

Отчёт о проделанной работе на кластере ИВЦ НГУ

1. Тема работы:

Разработка и исследование методов эволюционного синтеза нелинейных математических моделей на основе темплейтов.

2. Состав коллектива:

Монахов Олег Геннадьевич, ИВМиМГ СО РАН, в.н.с., к.т.н., руководитель,
Монахова Эмилия Анатольевна, ИВМиМГ СО РАН, с.н.с., к.т.н., исполнитель,
Токтошов Гулжигит Ысакович, ИВМиМГ СО РАН, м.н.с., к.т.н., исполнитель.

3. Информация о гранте:

«Разработка и исследование методов эволюционного синтеза нелинейных математических моделей на основе темплейтов с реализацией на суперЭВМ».
Проект РФФИ № 14-01-00031. Руководитель - к.т.н. Монахов О.Г., срок действия: 2014-2016.

4. Научное содержание работы:

а. Постановка задачи.

В рамках проекта будут разработаны и исследованы новые, основанные на эволюционных вычислениях, методы автоматического синтеза нелинейных математических моделей на основе экспериментальных данных. Эффективность данного подхода будет теоретически исследована и экспериментально апробирована на суперЭВМ при синтезе нелинейных математических моделей, возникающих при описании природных процессов и при проектировании инфокоммуникационных сетей связи.

б. Современное состояние проблемы.

В рамках проекта рассматривается проблема автоматического синтеза описаний нелинейных математических моделей природных процессов и технических объектов на основе экспериментальных данных. Данная проблема решается известными методами теории индуктивного вывода, методами группового учета аргументов, методами теории искусственных нейронных сетей, методами машинного обучения и установления закономерностей, обобщения и вывода на знаниях, а также одним из перспективных направлений в данной области - методами эволюционных вычислений и генетического программирования. Традиционный подход генетического программирования страдает от слабо-ограниченного слепого поиска в огромных пространствах параметров для реальных прикладных проблем и имеет большие временные затраты при увеличении сложности решаемых задач. Поэтому перспективным путем сокращения трудоемкости автоматического синтеза нелинейных математических моделей является использование типовых алгоритмических структур (темплейтов, шаблонов, скелетонов) и гибридизации методов генетического программирования и генетических алгоритмов. Другим перспективным путем сокращения времени поиска решения при автоматическом эволюционном синтезе является распараллеливание эволюционных вычислений, позволяющее значительно сократить время исполнения и получить линейное ускорение на высокопроизводительных вычислительных системах.

в. Подробное описание работы, включая используемые алгоритмы.

Для решения проблемы построения нелинейных моделей (математических выражений, функций, алгоритмов, программ) на основе заданных экспериментальных данных, множества переменных, базовых функций и операций разработан метод и алгоритм многовариантного эволюционного синтеза нелинейных моделей, который имеет (1) линейное представление хромосомы, (2) модульные операции при декодировании генотипа в фенотип для интерпретации хромосомы как последовательности команд, (3) многовариантный метод для представления множества моделей (выражений) с помощью одной хромосомы, (4) векторизацию вычислений при определении целевой (фитнес) функции.

d. Полученные результаты.

Проведено исследование предложенного метода многовариантного эволюционного синтеза для синтеза нелинейных математических моделей природных процессов, описываемых системами обыкновенных дифференциальных уравнений и дифференциальными уравнениями в частных производных, включая (1) разработку и исследование темплейтов для представления нелинейных математических моделей, (2) разработку оригинальных методов спецификации и представления данных для этих темплейтов и (3) разработку критериев отбора лучших решений для эволюционного алгоритма синтеза. Получены оценки влияния характеристик темплейтов, таких как степень специализации темплейта и количество неопределенных операторов в нем, на показатели эффективности эволюционного алгоритма синтеза рассматриваемого типа моделей. Предложенный подход с использованием темплейтов позволил получить существенное сокращение времени поиска аналитических описаний нелинейных математических моделей по экспериментальным данным (в несколько раз, в зависимости от степени специализации темплейта). Получены оценки эффективности предложенного метода и проведено его сравнение со стандартным алгоритмом генетического программирования, алгоритмом грамматической эволюции и алгоритмом декартового генетического программирования. Проведенные эксперименты показали существенное преимущество предложенного подхода по сравнению с указанными алгоритмами как по времени поиска решения (более чем на порядок в большинстве случаев), так и по вероятности нахождения заданной модели (во многих случаях более чем в два раза). Проведены эксперименты по распараллеливанию предложенного метода многовариантного эволюционного синтеза нелинейных моделей и реализации его на параллельных суперкомпьютерных системах и получены оценки эффективности предложенного параллельного алгоритма, демонстрирующие линейные ускорение и масштабируемость.

Предложенный эволюционный метод синтеза был успешно применен к решению экстремальных задач комбинаторной оптимизации для сетей связи вычислительных систем, моделируемых параметрически описываемыми регулярными графами. С помощью эволюционного метода синтеза экспериментально были получены аналитические описания более 70 семейств оптимальных циркулянтных графов (сетей) степени четыре, ранее неизвестные в литературе. Разрабатываемый метод эволюционного синтеза позволил получить также оптимальные описания для введенного исполнителями проекта класса гиперциркулянтных графов (сетей). При проведении сравнительного анализа структурных характеристик

гиперциркулянтных, тороидальных и циркулянтных сетей показано преимущество полученных оптимальных гиперциркулянтных сетей по таким структурным характеристикам, как диаметр, средний диаметр и ширина бисекции, при соизмеримых затратах на число узлов и число связей системы.

Эволюционный метод синтеза был успешно применен также к решению задачи построения больших циркулянтных инфокоммуникационных сетей, т.е. циркулянтных сетей (графов, структур), имеющих наибольшее число вершин при заданных степени вершин и диаметре. Данная задача является вариантом известной классической задачи - "Degree/Diameter problem". Для задачи построения больших циркулянтных сетей предложен и реализован параллельный алгоритм эволюционного синтеза таких циркулянтных сетей. Проведены вычислительные эксперименты на параллельных вычислительных системах ИВЦ Новосибирского Государственного Университета, получено линейное ускорение для предложенного алгоритма. Найдены новые циркулянтные сети, улучшающие порядки самых больших циркулянтов из таблицы рекордных значений в 28 случаях из 63 для разных степеней и диаметров, которые известны на время публикации по зарубежным источникам.

Таким образом, предложенный многовариантный эволюционный метод синтеза на основе темплейтов позволяет получать новые знания, в частности при решении сложных проблем построения нелинейных математических моделей, комбинаторной оптимизации, в теории графов и структур экзафлопных вычислительных систем.

е. Иллюстрации, визуализация результатов. Нет.

5. Эффект от использования кластера в достижении целей работы.

Проведенные эксперименты по распараллеливанию предложенного метода многовариантного эволюционного синтеза нелинейных моделей, реализация его на параллельных суперкомпьютерных системах и полученные оценки эффективности предложенного параллельного алгоритма демонстрируют линейные ускорение и масштабируемость.

6. Перечень публикаций, содержащих результаты работы.

1. Монахов О.Г., Монахова Э. А. Параллельный алгоритм многовариантного эволюционного синтеза нелинейных моделей // Сибирский журнал вычислительной математики. 2017. № 2. С. 169-180.
2. Монахова Э.А., Монахов О.Г. Поиск рекордных циркулянтных графов с использованием параллельного генетического алгоритма // Дискретный анализ и исследование операций.-Т. 22, № 6, 2015. С. 29-39.
3. Монахов О.Г., Монахова Э. А. Улучшение характеристик класса регулярных сетей с помощью алгоритма эволюционного синтеза // Наука и образование. №10, 2014. С.273-283.
4. Монахова Э. А. , Монахов О. Г. Эволюционный синтез семейств оптимальных двумерных циркулянтных сетей // Вестник СибГУТИ. 2014. № 2. С. 72-82.
5. Монахов О.Г., Монахова Э. А. О параллельном алгоритме многовариантного эволюционного синтеза нелинейных моделей. // Труды XII международной азиатской школы-семинара «Проблемы оптимизации сложных систем», г. Новосибирск, 12–16 декабря 2016 г. -- С. 151-156.

6. Монахов О.Г., Монахова Э.А. Параллельные алгоритмы эволюционного синтеза больших циркулянтных сетей с реализацией на суперЭВМ. // Труды Международной конференции "Актуальные проблемы вычислительной и прикладной математики 2015" (АПВПМ-2015), Новосибирск. 19-23 октября 2015г. - С. 501–505.

7. **Впечатления от работы вычислительной системы и деятельности ИВЦ НГУ.** Впечатления от работы с ИВЦ НГУ однозначно положительные, особенно следует отметить квалифицированные и своевременные консультации и помощь ведущего инженера Владислава Анатольевича Калюжного.