Шаблий Алексей Денисович

Исследование и разработка методов построения инструментальных

средств конфигурирования в плагинных системах

Специальность 2.3.5 - Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей

РЕФЕРАТ

диссертации на соискание ученой степени кандидата физико-математических наук

Москва — 2024

Работа выполнена в Московском авиационном институте (национальном исследовательском университете) на кафедре вычислительной математики и программирования.

|  |  |
| --- | --- |
| Научный руководитель: | **Судаков Владимир Анатольевич**, доктор технических наук, доцент, профессор кафедры вычислительной математики и программирования |
| Ведущая организация: | ПАО «ОАК» «ОКБ Сухого» |

**Общая характеристика работы**

**Актуальность темы исследования**

В настоящее время информационные технологии активно развиваются. Их развитие сопровождается неминуемым усложнением как самих информационных решений, так и технологий, применяемых для создания конечного продукта. Примером таких технологий являются инструментальные средства конфигурирования. С их помощью специалисты могут настраивать поведение информационных систем в допустимой конфигурации:

* для операционных систем - настройка системных ресурсов, энергозависимой и энергонезависимосй памяти, расписание работы компонентов системы;
* для прикладных программ - их поведение, правила журналирования своей работы, поддержка различных протоколов взаимодействия;
* для баз данных - управление доступом, шардирование и реплекация;
* и т.д.

В рамках одной предметной области объем решаемых задач разными специалистами может существенно отличаться. В следствие этого отличается и состав полезного функционала, используемого в инструментальном средстве конфигурирования.

Удовлетворение спроса на инструментальные программные средства с разным объемом функционала может быть реализовано поставкой решения в различных комплектациях.

Это может быть достигнуто применением технологии плагинных систем. В них результирующий объем функционала зависит от состава установленных расширений системы - плагинов.

Применение плагинов для решения задачи поставки функционала в различных комплектациях не лишено недостатков. Одним из недостатков является необходимость поставки бесполезного для заказчика функционала в следствие технической необходимости разрешения функциональных зависимостей.

Анализ показал, что на сегодняшний день применяются следующие методы построения решения в плагинной системе:

* разместить весь функционал в одном плагине - очевидно не обеспечивает возможность поставки функционала в различных комплектациях
* разместить каждую единицу функционала в своем плагине - при достаточно большом числе плагинов становится невожным управление ими
* разместить функционал по плагинам по некоторому критерию - существуют разные критерии и методы распределения функционала в соответствии с ними, в рамках диссертационного исследования сформулирован и описан один из них.

В рамках диссертационного исследования рассматривается задача поиска оптимальной декомпозиции функционала по плагинам. Критерием оптимальности распределения является минимизация стоимости постпродажного обслуживания (ППО). В рассматриваемой задаче стоимость ППО поставки, которая формируется в соответствии с заявленной комплектацией, зависит от состава вошедших в поставку требований. Стоимость ППО поставки рассчитывается как сумма стоимостей ППО каждого из реализованных в поставке требований. При этом стоимость ППО каждого отдельного требования зависит как от себестоимости его сопровождения так и от того, какие требования дополнительно были включены в поставку: включение их может как повышать, так и понижать стоимость.

**Степень разработанности темы исследования.**

В настоящее время существующие программные решения, выполненные в виде плагинных систем, в том числе инструментальные средства конфигурирования, могут быть поставлены в различных комплектациях с целью обеспечения конечного потребителя востребованным функционалом. Однако существующие решения разрабатываются с целью поставки в заранее сформулированных комплектациях и не учитываются разбиение уже существующих готовых решений. Кроме того они не устойчивы или слабо устойчивы к изменению состава этих заявленных комплектаций.

**Объект исследования.**

Инструментальные средства конфигурирования, осуществляющие работу в плагинной системе.

**Предмет исследования.**

Оптимальная декомпозиция функционала инструментального средства конфигурирования в плагинной системе по критерию минимизации бесполезного функционала в формируемой поставке заданной комплектации.

**Цель исследования**

Целью исследования является повышение эффективности поставляемых программных решений инструментальных средств конфигурирования в плагинных системах.

**Задачи исследования**

Для достижения поставленной цели были сформулированы и решены следующие основные задачи:

1. анализ предметной области плагинных систем для выявления значимых сущностей и описание их ограничений;
2. разработка методики построения математической модели, включающей целевую функцию и действующие ограничения на значения ее параметров;
3. решение задачи оптимизации в том числе и методами машинного обучения с последующим сравнением результатов решений;
4. вывод о целесообразности применения того или иного метода в вопросе решения конкретной практической задачи.

**Границы (рамки) исследования.**

Исследование ограничено функциональными возможностями имеющихся плагинных систем. А так же основывается на экспериментах, выполненных не на всех существующих плагинных системах, а на самых распространенных и актуальных на сегодняшний день решениях. Кроме того, в рамках диссертационного исследования используются существующие на сегодняшний день программные решатели оптимизационных задач с имеющимися у них достоинствами и недостатками.

**Научная новизна**

На сегодняшний день в диссертационном исследовании научная новизназаключается в следующем:

1. Разработан метод приведения выражений к линейному виду в математической модели распределения функционала в плагинной системе.
2. Сформулирована методика оценки стоимости ППО сформированной поставки в соответствии с заданной комплектацией.
3. Реализован комплекс программных решений применяемый как для проведения экспериментальных рассчетов гипотетических моделей, так и реальных, построенных на основе открытых данных.

**Теоретическая значимость работы**

Теоретическая значимость работы заключается в математической формализации задачи, в рамках которой осуществляется сведение к задаче на графе, описание ограничений на числовые значения элементов модели. Кроме того описан метод, позволяющий свести сформулированную задачу к линейному виду с целью увеличения числа решателей, которыми задача может быть решена.

**Практическая значимость**

Практическая значимость работы заключается в реализации предложенного метода формирования математической модели с вычисления стоимости ППО поставки, сформированной в соответствии с заданной комплектацией.

**Методология и методы исследования.**

Эксперименты с применением различных алгоритмов решения задачи оптимальной декомпозиции на графовых математических моделях. Сравнительный анализ результатов применения алгоритмов по критерию минимально необходимого объема компонентов плагинной системы выраженного в количественном отношении.

**Эмпирическая база исследования.**

Эмпирической базой исследования являются данные об объемах программных решений инструментальных средств конфигурирования: количество реализуемых ими требований, объем кодовой базы и количество плагинов, использованных в этих средствах.

**Основные положения, выносимые на защиту:**

* Сформулированные механизмы построения целевой функции и ограничений математической оптимизационной модели и формализация задачи.
* Результаты экспериментов по оптимальности применения того или иного механизма поиска оптимальной декомпозиции, а так же данные о достоверности результатов применяемого решения.

**Степень достоверности и апробация результатов.**

Основные положения диссертации обсуждались в научных конференций и отражены в статьях в рецензируемых журналах, включенных в список ВАК. Основные результаты диссертационного исследования также апробированы в ходе практической деятельности автора и внедрены на авиационном предприятии.

**Содержание работы.**

**Во введении** показана актуальность темы исследования, сформулированы его цели и задачи, новизна, а также достоверность и обоснованность результатов. Кратко охарактеризованы методы исследования. Представлены теоретическая и практическая значимость работы.

**В первой главе** приведен обзор актуальных на сегодняшний день проблем программных инструментов конфигурирования, выполненных в виде комплекса плагинов и дана оценка перспективности применяемого подхода к решению выявленных проблем.

Анализ показал, что стоимость решения может зависеть от доступности и стоимости применяемых технологий, характера реализации программного обеспечения (ПО), а так же от требований к квалификации оператора для взаимодействия с ПО. Так, например, дополняющий друг друга функционал может снижать стоимость решения. Повышать стоимость могут, например, конкурирующие за разделяемый ресурс единицы функционала.

Была выдвинута гипотеза, что снижение стоимости может быть достигнуто за счет оптимизации расходов на послепродажное обслуживание (ППО). В рамках оптимизации из поставки решения исключается бесполезный для заказчика функционал. Такой способ был бы актуален для решений, которые динамически формируют конечный функционал программного комплекса, вносят в него новый и изменяют существующий. Примером таких решений являются плагинные системы. В них интеграционными функциональными единицами являются плагины.

После изучения работ был сделан вывод, что в решаемой задаче необходимо учесть:

* число требований - в них описываются функциональные возможности ПО
* число файлов исходного кода - в них на языке программирования реализованы требования
* число плагинов - они включают в себя файлы исходного кода
* трассируемость требований на файлы исходного кода
* зависимости между файлами исходного кода
* распределение файлов исходного кода по плагинам

Благодаря этим сведениям можно построить математическую модель, отвечающую на вопрос, какие требования будут реализованы в поставке. Однако, если имеющуюся информацию дополнить сведениями о стоимости ППО требований и механизмами формирования результирующей стоимости, тогда может быть решена задача оптимизации внутренней структуры программного решения по критерию минимальной стоимости ППО.

**Во второй главе** приведено описание решаемой задачи, математически описана и формализована задача. Для ее описания выполнено:

* Описаны ограничения, выполняемые для элементов матрицы трассируемости требований к ПО на файлы исходного кода:
* Описаны ограничения, выполняемые для элементов матрицы распределения файлов по плагинам:
* Сформулирована целевая функция

Описан механизм исключения из задачи нелинейности. Для этого:

* Описана система неравенств, выполнение которых достигается при корректном умножении бинарных параметров:
* Описана система неравенств, выполнение которых достигается при корректной работе функции :
* Описана система неравенств, выполнение которых достигается при корректной работе функции :

Так же приведен пример математической модели, построенной для 1 комплектации, 5 требований, 4 файлов исходного кода и 3 плагинов. Полученная математическая модель включает 71 ограничение и 37 параметров. Из них 12 являются искомыми, а 25 дополнительными параметрами.

На рисунке 1 приведена графовая модель рассматриваемой в рамках примера задачи.

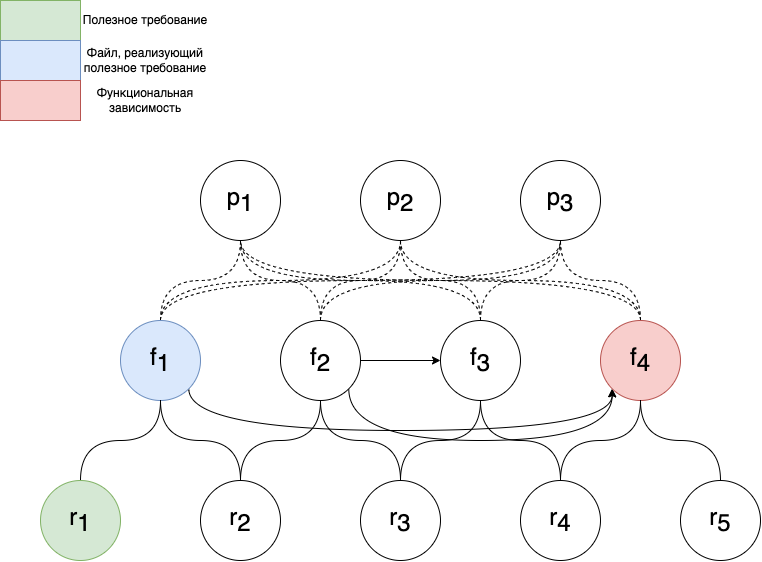


Рисунок 1. Графовая модель

**В третьей главе** описанная модель реализована на языке программирования Python. Реализация сочетает достоинства применения как парадигмы ООП, так и функционального программирования за счет активного использования лямбда-выражений.

Для реализации на языке программирования был использован внешний модуль Pyomo. В части решаемой задачи этот модуль был задействован для построения математической модели и взаимодействия с решателями. Построение математической модели заключалось в динамической генерации параметров модели и ограничений. Взаимодействие с решателями - вызов их из переменных окружения среды, задание входных аргументов для их работы и считывание результатов вычислений.

С целью понижения когнитивной сложности применяемого решения и достижения управляемости поведением каждой отдельной единицы функционала весь код разбит на обособленные модули, представленные в виде классов языка программирования Python.

Компоновка их друг с другом, последовательное разрешение зависимостей происходит в едином месте - в главной функции программы. В такой схеме отсутствуют передача локальных неуправляемых параметров по цепочке вызовов в иерархии функций, а каждая отдельная единица функционала работает в своем информационном окружении.

Диаграмма классов приведена на рисунке 2.

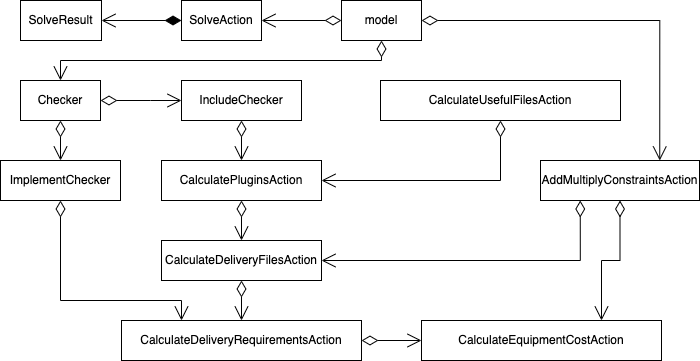


Рисунок 2. Диаграмма классов

**В четвёртой главе** дано описание проведенным экспериментам и описан анализ полученных результатов.

Так, рассмотренная в примере модель была обсчитана рядом решателей: glpk, gurobi, xpress, gams, gdpopt, mindtpy, cbc, conopt, copt, cplex, ilogcp и minos.

В рамках проведения экспериментов каждый решатель обсчитал модель по раз. При каждом обсчете вычислялось результирующее значение стоимости ППО от применения вычисленных значения элементов матрицы распределения (таблица 1), а так же замерялось время, необходимое решателю для выполнения вычислений (рисунок 3).

**Таблица** 1

**Стоимость ППО при применении результата работы решателей**

|  |  |
| --- | --- |
| Решатель | Стоимость ППО |
| glpk | 1 |
| gurobi | 1 |
| xpress | 1 |
| gams | 1 |
| gdpopt | 1 |
| mindtpy | 1 |
| cbc | 1 |
| conopt | 25 |
| copt | 1 |
| cplex | 1 |
| ilogcp | 1 |
| minos | 25 |

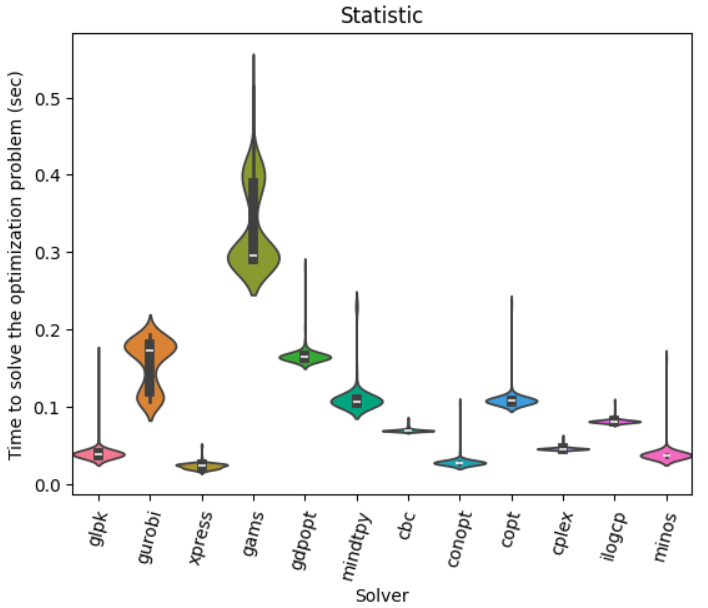


Рисунок 3*.* Данные о потребном времени работы решателей

Результаты экспериментов показывают, что время, потребное различным решателям для вычисления одних и тех же параметров может отличаться в разы. А так же, что нельзя идентифицировать полученный ответ от решателя как единственно верный. Лучше проводить серию экспериментов с разными решателями для получения достоверного результата.

Наименьшее значение стоимости ППО, полученное в ходе экспериментов, составляет 1. Это значение было достигнуто при применении результатов работы 10-ти решателей из 12. Наибольшее значение стоимости ППО составляет 25 и оно было достигнуто при применении в результатов работы 2 х решателей: conopt и minos. Наименьшее медианное рабочее время потребовалось решателям glpk, xpress и conopt. Наибольшее время потребовалось решателю gams.

**Заключение**

Разработана графовая модель, описывающая программное решение, выполненное в виде комплекса плагинов. Приведено описание метода, позволяющего не только сформулировать задачу оптимизации математической модели с целевой функцией, перечнем параметром и набором ограничений на них, но и свести задачу к линейному виду, что позволяет выполнять ее решение на ряде стандартных программных решателей.

Разработано программное решение, позволяющее выполнять задачи по формированию и решению описанной задачи оптимизации различными программными решателями.

**Работы, опубликованные автором по теме диссертации.**

В настоящее время опубликована одна статья в журнале, рецензируемом в ВАК:

Судаков В.А., Шаблий А.Д. Разрешение циклических зависимостей графовой модели взаимосвязи требований к программному обеспечению // Препринты ИПМ им. М.В.Келдыша. 2024. No 31. 13 с.

<https://doi.org/10.20948/prepr-2024-31>

<https://library.keldysh.ru/preprint.asp?id=2024-31>

В настоящее время ведется работа еще над двумя статьями и подготовка их публикации в ИПМ им. Келдыша и труды МАИ.