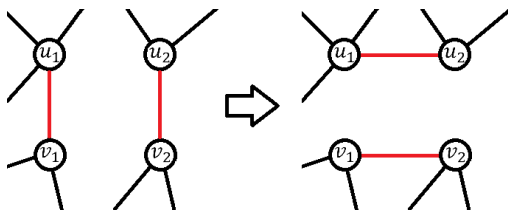


Рисунок 2. Переключение связей без сохранения степеней узлов

результате данного переключения сохраняется не только общее количество ребер, но и распределение степеней. Ниже представлен пример подобного переключения (рис. 2).

Очевидно, что в данном примере условиям будет удовлетворять также добавление ребер (u_1, v_2) и (v_1, u_2) после удаления изначально выбранных ребер.

Раздел 2.4 содержит описание основных технологий параллельного и распределенного программирования, рассмотренных при разработке программной среды моделирования.

Глава 3 описывает алгоритмы и программные методы, разработанные для моделирования направленных процессов в сетях. Представлена программная

систему, разработанную для проведения необходимых исследований. Глава состоит из пяти разделов.

Раздел 3.1 представляет основные требования к программной системе для проведения исследований, которые включают следующие пункты:

1. Возможность моделирования процесса направленного развития с поддержкой двух описанных ранее вариантов.
2. Возможность работы с большими ($> 10^3$) ансамблями сетей.
3. Возможность получения длинных ($> 10^7$ шагов), достигающих стабилизации траекторий.
4. Возможность получение сетей по ходу процесса.
5. Поддержка графического представления полученных данных.
6. Эффективное использование доступных ресурсов.

В разделе 3.2 приведено обоснование необходимости проектирования и создания специализированной программной системы с использования параллельных и распределенных технологий. Далее приведены основные аспекты проводимых исследований и представлены соответствующие примененные методы решения возникающих вычислительных задач.

Первый этап исследования заключается в определении и изучении критических точек параметра μ . На данном этапе выполняются следующие шаги:

1. Нахождение критического порога параметра μ .
2. Построение траекторий вблизи критического порога.
3. Анализ стабильных хвостов траектории.
4. Анализ основных свойств сетей
5. Получение результатов для различных значений параметров.

Первый шаг предполагает для каждого фиксированного начального состояния (N, p) просмотр широкого интервала значений параметра μ с целью выявления критического порога. В интервале равномерно выбираются несколько значений параметра и для каждого из них генерируется траектория. Количество шагов, в зависимости от начальных параметров, может меняться от нескольких миллионов до нескольких десятков миллионов шагов. Данный процесс является итеративным. На каждой следующей итерации, в соответствии с результатами предыдущей итерации, рассматривается либо сдвинутый, либо более узкий интервал значений для получения значения критического порога с необходимой точностью.

Генерация траекторий без применения техник параллелизации – процесс ёмкий по времени выполнения. Для сети с параметрами $N = 256, p = 0.3$ и траектории длиной 10 миллионов шагов такая генерация длится порядка часа (все приведенные эксперименты проведены с использованием процессора Intel Core i7 3960X). Рассмотрим примененные техники на примере приближенного времени расчета критического значения параметра в случае сети с упомянутыми параметрами. Предположим проводится 3 итерации упомянутого алгоритма и на каждом шаге рассматривается 10 траекторий. Такой процесс позволит получить значение

параметра с точность порядка нескольких сотых долей единицы. Кроме того, получение критической точки для фиксированных значений входных параметров необходимо подтвердить на ансамбле сетей. Предположим размер ансамбля равен 20, что является маленьким числом – реальные исследования требуют более 10^3 сетей.

Итак, необходимо последовательно сгенерировать 3 раза по 10 траекторий для 20 сетей (600 траекторий). Без какой-либо параллельности общее время расчетов составляет 600 часов (или, 25 дней). Ясно, что эта величина неприемлема для проведения исследований и необходима разработка и применение отвечающих временным требованиям вычислительных методов и программных решений. Основной целью разработки новой системы является быстрая генерация длинных траекторий для ансамбля сетей и получение всех необходимых свойств без каких-либо дополнительных манипуляций с данными и переключения между различными программными средами.

Первый подход, примененный в разработанной системе для решения данной задачи – одновременное выполнение генерации траекторий для различных значений параметра модели и для различных сетей. В приведенном примере – параллельная генерация 10 траекторий для 20 сетей (200 траекторий) на каждой итерации – позволит, в идеальном случае назначения каждому процессу отдельного ядра, сократить время генерации в 200 раз. Однако локальные компьютеры ресурсами таких объемов не обладают. В основном, кластеры также не обладают подобным количеством доступных ядер. Ввиду этого, в системе запроектирована ключевая возможность работы на заданном количестве ядер. Траектории равномерно распределяются по очередям, каждая из которых выполняется на выделенном для этой очереди ядре. Результат – при наличии k ядер, общее время выполнения сокращается в $\left\lceil \frac{200}{k} \right\rceil$ раз (в описанном случае). Учитывая независимость отдельных процессов по входным данным, можно запускать параллельные процессы на различных машинах. Известной и распространенной технологией, позволяющей выполнять подобное распределение задач по вычислительным сетям (кластерам) является стандарт MPI. Стандарт определяет основные аспекты управления процессов и методы организации взаимодействия между ними. Существуют различные библиотеки реализации стандарта и различные кластеры используют различные реализации. Для абстракции от конкретной реализации в разработанной системе и возможности легкого переноса кода с одной конфигурации на другую, система спроектирована с использованием оболочки, использующей только методы стандарта и имеющей возможность использовать различные библиотеки реализации без значительной потери производительности.

Применение упомянутой техники позволяет снизить время выполнения до 3 часов в случае назначения каждому набору параметров одного ядра. Однако, данная величина также является достаточно большой с точки зрения обычного пользователя.

На следующем этапе были рассмотрены параллельные технологии для увеличения скорости генерации одной траектории. Анализ технологии CUDA в контексте процесса направленного развития выявил сложности её применения ввиду последовательной природы процесса развития. Поэтому были детально проанализированы действия на каждом шаге направленного развития случайных сетей и были выявлены критические для производительности участки. Далее были разработаны соответствующие алгоритмы, позволяющие уменьшить общее время выполнения за счет уменьшения количества действий и параллельного выполнения некоторых из них. Результатом данных модификаций стало сокращение времени расчета одной траектории в рассматриваемом случае до 10-12 минут, что является ускорением в 5-6 раз.

Итак, финальное время вычисления критического порога в рассмотренном случае при наличии соответствующей аппаратной составляющей составляет 30-40 минут, и это является приемлемым временем расчета (учитывая использование ансамбля из 20 сетей). Зависимость времени выполнения от входных параметров представлена в разделе 3.7.

Второй этап исследования предполагает получение траекторий около критического порога и анализ свойств сетей. Для этого необходимо иметь состояние сети в необходимые моменты времени и, следовательно, соответствующая функциональность должна поддерживаться программной средой. Количество сетей, необходимых для проведения исследований, зависит от длины траектории и величины интервала времени, через которые необходимо сохранять состояние сети. Важным аспектом сохранения сетей является быстрота производимых действий – сохранение не должно значительным образом влиять на время генерации траекторий. Данное требование учтено при конструировании программной системы.

Объём вычислений на следующих этапах исследований меньше. Основные задачи – получение и графическое представление траекторий, свойств, сетей, матриц смежности, спектров и спектральных плотностей матриц смежности. Данные возможности реализованы в системе при помощи применения известных алгоритмов, адаптированных и модифицированных для работы с остальными частями системы.

```
1: PROCEDURE Randomization(G; m)
2:   trajectory = []
3:   prop = CalculateObservedProperty(G)
4:   trajectory.append(prop)
5:   WHILE !Stabilized(trajectory) DO
6:      $D_{prop} = \text{RandomizationStep}(G; m)$ 
7:     prop = prop +  $D_{prop}$ 
8:     trajectory.append(prop)
9:   RETURN trajectory
```

Алгоритм 1. Общий алгоритм рандомизации

Раздел 3.3 описывает детальный анализ вопросов применения некоторых параллельных и распределенных технологий, в частности обосновывается применение технологии MPI.

Раздел 3.4 содержит описание разработанных алгоритмов моделирования процесса эволюции для двух рассмотренных вариантов переключения связей. Алгоритм 1 представляет псевдокод общего алгоритма направленного развития сетей; алгоритм 2 приводит последовательность действий одного шага процесса в случае сохранения степеней вершин.

```
1: PROCEDURE RandomizationStep(G; m)
2:   DO
3:     e1 = GetRandomEdge(G)
4:     e2 = GetRandomEdge(G)
5:     WHILE EdgeExists(e1.source; e2.source; G) OR EdgeExists(e1.target; e2.target; G)
6:       RemoveEdge(e1; G)
7:       RemoveEdge(e2; G)
8:       D = CalculatePropertyChange(G; e1; e2)
9:       AddEdge(e1.source; e2.source; G)
10:      AddEdge(e1.target; e2.target; G)
11:      p = CalculateProbability(D; m)
12:      IF 0 == BernoulliDistribution(p) THEN
13:        RevertChanges(G)
14:        D = 0
15:      RETURN D
```

Алгоритм 2. Шаг процесса эволюции с сохранением степеней

Далее представлены ключевые детали процесса реализации, наладки и оптимизации различных фрагментов; представлена архитектуры системы, описаны её уровни и составляющие модули и механизмы взаимодействия между ними. Детально представлена архитектура ядра системы и представлены потоки выполнения и данных. В конце главы приводятся временные характеристики и особенности разработанной системы (рис. 3).

Глава 4 диссертационной работы посвящена основным результатам моделирования процесса направленного развития. Глава состоит из трех разделов.

Раздел 4.1 содержит описание основных этапов проведенных исследований. Выделяются три основных этапа:

1. Получение и исследование траекторий для сетей с различными параметрами.
2. Исследование структуры и структурных изменений в наблюдаемых сетях.
3. Исследование структуры и спектра матриц смежности сетей.

Раздел 4.2 представляет основные результаты каждого из этапов, описанных в разделе 4.1. Далее представлены основные результаты для случая с сохранением степеней вершин.

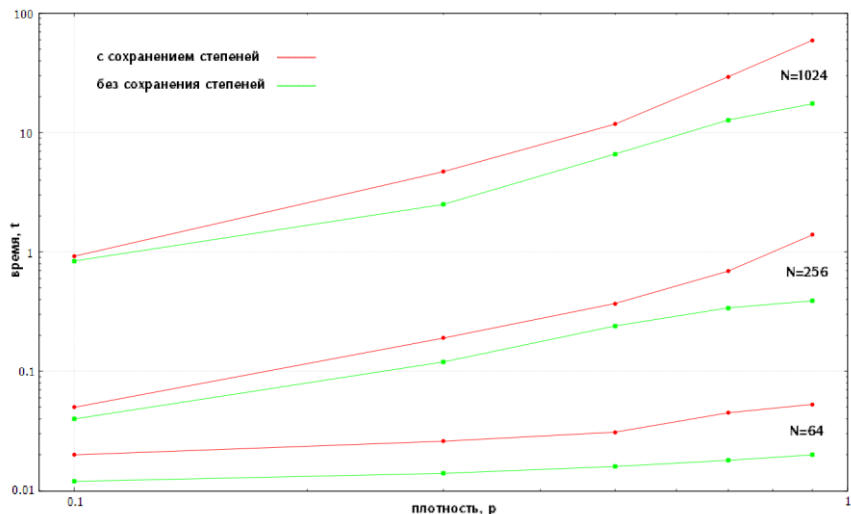


Рисунок 3. График зависимости времени выполнения от входных параметров

Компьютерное моделирование показало, что при значениях, больших критического порога – $\mu \geq \mu_c$, эволюция с сохранением степеней приводит к последовательному образованию множества плотных кластеров. Образование кластеров сопровождается увеличением среднего коэффициента кластеризации и среднего расстояния между узлами.

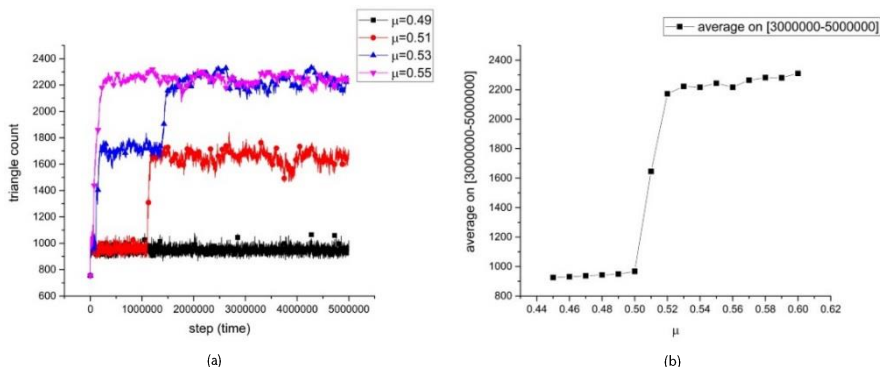


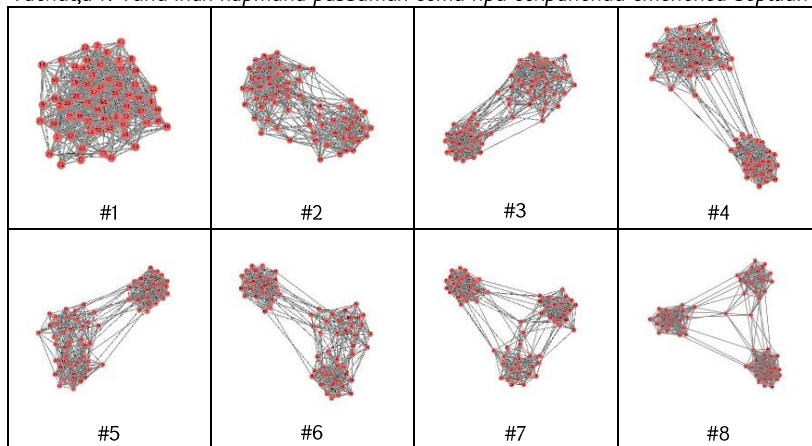
Рисунок 4. Эволюция случайной сети ($N=64$, $p=0.3$) с сохранением степеней вершин. (a) – Траектории для $\mu=0.49$, $\mu=0.51$, $\mu=0.53$, $\mu=0.55$. Переход наблюдается при $\mu=0.51$ и $\mu=0.53$. (b) – Равновесное значение $T(G)$ как функции от μ .

Далее показаны типичные результаты разделения изначально однородной сети в множество кластеров на примере сети размером 64. Такие же результаты наблюдаются и для сетей большего размера (128, 256 и 512 узлов). Начальные сети были сгенерированы случайным образом при вероятности связей $p = 0.3$, что намного выше порога перколяции.

На рис. 4(a) показаны эволюционные траектории, полученные из одной и той же начальной сети G при различных значениях параметра μ ниже и выше критической точки μ_c . Внимания заслуживает тот факт, что некоторые траектории обладают ступенчатыми переходами. Рис. 4(b) показывает равновесное значение $T(G)$ как функцию от μ . Здесь ясно виден скачок значения $T(G)$ вблизи $\mu = 0.5$.

Траектории показывают, что при значениях параметра, больших критического порога, например, при $\mu = 0.53$, происходит ряд переходов, которые сопровождаются увеличением количества циклов длины три. Данное явление объясняется структурным анализом сетей, сгенерированных по ходу процесса эволюции (таблица 1).

Таблица 1. Типичная картина развития сети при сохранении степеней вершин



Структуризация сети над критическим порогом обусловлена разделением сети на плотные связные компоненты. На проиллюстрированном примере, сеть сначала разделяется на два кластера, один из которых меньше, но плотнее другого, и содержит узлы с большими степенями. На следующем этапе эволюции, плотный кластер почти не подвергается каким-либо структурным изменениям, а больший второй кластер расщепляется на два меньших кластера. Дальнейшее развитие к образованию новых кластеров не приводит. На последнем рисунке (рисунок #8) показана финальная структура рассмотренной сети – три маленьких, плотных кластера связанные между собой гораздо (на порядок) слабее, чем внутри кластеров.

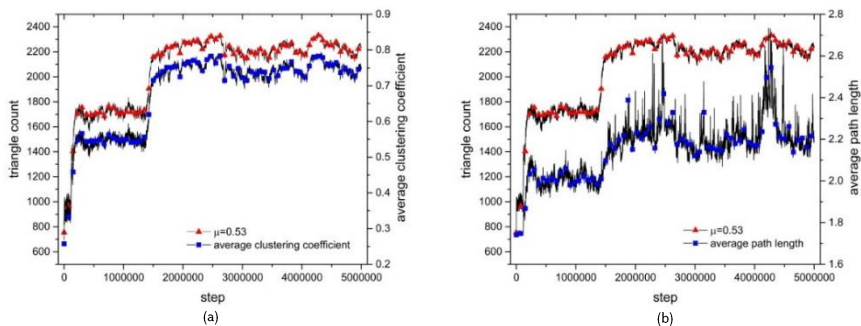


Рисунок 5. Сравнение различных топологических свойств в случае эволюции с сохранением степеней. Верхние кривые соответствуют количеству циклов длины три; нижние кривые соответствуют среднему коэффициенту кластеризации (a) и среднего расстояния между вершинами (b), соответственно.

Кроме эволюционных траекторий были изучены изменения в топологических характеристиках сети по ходу образования кластеров. Далее представлены данные, соответствующие траектории, для которой структурная картина представлена в таблице 1. Рис.5 представляет сравнение изменения количества циклов длины три по времени и соответствующие изменения среднего расстояния между вершинами и коэффициента кластеризации сети. Ясно видно, что обе приведенные характеристики сети коррелируют с изменением количества циклов длины три в сети – все три величины ведут себя синхронно.

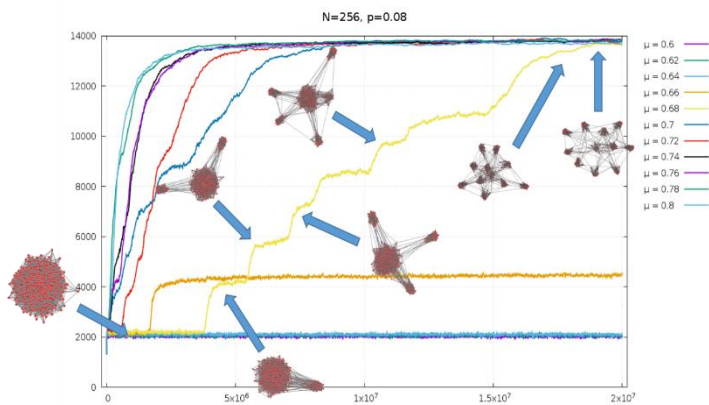


Рисунок 6. Картина структурообразования для случая $N=256$, $p=0.08$

Эволюция с сохранением степеней вершин приводит к реструктуризации сети и появлению кластерной структуры, причем количество кластеров обратно пропорционально вероятности связи p (рис. 6). Данный процесс можно рассматривать как образование многомасштабной архитектуры (в данном случае -

двухуровневой). На первом уровне находятся плотные кластеры со слабыми межкластерными связями; на втором уровне узлами являются кластеры, а связи уже представляют связи между этими вершинами. Данная архитектура является отображением некой оптимальной структуры, необходимой для достижения поставленной цели – максимального количества циклов длины три в сети при наличии «вмороженного» беспорядка.

Была выполнена *визуализация матриц смежности* в виде квадратной сетки с закрашенными клетками, соответствующими связям в сети. Матрицы получены по ходу процесса с определенным интервалом.

Случай с сохранением степеней обладает следующей картиной. На диагонали образуются примерно равные по размерам блоки, каждый из которых представляет отдельный кластер (рис. 7). Блоки, как и в первом случае, почти полностью закрашены, что свидетельствует о большой плотности связей в блоках. Вне блоков очень мало закрашенных клеток и это свидетельствует о том, что связи между блоками слабые.

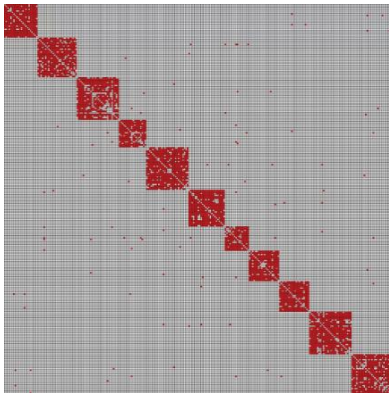


Рисунок 7. Финальная структура матрицы смежности в ходе процесса эволюции с сохранением степеней вершин.

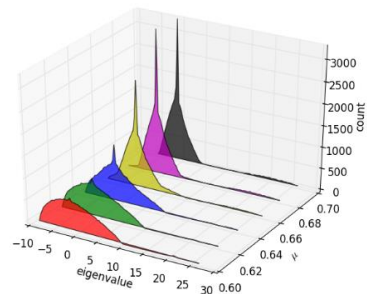


Рисунок 8. Динамика изменения спектральной плотности в случае сохранения степеней вершин

Визуализация полученных для различных значений параметра μ объединенных спектров показала следующее (динамика спектральной плотности изображена на рис. 8):

- основная зона финальной спектральной плотности обладает *треугольной формой*, специфичной для масштабно-инвариантных сетей.
- вторая зона – зона отделенных собственных значений, преобразуется из скопления нескольких значений вокруг одной точки в *обширное плато*, которое по размерам близко к основной зоне, однако обладает значительно меньшей плотностью. Данная зона отображает кластерную структуру финальных сетей – как было сказано выше, оторванные собственные

значения образуются параллельно с отщеплением кластеров от основной массы вершин в сети.

Для понимания природы наблюдаемого изменения формы спектральной плотности были рассмотрены спектральные плотности каждого из образовавшихся в сети кластеров. Эксперименты показали, что спектральная плотность матриц смежности каждого из кластеров подчиняется закону полуокружностей Вигнера. Следовательно, изменение формы спектральной плотности обеспечивается межкластерными связями.

Заключение содержит краткое описание полученных в работе результатов и их возможные применения.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ ДИССЕРТАЦИОННОЙ РАБОТЫ

1. Разработаны алгоритмы и программные методы моделирования направленных процессов в сетях [3].
2. Разработана программная система, являющаяся универсальным инструментом для проведения экспериментальных исследований и разностороннего анализа процесса направленного развития сетей [1,3].
3. Исследован процесс направленного развития сетей с сохранением степеней и без сохранения степеней; получены основные характеристики сетей и их динамики в ходе двух рассматриваемых вариантов процесса развития сетей и проведен их сравнительный анализ. Получена структурная картина, образующаяся в сетях при направленном развитии [1,2].
4. Исследованы матрицы смежности рассматриваемых сетей, получена общая картина наблюдаемой динамики. Исследованы спектры и спектральные плотности матриц смежности сетей [2].

Выражаю глубокую благодарность профессору В.А. Аветисову за постановку задачи, ценные советы, продуктивные обсуждения и помощь в процессе работы над диссертацией.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Hovhannisyan M., Avetisyan S. Clustering of random networks under topological constraints // 2015 Computer Science and Information Technologies (CSIT), Yerevan, 2015, pp. 156-159, DOI: 10.1109/CSITechnol.2015.7358272
2. Avetisov V., Hovhannisyan M., Gorsky A., Nechaev S., Tamm M., Valba O. Eigenvalue tunneling and decay of quenched random network // Physical Review E, vol. 94, issue 6, id 062313, pp. 1-6, 2016. DOI: 10.1103/PhysRevE.94.062313
3. Hovhannisyan M. Parallel technologies in some problems of network modelling // Proceedings of the YSU, Physical and Mathematical Sciences, 2017, v. 51, #1, pp. 53-59

ԱՄՓՈՓՈՒՄ

Մինաս Հ. Հովհաննիսյան

«Բազմամասշտաբային ճարտարապետությամբ բարդ ցանցային համակարգերի նպատակային նախագծումը բաշխված հաշվողական ռեսուրսների օգտագործմամբ»

Սույն ատենախոսությունը նվիրված է պատահական ցանցերի ուղղորդված զարգացման պրոցեսի մոդելավորման ծրագրային մեթոդներին և նրանց կիրառությանը պրոցեսի հետազոտության համար: Աշխատանքը ներկայացնում է մշակված ալգորիթմները և ծրագրային համակարգը և նրա միջոցով կատարված զարգացման պրոցեսի բազմակողմանի հետազոտությունների արդյունքները:

Ատենախոսության նպատակն է պատահական ցանցերի տոպոլոգիական սահմանափակումներով ուղղորդված զարգացման պրոցեսի մոդելավորման համար նախատեսված ալգորիթմների և ծրագրային մեթոդների մշակումը և դրանց օգտագործումը նշված պրոցեսի բազմակողմանի հետազոտության համար:

Աշխատանքի նորույթը կայանում է ցանցերի ուղղորդված զարգացման պրոցեսի մոդելավորման համար մշակված ալգորիթմների և ծրագրային մեթոդների մեջ: Նրանց հիման վրա ստեղծված ծրագրային համակարգը հնարավորություն է տվել կատարել Էդդշ-Ռենյիի պատահական ցանցերի տոպոլոգիական սահմանափակումներով ուղղորդված զարգացման պրոցեսի փորձարարական հետազոտություններ, որոնց արդյունքում ստացված են հետևալ նոր արդյունքները.

1. Դիտարկվող պրոցեսի ընթացքում տեղի է ունենում ֆազային անցում, որը ուղեկցվում է ցանցի կառուցվածքի փոփոխությամբ՝ ցանցը տրոհվում է մի քանի խիտ կլաստերների, որոնց միջև կապերը թույլ են:
2. Ցանցերի կցության մատրիցներում դիտարկվո առաջանում է բլոկ-անկյունագծային կառուցվածք:
3. Սպեկտրալ խտության տեսքը փոփոխվում է՝ կիսաշրջանաձևից դեպի եռանկյունաձև:

Աշխատանքի կառուցվածքը հետևյալն է՝ ներածություն, չորս գլուխ և եզրակացություն:

Ներածությունը համառոտ նկարագրում է դիտարկվող ոլորտը և աշխատանքում ստացված արդյունքները:

Առաջին գլուխը ներկայացնում հետազոտվող ոլորտի հիմնական խնդիրները, հիմնավորում է դիտարկվող պրոցեսի արդիականությունը, նրա ծրագրային մոդելավորման համար գոյություն ունեցող համակարգերը և նրանց թերությունները:

Երկրորդ գլուխը ներկայացնում է հետազոտությունների տեսական հիմունքները և օգտագործված ծրագրային տեխնոլոգիաների նկարագրությունը:

Նկարագրված է պատահական գրաֆների Էդդո-Ռենյի մոդելը և ցանցերի հիմնական հատկությունները: Ներկայացված են ցանցերի պատկերման ալգորիթմները: Նաև նկարագրված է Մետրոպոլիս-Հաստինգսի ալգորիթմը, որը հիմք է կազմում ուղղորդված զարգացման պրոցեսի: Հաջորդիվ ներկայացված է ուղղորդված զարգացման պրոցեսի մոդելը: Այնուհետև բերված են դիտարկված ծրագրավորման զուգահեռ հիմնական տեխնոլոգիաների համառոտ նկարագրությունները:

Երրորդ գլուխը ներկայացնում է հետազոտությունների անցկացման համար անհրաժեշտ ծրագրային միջավայրի նկատմամբ առաջադրվող պահանջները և հիմնավորում է նոր ծրագրային համակարգի մշակման անհրաժեշտությունը: Այնուհետև բերված է մշակված ծրագրային համակարգի կառուցվածքը, բաղկացուցիչ մասերի և մոդելավորման համար նախագծված ալգորիթմների նկարագրությունները և տրամադրվող ֆունկցիոնալությունները:

Չորրորդ գլուխը պարունակում է մշակված ծրագրային համակարգով կատարված հետազոտությունների հիմնական արդյունքները: Սա իր մեջ է ներառում չորս հիմնական պատկերների նկարագրությունը կապերի փոփոխման դիտարկված երկու ձևերի համար՝

1. Էվոլյուցիոն հետազոտի և հիմնական հատկությունների համեմատական ներկայացում, ցանցերում կատարվող կառուցվածքային փոփոխությունների նկարագրում:
2. Ցանցերում կառուցվածքային պատկերի ձևավորման օրինաչափությունների ներկայացում:
3. Կցության մատրիցների կառուցվածքի ներկայացում:
4. Սպեկտրալ խտության տեսքի փոփոխության պատկերի ներկայացում:

Եզրակացությունը ներկայացնում է կատարված աշխատանքի համառոտ նկարագրությունը, այն է՝

1. Մշակված են ուղղորդված պրոցեսների մոդելավորման ալգորիթմներ և ծրագրային մեթոդներ:
2. Մշակված է ուղղորդված պրոցեսների բազմակողմանի հետազոտության համար նախատեսված ծրագրային համակարգ:
3. Հետազոտված է ուղղորդված զարգացման պրոցեսը և ստացված են ցանցերի հիմնական հատկությունները պրոցեսի ընթացքում: Հետազոտված է ցանցերի կառուցվածքային պատկերի ձևավորումը:
4. Հետազոտված է ցանցերի կցության մատրիցների կառուցվածքի դինամիկան: Հետազոտված է ցանցերի կցության մատրիցների սպեկտրը և սպեկտրալ խտությունը:

Վերջում բերված են ստացված արդյունքների հնարավոր կիրառությունների նկարագրությունը:

ABSTRACT

Minas H. Hovhannisyan

“Target designing of complex network systems with multi-scale architecture using distributed computing resources”

The thesis is devoted to the programmatic methods of modelling of the process of directed evolution of networks and their usage for the research of considered evolution process. This work presents the developed algorithms and software system and the results of comprehensive analysis of the directed evolution process performed with the system.

The aim of the thesis is the development of algorithms and programmatic methods for modeling of directed evolution process with topological constraints and their application for in-depth analysis of the evolution process.

Developed algorithms and programmatic methods of directed evolution modelling represent **the novelty of the thesis**. Developed software system allowed to perform experimental research of directed evolution of Erdős-Rényi random networks with topological constraints, which resulted in discovering of following new phenomena:

1. Directed evolution with node degree conservation shows phase transition, which is accompanied with network restructure – network disintegrates into a set of dense clusters with weak inter-cluster links.
2. Block-diagonal structure is emerging in adjacency matrices of networks.
3. Spectral density shape changes from semicircle to triangle.

The structure of the thesis is following: introduction, four chapters and conclusion.

Introduction presents the overview of the area of research and the results, achieved in this work.

First chapter presents main problems, which are considered in network theory and presents the actuality of the process of directed evolution of networks. After, available software systems are presented and the major problems of their application for modelling the described process.

Second chapter presents theoretical basics of research process and description of used programmatic technologies. Erdős-Rényi random graphs model is described along with the description of the main properties of networks. Network visualization and community detection algorithms are presented. Depicted the Metropolis-Hastings algorithm, which is the basis of directed evolution process. Next, the model of directed evolution process is described. Finally, the major parallel and distributed technologies are presented in short.

Third chapter presents the necessary preconditions for software system to perform required research and states the necessity for development of new system with application of parallel and distributed calculations. Following, the architecture of system is presented and the structural components are depicted. After this, the developed algorithms for the evolution modelling are presented. Lastly, the key features of system are described.

Fourth chapter contains the main results of conducted research. This includes four major topics for two variants of link switching.

1. Presentation of evolutionary trajectories and comparison with the main properties of networks. Description of the structural changes, observed in networks during the evolution.
2. Description of pattern formation in networks.
3. Presentation of structure of the adjacency matrices of networks.
4. Presentation of spectra and spectral density shape dynamics.

Conclusion presents short description of performed work:

1. Algorithms and programmatic methods are developed for modelling directed processes in networks.
2. Software system is developed to allow performing experimental research and comprehensive analysis of directed evolution process of networks.
3. Analysis of directed evolution process is performed. Main properties of networks during evolution process are retrieved and analyzed. Analysis of network structure is performed and the pattern formation is described.
4. Analysis of the structure of network adjacency matrices is performed. The spectra and the spectral density of adjacency matrices are analyzed.

Finally, possible applications of observed phenomena are described.