

## **ОТЗЫВ**

официального оппонента на диссертацию

Дробышевского Михаила Дмитриевича

на тему: «Методы и программные средства моделирования и генерации

сложных сетей с сохранением графовых свойств»

по специальности 05.13.11 «Математическое и программное обеспечение

вычислительных машин, комплексов и компьютерных сетей»

на соискание ученой степени кандидата физико-математических наук

### **Актуальность темы**

Графовые структуры данных являются часто используемой моделью описания данных в различных предметных областях. Одними из наиболее востребованных объектов в цифровой экономике являются графы социальных взаимодействий: связи между пользователями социальных медиа, графы взаимных звонков в мобильных сетях и т.п. При этом возникающие графы имеют высокое разнообразие, динамическую изменчивость, ограничения конфиденциальности, что затрудняет исследование свойств реальных графов. Поэтому для исследования используются случайные графы, моделирующие свойства реальных.

Диссертационная работа Дробышевского Михаила Дмитриевича посвящена исследованию методов построения случайных графов, которые близки моделируемым графам в смысле близости выбранных характеристик, включая характеристики направленности, весов на ребрах, структуры сообществ. При этом рассматривается также задача порождения разнообразных графов, что важно для моделирования разнородных фрагментов реальных графов, оценки устойчивости методов анализа сложных графов.

В диссертационной работе автором предложен подход ERGG (от англ. - Embedding based Random Graph Generation) - генерация случайных графов на основе вложения в векторное пространство меньшей размерности по сравнению с размерностью множества вершин исходного графа. На основе данного подхода предложен метод ERGG-dwc (от англ. - directed, weighted, communities), предназначенный для сохранения характеристик направленности ребер, весов ребер, структуры сообществ исходного графа.

## **Достоверность и степень обоснованности научных положений, выводов и рекомендаций, сформулированных в диссертации**

Достоверность исследования основана на проведении теоретического обоснования свойств предложенного алгоритма и получающихся в результате графов, а также путем проведения вычислительного эксперимента с использованием разработанной программной системы для анализа графов различной структуры.

Все положения, выводы и рекомендации диссертации имеют достаточно строгое обоснование. Обоснованность научных и практических результатов диссертации определяется проведенным автором подробным анализом публикаций по разрабатываемой теме. Приведенный в диссертации список проанализированных источников включает основные работы отечественных и зарубежных авторов в данной области.

## **Новизна исследования, полученных результатов, выводов и рекомендаций, сформулированных в диссертации**

Следующие основные результаты диссертации являются новыми:

1. Предложен новый подход ERGG к генерации случайных направленных графов, похожих на данный, основанный на вложении графа в пространство размерности, много меньшей числа его вершин.

2. В рамках подхода ERGG предложен метод ERGG-dwc, решающий задачу генерации графов, похожих на данный и удовлетворяющих требованиям: автоматическое обучение на заданном графе, контролируемый размер генерируемых графов, поддержка направленных ребер, взвешенных ребер и структуры сообществ.

Соответствие метода ERGG-dwc требованиям похожести и вариабельности генерируемых графов подтверждено экспериментально, также показана корректность назначения структуры сообществ и весов ребер. Доказаны теоремы о вычислительной сложности и масштабировании. Налагаемые ими ограничения не являются существенными для применимости метода.

3. Создана программная система, в которой реализован прототип ERGG-dwc и проведено его экспериментальное сравнение с другими современными методами. Показано, что ERGG-dwc не уступает другим методам в похожести генерируемых графов, но превосходит их по вариабельности.

Работа представляет собой первое диссертационное исследование, когда техника вложения графов используется не только для анализа свойств графов, но и для порождения множества случайных графов, моделирующих исходный.

### **Значимость для науки и практики полученных автором результатов**

**Теоретическая значимость** исследования заключается в предложенном структурировании сложной и разнообразной области генерации случайных графов, предложенном и теоретически обоснованном новом методе генерации случайных графов на основе использования представления анализируемых графов во вложенном пространстве.

**Практическая значимость** исследования заключается в доказанной на практике возможности реализации предложенных методов в программной системе, позволяющей анализировать свойства реальных графов, в том числе большого размера.

### **Конкретные рекомендации по использованию результатов и выводов диссертации**

Полученные результаты могут найти применение при решении задач анализа реальных графов, прежде всего графов социальных взаимодействий, не нарушая при этом конфиденциальность участников социальных взаимодействий.

Кроме того, аналогичные задачи встречаются в биологии (метаболические и пищевые цепочки), наукометрии (графы цитирований), автономных системах (сеть Интернет, граф взаимодействия компонентов программ) и т. д.

### **Оценка содержания диссертации, ее завершенность.**

Диссертация написана ясным и четким языком и состоит из введения, трех глав, заключения и двух приложений. Библиография включает 179 наименования.

**В первой главе** дается обширный обзор существующих подходов к генерации случайных графов, в том числе рассмотрены имеющиеся в литературе обзоры, классификационные подходы. Предложена собственная система классификации подходов к моделированию случайных графов.

**Во второй главе** предложен подход ERGG построения случайных графов с использованием техники моделирования исходного графа путем вложения в пространство меньшей размерности, когда характеристики графа кодируются векторами, ассоциированными с вершинами графа.

Описан алгоритм порождения случайных графов ERGG-dwc, который имеет цель сохранения таких характеристик исходного графа, как направленности ребер, весов на ребрах, структуры сообществ. Данный алгоритм базируется на предложенном методе COMBO получения вложения графа, который является синтезом известных в литературе методов BLM и LINE. Метод позволяет строить случайные графы с масштабированием относительного исходного.

Приведено теоретическое обоснование свойств предложенного метода. Доказано утверждение (Теорема 1, страницы 91-92) о квадратичном характере роста количества ребер в порождаемых графах. Также приведена теоретическая оценка (Теорема 2, раздел 2.4.4, страницы 94-95) вычислительной сложности алгоритма ERGG-dwc.

**В третьей главе** описана программная реализация алгоритма ERGG-dwc и представлено экспериментальное исследование, цель которого заключается в проверке возможностей формируемых случайных графов моделировать характеристики реальных графовых структур из различных предметных областей.

Определены оптимальные параметры метода вложения COMBO, измерена производительность алгоритма ERGG-dwc.

Проведены численные эксперименты по сравнению предложенного алгоритма ERGG-dwc с другими методами для различных типов графов с оценкой известных графовых характеристик (степень вершин, количество подграфов, связность, спектральные характеристики). В результате проведения вычислительных экспериментов было установлено, что предложенный метод ERGG-dwc принципиально не уступает известным методам по возможности воспроизведения рассматриваемых графовых характеристик, при этом показывает значительно лучшие результаты по возможности моделировать вариативность реальных графов, которые часто имеют разнородную структуру.

Диссертацию Дробышевского М.Д. можно квалифицировать как завершённую научно–исследовательскую работу, содержащую новые результаты, имеющие научную и практическую ценность, вносящие заметный вклад в разработку методов анализа сложных графовых структур, встречающихся в различных сферах деятельности.

### **Достоинства и недостатки в содержании и оформлении диссертации**

В качестве **замечаний** к тексту диссертации следует отметить следующие:

1. Автору следовало бы более разнообразно подойти к используемым обозначениям, так при доказательстве Леммы 1 (раздел 2.4.4 на странице 94) одновременно встречаются схожие по написанию факторы, имеющие существенно разный смысл:  $d$  – размерность пространства вложения и  $d_i^{out}$  – исходящая степень вершины.
2. В диссертации используется термин «домен», вместо которого целесообразно было бы использовать более подходящий термин «предметная область».
3. Список сокращений и условных обозначений (страницы 127-128) не отсортирован по алфавиту.
4. Несколько нетипичным является использование при изложении алгоритма стиля формулировки вопросительного высказывания и дальнейшего описания соответствующего фрагмента алгоритма в виде ответа на поставленный вопрос.
5. Приведенные графики (Глава 3, Приложения А и Б) недостаточно хорошо различимы при черно-белой печати.

Указанные недостатки не являются принципиальными и не умаляют достоинств диссертации.

### **Заключение о соответствии диссертации критериям, установленным Положением о порядке присуждения ученых степеней**

Результаты диссертационной работы представлены в 3 печатных работах, опубликованных в ведущих рецензируемых изданиях, входящих в перечень ВАК, а также в одном патенте.

Материал диссертации изложен последовательно и логично. Структурные составляющие диссертационной работы (введение, главы, заключение, библиогра-

фический список, приложения) позволяют получить полное представление о проделанных исследованиях и полученных результатах.

Автореферат соответствует диссертации, отражает её содержание и дает представление об актуальности темы, целях, задачах, объекте и методах исследования, научной новизне, практической ценности, реализации, апробации, объеме, кратком содержании и результатах работы.

Таким образом, диссертация Дробышевского Михаила Дмитриевича является научно-квалификационной работой, в которой **содержатся подходы к решению задачи** моделирования и генерации сложных сетей с сохранением графовых характеристик, имеющие значение для развития вычислительной математики, что соответствует требованиям п. 9 «Положения о порядке присуждения ученых степеней», утвержденного постановлением Правительства Российской Федерации от 24.09.2013 г. № 842, предъявляемым к диссертациям на соискание ученой степени кандидата наук, а ее автор заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 05.13.11 «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Официальный оппонент,  
Добров Борис Викторович, к.ф.-м.н.,  
заведующий лабораторией анализа информационных ресурсов  
Научно-исследовательского вычислительного центра  
МГУ имени М.В. Ломоносова  
119899, Москва, Ленинские горы 1, стр. 4

Добров Б.В., 1

Дата « 25 » ноября 20 19 г.