

На правах рукописи

ГУБАЙДУЛЛИН Ильдар Ирекович

**КОМПЛЕКСИРОВАНИЕ МОДУЛЕЙ
ПРИКЛАДНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ
НА ОСНОВЕ МЕТРИЧЕСКИХ ОЦЕНОК ЭФФЕКТИВНОСТИ**

**Специальность 05.13.11 –
Математическое и программное обеспечение вычислительных
машин, комплексов и компьютерных сетей**

**АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук**

Уфа - 2010

Работа выполнена на кафедре вычислительной техники и защиты информации
ГОУ ВПО «Уфимский государственный авиационный технический университет»

Научный руководитель

д-р техн. наук, проф.
ФРИД Аркадий Исаакович

Официальные оппоненты

д-р техн. наук, проф.,
МАРТЫНОВ Виталий Владимирович,
каф. экономической информатики
Уфимского государственного авиационного
технического университета

канд. техн. наук,
МАКСИМОВ Сергей Владимирович,
каф. программирования и
вычислительной математики Башкирского
государственного педагогического
университета им. М. Акмуллы

Ведущая организация

Башкирский государственный университет,
кафедра математического моделирования,
г. Уфа

Защита состоится «29» декабря 2010 года в 10:00 часов
на заседании диссертационного совета Д-212.288.07
при Уфимском государственном авиационном техническом университете
по адресу: 450000, г. Уфа, ул. К. Маркса, 12

С диссертацией можно ознакомиться в библиотеке университета

Автореферат разослан «24» ноября 2010 года

Ученый секретарь
диссертационного совета
д-р техн. наук, проф.



С. С. Валеев

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы

Согласно стандартам в области проектирования программного обеспечения (ПО) вычислительных машин, комплексов и компьютерных сетей, одной из основных характеристик качества ПО является его эффективность. Под эффективностью ПО понимают набор атрибутов, относящихся к соотношению между уровнем качества функционирования ПО и объемом используемых ресурсов при установленных условиях.

В настоящее время технологические процессы в промышленности управляются вычислительными машинами, характеризуемые как программной, так и аппаратной составляющей. Эффективность управления процессом, решением прикладной задачи, во многом определяется эффективностью ПО. Сложность оценки эффективности программ, необходимость учета специфики прикладной задачи, различные требования к уровню качества функционирования программ приводят к необходимости разработки новых методов оценки и обеспечения эффективности программного обеспечения.

В работах как российских, так и зарубежных авторов приведено множество различных методов и моделей качества и метрик ПО. При этом нередко отмечаются проблемы в их использовании. Во многих имеющихся методах обеспечения и оценки эффективности ПО не учитывается влияние исходных данных на результаты работы программы. Например, модели Липова, Джелинского–Моранды, Шика–Волвертона, Литтлвуда–Верралла, Миллса, Шумана и др. основываются на данных, полученных в процессе проектирования ПО. Сложность использования этих моделей заключается в том, что они построены на основе известных опытных данных о процессах разработки и не учитывают работу ПО в динамике. Универсальных метрик качества ПО не существует. В настоящее время их насчитывается более 200. При этом для некоторых специфических программ метрики ПО разработаны недостаточно. Например, имеется целый класс программ, в которых заметную роль играет не только эффективность технологии программирования, но и эффективность математического обеспечения. В таких программах необходим метод оценки, учитывающий специфику прикладной задачи и возможные качественные особенности, связанные, например, с плохой обусловленностью, с неустойчивостью полученных результатов при изменении входных данных. К примеру, в химической промышленности проблема повышения эффективности ПО приобретает особую актуальность в программном модуле определения кинетических параметров каталитических реакций. Обнаружено, что в существующей версии, как правило, в 30% запусков происходит останов дальнейших вычислений, причина которых является невысокая эффективность используемого ПО. При работе в реальных системах остановки могут приводить к огромным экономическим потерям, что связано с потерей дорогостоящих реагентов и катализаторов.

Для специфических задач особую актуальность приобретают динамические метрики, характеризующие такие параметры ПО, как вероятность достижения цели программы, время выполнения, объем занимаемой памяти. Как

правило, эти метрики приобретают разные значения в зависимости от законов распределения исходных данных. Поэтому изучение вопроса изменения метрик в зависимости от законов распределения исходных данных с последующей организацией управления этими метриками в ходе выполнения программного модуля, в частности, на основе идей комплексирования, является актуальной задачей. Хотя комплексирование при проектировании ПО используется, например, в работах Андреева Е. В., Аникина А. А., Зинченко В. П., однако комплексирование здесь рассмотрено с точки зрения прикладной области, и особенности внутренних свойств алгоритмов не учитываются. Механизм реализации идей комплексирования на основе оценивания метрик ПО не разработан.

Выявленные недостатки в существующих методах требуют разработки новых методов и алгоритмов, учитывающих специфику рассматриваемой программы и обеспечивающих требуемый уровень эффективности. В связи с этим, особую актуальность приобретает метод оценки эффективности ПО в зависимости от законов распределения исходных данных. На основе него в диссертационной работе предлагается метод управления эффективностью прикладного ПО путем комплексирования программных модулей.

Объектом исследования является эффективность прикладного ПО.

Предмет исследования – алгоритмы комплексирования программных модулей вычислительных машин на основе оценки эффективности программного обеспечения.

Целью работы является повышение эффективности прикладного ПО на основе комплексирования программных модулей.

Задачи исследования

1. Разработка метода оценки эффективности ПО в зависимости от законов распределения исходных данных.
2. Разработка метода и алгоритма управления эффективностью прикладного ПО на основе комплексирования программных модулей.
3. Разработка алгоритмического и программного обеспечения для оценки эффективности прикладного ПО.
4. Разработка алгоритма и методики проектирования программ с заданным уровнем эффективности.
5. Экспериментальное обоснование полученных результатов в системах реального времени, проведение вычислительных экспериментов на примере задачи определения кинетических параметров каталитических реакций.

Методы исследования

При решении поставленных в диссертационной работе задач использовались методы теории информации, системного анализа, вычислительной математики, моделирования сложных объектов, объектно-ориентированного программирования, теории алгоритмов, принципов и методов алгоритмизации.

Результаты, выносимые на защиту

1. Метод оценки эффективности программного обеспечения в зависимости от законов распределения исходных данных.

2. Метод и алгоритм управления эффективностью ПО на основе комплексирования программных модулей в зависимости от законов распределения исходных данных.
3. Алгоритм и методика проектирования программ с заданным уровнем эффективности.
4. Программный комплекс по оценке и управлению эффективностью ПО в зависимости от законов распределения исходных данных.
5. Результаты вычислительных экспериментов на примере задачи определения кинетических параметров каталитических реакций

Научная новизна результатов работы:

1. Метод оценки эффективности прикладного ПО, основанный на статистических данных, полученных в ходе моделирования работы ПО, отличающийся тем, что используя аппроксимацию законов распределения исходных данных, на основе предлагаемого метода получают статистические зависимости, характеризующие вероятность достижения цели программой.
2. Алгоритм и методика проектирования программ с заданным уровнем эффективности, отличающиеся тем, что на основе выбранных метрик определяются области, характеризующие эффективность ПО в зависимости от законов распределения исходных данных.
3. Метод и алгоритм управления эффективностью ПО на основе комплексирования альтернативных программных модулей, эффективность которых зависит от законов распределения исходных данных, отличающиеся использованием метода оценки эффективности ПО с учетом законов распределения исходных данных при реализации алгоритма комплексирования, и позволяющие повысить эффективность прикладного ПО в целом.

Практическая значимость и внедрение результатов работы

Практическая ценность результатов, полученных в диссертации:

1. Разработанный метод оценки эффективности программных продуктов и программного комплекса с учетом законов распределения исходных данных позволяет количественно оценить эффективность прикладного ПО и организовать процесс комплексирования.
2. Разработанные метод и алгоритм управления эффективностью ПО на основе комплексирования программных модулей в зависимости от законов распределения исходных данных позволяют увеличить вероятность достижения цели программы.
3. Разработанные алгоритм и методика проектирования прикладных программ с заданным уровнем эффективности позволяют определить диапазоны распределения параметров исходных данных, при которых прикладное ПО работает с заданным уровнем эффективности.

Разработанный программный продукт по расчету кинетических параметров каталитических реакций внедрен в практику работы лабораторий Института нефтехимии и катализа (ИНК) РАН и применяется для построения

кинетических моделей промышленно значимых реакций, используемых в программных модулях системы управления технологическими процессами.

Апробация работы

Основные положения и результаты работы докладывались и обсуждались на следующих научных конференциях и семинарах: XVII Международная конференция по химическим реакторам CHEMREACTOR (Афины–Крит, 2006); Всероссийская молодежная научная конференция «Мавлютовские чтения» (Уфа, 2008); III, IV, V Всероссийская зимняя школа-семинар аспирантов и молодых ученых (Уфа, 2008-2010); 9-я, 10-я и 12-я Международная конференция «Компьютерные науки и информационные технологии» (Уфа, 2007, Анталия, 2008, Москва–Санкт-Петербург, 2010); IX Молодежная научная конференция «Королёвские чтения» (Самара, 2007); студенческие научные конференции УГАТУ (Уфа, 2006-2010); научные семинары лаборатории математической химии ИНК РАН.

В 2009-2010 гг. диссертант являлся стипендиатом Президента Республики Башкортостан.

Связь с научными программами

Работа поддержана грантом РФФИ № 08-08-00357 «Методы проектирования отказоустойчивых программно-аппаратных вычислительных комплексов для систем управления сложными техническими объектами с использованием средств искусственного интеллекта» и грантом Президента Российской Федерации для государственной поддержки ведущих научных школ Российской Федерации № НШ-65497.2010.9

Публикации

Основные положения и результаты исследований по теме диссертации опубликованы в 13 работах, включая 3 статьи в научных изданиях из списка ВАК, 10 публикаций в центральных журналах, материалах Всероссийских и Международных конференций.

Структура и объем работы.

Диссертация состоит из введения, четырех глав основного текста, заключения, списка использованной литературы из 110 наименований, приложений, содержит 43 рисунка и 15 таблиц. Общий объем диссертации составляет 150 страниц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность темы, формулируются цель работы и задачи исследования, отмечается научная новизна и практическая ценность выносимых на защиту результатов.

В первой главе проводится анализ известных методов повышения эффективности прикладного программного обеспечения, выявляются их достоинства и недостатки. Описывается специфика задачи и обосновывается необходимость комплексирования программных модулей вычислительных машин в зависимости от законов распределения исходных данных.

Формулируется проблема выбора метода решения в зависимости от области распределения исходных данных программного модуля. Проблема актуальна, когда априори, до решения задачи, невозможно по наличию исходных данных определить наиболее эффективный метод решения. В качестве примера приведена задача минимизации нулевого порядка при неизвестной поверхности отклика.

Задача, решаемая с помощью ЭВМ, в общем случае может иметь несколько методов решения, каждый из которых может быть представлен определенным множеством алгоритмов. Алгоритмы могут быть реализованы множеством программ. Наличие многих алгоритмов решения одной задачи предоставляет возможность комплексирования программных модулей для обеспечения наибольшей эффективности решения прикладной задачи. Проводится анализ зависимости эффективности ПО от свойств алгоритмов. Среди основных свойств выделены следующие: детерминированность, конечность, наличие системы команд и области допустимых значений исходных данных.

Делается заключение о необходимости исследования задачи на устойчивость, сходимость и на отсутствие неопределенности. В связи с тем, что проверка подобных свойств является практически невыполнимой на этапе проектирования программы, приобретают актуальность методы анализа программ, основанные на их моделировании, в ходе которого проводится анализ программы на устойчивость к малым возмущениям и на отсутствие неопределенности.

Для обеспечения эффективности ПО необходима ее количественная оценка, поэтому в работе приводится классификация метрик ПО и анализ зависимости значений некоторых метрик от распределения исходных данных.

В случае комплексирования программных модулей формирование конфигураций итоговой программы происходит за счет использования нескольких методов решения задачи. При этом каждому методу решения соответствует определенный программный модуль, имеющий свои достоинства и недостатки, характеризуемые эффективностью, надежностью, сложностью, точностью. Конфигурирование общей системы основано на оценке критерия, приобретающего наилучшее значение выбранной метрики эффективности ПО за счет допустимого ухудшения остальных метрик.

На основе выполненного анализа, сформулированы цель и задачи исследования.

Вторая глава посвящена разработке метода управления эффективностью прикладного ПО на основе комплексирования программных модулей. Рассматриваются теоретические основы и приводятся примеры применения разработанного метода для некоторых задач.

Алгоритм управления на основе комплексирования состоит из двух этапов: динамическая оценка эффективности модулей и реконфигурирование программы. Оценка эффективности ПО производится путем моделирования работы исследуемой программы, когда ее исходные данные распределяются согласно определенному закону распределения. Особенностью предлагаемого подхода является проверка программы не на дискретном наборе исходных данных, а на некоторой распределенной области данных, что является статистически оп-

равданным. Методика оценки может использоваться для получения численной метрики эффективности. Основные блоки метода представлены на рис. 1. Программы представляются в виде функциональных элементов, на входы которых подаются величины, подчиняющиеся определенному закону распределения. Недостижимостью цели программы в широком смысле будем считать невыполнение блоком своих функций. Появление событий, приводящих к недостижимости целей, отслеживается блоками оценки и анализа. После достаточного количества прогонов методом оценки по частоте успешного выполнения задачи делается заключение об оценке вероятности достижимости цели, и при промышленной эксплуатации анализируемой программы модуль реконфигурирования выбирает оптимальный путь выполнения с максимальной эффективностью. Метод оценки эффективности состоит из следующих этапов: установление законов распределения исходных данных; определение метрик программы и организация их расчета; моделирование программы на основе заданных законов распределения; оценка эффективности программы на основе оценки вероятности по частоте достижения цели.

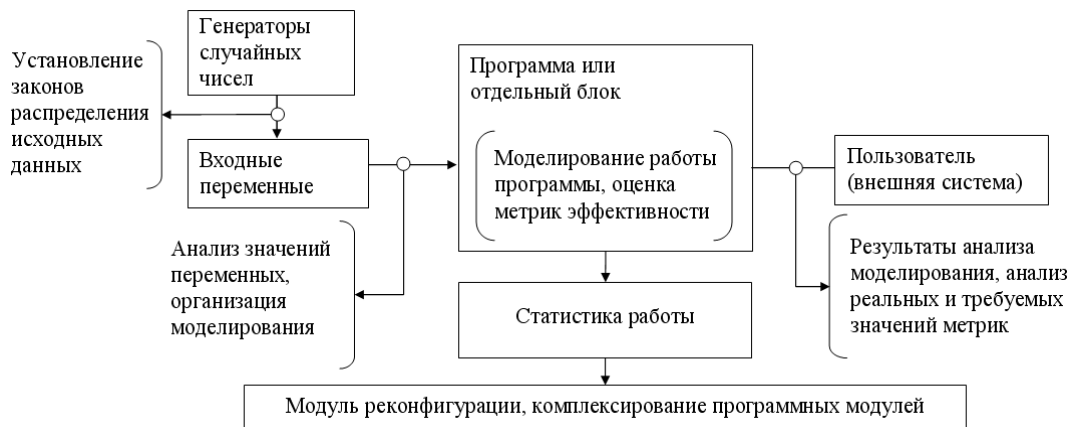


Рисунок 1 – Модель метода оценки эффективности

Описывается программный модуль в среде MATLAB для расчета вероятности достижения цели программ методом моделирования.

Разрабатывается метод управления эффективностью на основе комплексирования программных модулей. Для этого предлагается процедура комплексирования, которая формально может быть представлена с помощью теории множеств:

$$\Rightarrow K_i \begin{bmatrix} M_i \\ T_i \\ A_i \\ I_i \end{bmatrix} \rightarrow M_i \rightarrow R_1[M_i, \zeta] \rightarrow R_2[M_i, t] \rightarrow R_3[M_i, m] \rightarrow R_4[M_i, c] \rightarrow \{M\} \rightarrow Y\{M\} \rightarrow Q \Rightarrow \quad (1)$$

Здесь K_i – классификация программных модулей (M_i), задач (T_i), алгоритмов решения каждой задачи (A_i) и их функций; R_i – критерий выбора M_i в зависимости от необходимого параметра и от входных данных; Y – комплексирование программных модулей; Q – получение выходных результатов с наилучшим сочетанием выбранных метрик. Параметры программных модулей: ζ – величина отклонения от точного решения; t – время работы модуля; m – объем исполь-

зуемых ресурсов; c – число вызовов целевой функции. В отличие от известных, предложенная идея комплексирования опирается на использование метрик ПО.

Предлагается комплексный показатель эффективности модуля:

$$C(M_k) = C(\zeta, t, m, c) \rightarrow \max, \text{ при ограничениях } \begin{cases} \zeta < \zeta_0; \\ t < t_0; \\ c < c_0. \end{cases} \quad (2)$$

Приводится алгоритм реализации идеи комплексирования программных модулей. Алгоритм состоит из следующих основных этапов: выбор критерия по конфигурированию общей системы программных модулей; ранжирование списка конфигурируемых параметров по важности; сопряжение вычислительных модулей при комплексировании по входным и выходным переменным и функциям; обеспечение наилучшего сочетания выбранных параметров за счет допустимого ухудшения остальных показателей. Блок-схема алгоритма управления представлена на рис. 2. На первоначальном этапе управления необходимо накопление статистической информации об эффективности каждого блока. Эффективность модулей в зависимости от законов распределения исходных данных оценивается по ранее предложенному методу – работа программы или отдельных модулей моделируется в зависимости от законов распределения исходных данных и производится подсчет вероятности достижения цели модуля. Полученная информация используется при функционировании программы в реальном времени. Программа, достигнув определенного места в алгоритме и располагая параметрами закона распределения входных данных, выбирает наиболее эффективное решение. Управление ходом решения представлено с помощью линии управления. Получив выходные данные модуля 1, программа на основе текущих параметров закона распределения производит выбор наиболее эффективной альтернативы модуля A . Допустим, таким модулем оказался $A1$. Тогда другие пути, ведущие к модулям $A2$ и $A3$, временно отключаются до тех пор, пока выходные данные модуля 1 не попадут в промежуток, в котором эффективность отключенных модулей окажется выше эффективности модуля $A1$.

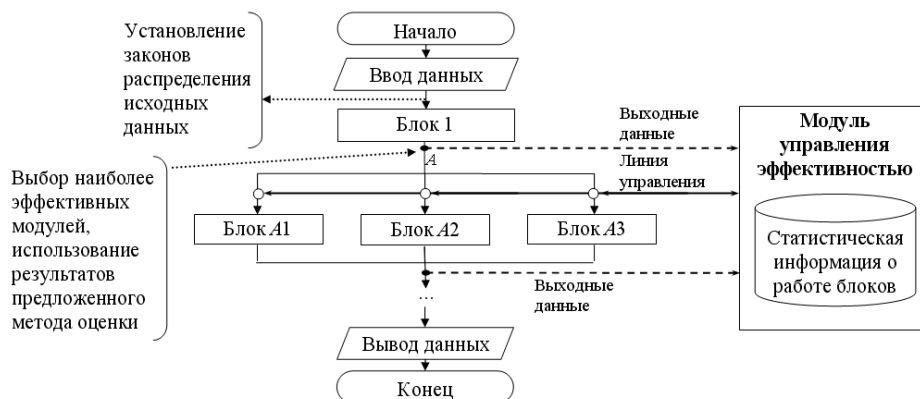


Рисунок 2 – Блок-схема алгоритма управления эффективностью

Принцип комплексирования заключается в выборе оптимального сочетания существующих модулей, при котором программа выполняется с максимальной эффективностью. Математически принцип комплексирования можно описать следующей функцией:

$$C = \sum_{j=1}^N \left(\max_{i=1}^k (C(M_i^j)) \right), \quad (3)$$

где C – общая комплексная эффективность программы; N – количество модулей; M_i^j – i -й альтернативный блок модуля M^j , $i \in 1 \dots k$, $j \in 1 \dots N$, k – количество альтернатив модуля M^j . В качестве примера предлагаемого алгоритма комплексирования рассматривается работа программы, реализующей решение системы уравнений Лотка–Вольтерра. Для численного решения задачи Коши в MATLAB реализованы семь различных вычислительных модулей (ВМ): ode45, ode23, ode113, ode15s, ode23s, ode23t, ode23tb. При выборе ВМ необходимо учитывать свойства решаемой системы. В противном случае можно получить неточный результат или затратить слишком много времени. Выбирая различные коэффициенты системы Лотка–Вольтерра, получают различные системы дифференциальных уравнений и соответствующие им решения. Для решения системы уравнений были выбраны три ВМ: ode45 (метод Рунге-Кутты 4), ode23s (метод Рунге-Кутты 2), ode113 (метод Адамса-Бэшфорда-Милтона), коэффициенты генерировались согласно нормальному закону распределения с дисперсией 0.25 и математическим ожиданием (МО) в диапазоне от 0 до 10. В задаче были приняты следующие ограничения: недопустимое время решения системы $t > T_0 = 10$ сек; недопустимое отклонение от точного решения $\zeta > \zeta_0 = 0.5$. Вероятность достижения цели программы определялась вероятностью нахождения решения с учетом ограничений.

В результате моделирования работы вычислительных методов ode23s (модуль A1), ode45 (модуль A2), ode113 (модуль A3) были получены графики вероятности достижения цели $P_{\text{дц}}$ в зависимости от математического ожидания (МО) S и R системы уравнений (рис. 3). На графиках светлые области соответствуют вероятностям достижения цели, близким к 0.9. Графики являются трехмерными, третья ось содержит уровень эффективности, характеризуемый вероятностью достижения цели. Линиями на графиках обозначены уровни-срезы с определенными значениями эффективности.

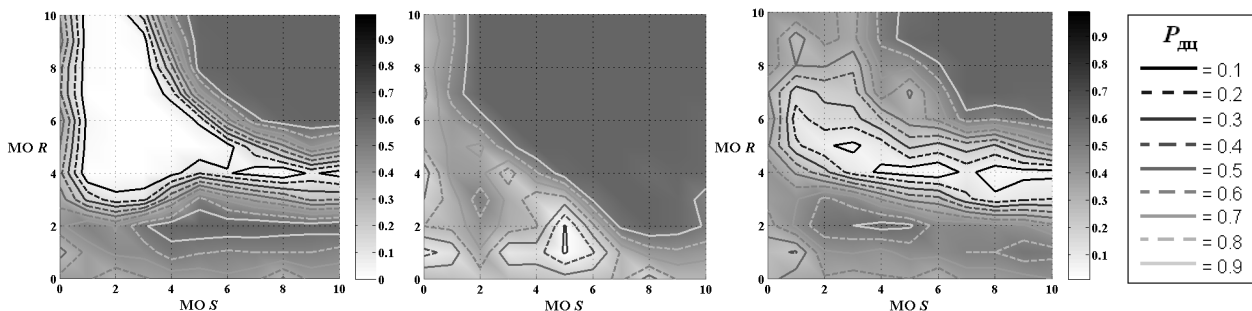


Рисунок 3 – Графики $P_{\text{дц}}(\text{МО } S, \text{МО } R)$ для ВМ ode23s, ode113 и ode45

При моделировании двумерного варианта ($S = \text{const}$) получены графики эффективности в зависимости от МО R (рис. 4). Согласно результатам моделирования, каждый ВМ имеет области, в которых вероятность достижения цели $P_{\text{дц}}$ менее 0.4 (на рис. 4 таким ВМ является ode113). Но при использовании комплексирования значение $P_{\text{дц}}$ общей системы – мажоранта, в соответствии с принятыми ограничениями, не опускается ниже уровня 0.6.

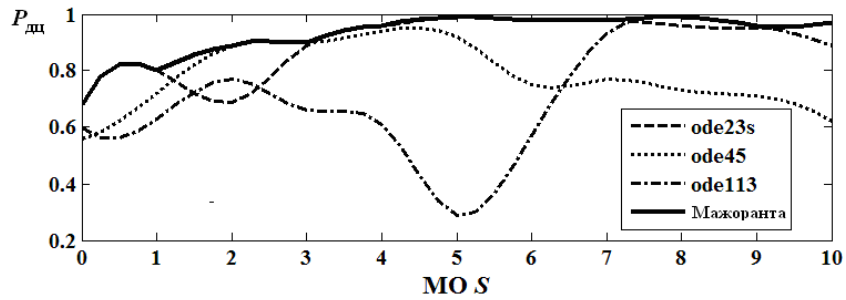


Рисунок 4 – Графики $P_{дц}$ в зависимости от $МО S$

Таким образом, предложен метод и разработан алгоритм управления эффективностью программы на основе комплексирования. Комплексирование заключается в выборе оптимального сочетания существующих вычислительных блоков, при котором программа выполняется с максимальной эффективностью.

В третье главе рассматривается проблема подготовки исходных данных с целью уменьшения затрат при использовании метода управления эффективностью на основе комплексирования программных модулей. Разрабатывается методика получения максимальной оценки эффективности в заданной области исходных данных и методика проектирования ПО с определенным уровнем эффективности. Методика состоит из 4 этапов: определение законов распределения исходных данных; формулирование целей программы и определение событий, препятствующих этому; моделирование программы с определенными законами распределения исходных данных, оценка эффективности ПО; определение допустимых параметров законов распределения, при которых программа будет обладать заданной эффективностью.

Первым шагом для реализации предложенной методики оценки эффективности ПО является установление законов распределения исходных данных. Это подразумевает определение вида закона распределения и его основных моментов. Для этого можно воспользоваться известными методами. Определение закона распределения в автоматическом режиме, т.е. в режиме выполнения программы, приводит к реконфигурированию системы для достижения максимальной эффективности. Для этого производится фоновое отслеживание поступающих значений входных переменных и при получении очередной порции данных производится расчет новых параметров закона распределения по предложенному алгоритму. Для применения описанного алгоритма в среде MATLAB разработаны соответствующие программные модули. Предложено альтернативное представление вложенных циклов, обладающее лучшим масштабированием по сравнению с традиционным представлением.

Ставится задача определения максимального уровня эффективности ПО в заданной области изменения исходных данных. Рассматриваются вопросы нахождения оптимального решения поставленной задачи.

Предложена методика подготовки исходных данных с целью уменьшения затрат при использовании метода управления эффективностью на основе комплексирования. Методика состоит из следующих основных этапов: определение законов распределения исходных данных; выявление ограничений в исследуемой программе; организация моделирования с накоплением статистической

информации о работе каждого программного модуля; комплексирование результатов моделирования, получение мажоранты, характеризующей эффективность программы в зависимости от законов распределения исходных данных.

В четвертой главе рассматривается применение метода управления эффективностью ПО на основе комплексирования программных модулей к программе поиска скоростных коэффициентов химической реакции. Существующая версия программы обладает низкой эффективностью, что приводит к частым остановам вычислений. Поэтому возникла необходимость в исследовании алгоритмов, заложенных в программу, и в проектировании новой версии на основе разработанного метода управления.

Задачу поиска скоростных коэффициентов химической реакции можно отнести к задачам подбора параметров. Для таких задач обычно имеется набор экспериментальных данных, для которых с помощью некоторой функции $f([x])$, где $[x]$ – вектор подбираемых коэффициентов, получают соответствующие наборы теоретических данных, при которых разность между теоретическими и экспериментальными данными будет удовлетворять некоторой наперед заданной величине. Рассматривая подобные задачи на системном уровне, можно выделить два основных блока:

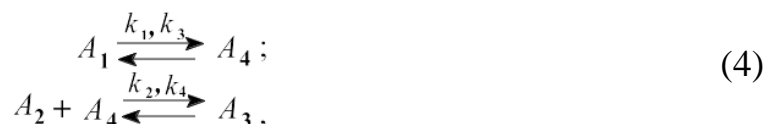
- 1) блок расчета теоретических данных, т.е. расчета функции $f([x])$;
- 2) блок подбора параметров, осуществляющий поиск минимума некоторой дискретной N -мерной поверхности, точками которой являются возможные наборы значений первого блока.

Подобные задачи решаются в химической кинетике при определении скоростных коэффициентов многостадийной химической реакции. При этом первый блок занимается решением системы обыкновенных нелинейных дифференциальных уравнений (СОНДУ), а второй блок реализует некоторый оптимизационный метод. Проводится исследование определенной реакции, в процессе которой через заданные промежутки времени фиксируют выход некоторых веществ, причем выход остальных веществ не наблюдается.

Задача отыскания констант может быть сформулирована следующим образом: на основе данных о скорости изменения концентраций найти такие значения k , которые наилучшим образом отвечали бы эксперименту. Математически эта задача сводится к минимизации некоторой функции отклонения экспериментальных данных от расчетных значений. В рассматриваемой задаче нахождения скоростных констант химических реакций выделяются две задачи:

- 1) задача решения СОНДУ;
- 2) задача минимизации разностей между теоретическими и экспериментальными данными, т.е. некоторой функции отклонения экспериментальных данных от расчетных значений.

В качестве примера рассматривается модель реакции, используемая в программном модуле системы управления технологическим процессом:



где A_j – определенные вещества (конкретный химический состав каждого вещества в данной работе не приводится). Выход веществ A_1, A_3 – наблюдаем, остальные – нет; k_i (с^{-1}) – скоростной коэффициент i -ой стадии.

В общей сложности, модели химических реакций могут быть представлены большим количеством уравнений СОНДУ, что ведет к увеличению переменных, исходных данных и соответствующих ограничений.

Математическое описание модели (4), согласно теории химической кинетики, представляется определенной СОНДУ, решение которой зависит от параметров $k_1..k_4$. Меняя эти константы и решая систему уравнений, определяют изменения во времени концентрации веществ $x_1..x_4$. После этого, найденные концентрации сравнивают с экспериментальными данными. Конечная цель вычислений – определить такие $k_1..k_4$, при которых разность между экспериментальными и расчетными данными будет пределах допуска.

Рассматриваемая программа состоит из двух основных блоков: 1) блоки решения системы дифференциальных уравнений; 2) блок, реализующий оптимизацию некоторой целевой функции, аргументы которой зависят от решения первого блока. Целевая функция рассматриваемой задачи имеет вид:

$$\zeta = \min \left(\sum_i^m (|y_i - F(x_i, [k])|) \right), \quad (5)$$

где y_i – ряд экспериментальных точек, определенных для соответствующих интервалов времени x_i ; m – количество экспериментальных точек; $[k]$ – вектор скоростных коэффициентов химических реакций; $F(x_i, [k])$ – функция теоретических значений при определенном наборе k . Исходными данными являются: схема превращений, т.е. математическая модель реакции, первоначальные мольные доли веществ и экспериментальные данные, полученные в процессе проведения реакции через фиксированные промежутки времени.

Подбирается такой набор скоростных коэффициентов $[k]$, при которых абсолютная разность между теоретическими и экспериментальными данными не будет превышать некоторую наперед заданную величину ζ_0 . В начале работы оптимизационный алгоритм (ОА) выбирает первоначальный вектор скоростных коэффициентов. Подставив этот вектор в СОНДУ, находится решение, из которого извлекается информация о количественном соотношении выхода веществ. Найденные процентные доли веществ, которые являются теоретическим решением, сравниваются с данными, полученными в процессе эксперимента.

Блок-схема программы расчета скоростных коэффициентов представлена на рис. 5. Исходными данными для программного модуля являются: схема превращений, т.е. математическая модель реакции, первоначальные мольные доли веществ и экспериментальные данные, полученные в процессе проведения реакции через фиксированные промежутки времени.

После ввода исходных данных начинается основная работа программы – поиск скоростных коэффициентов. В начале работы оптимизационный алгоритм (ОА) выбирает первоначальный вектор скоростных коэффициентов. Подставив этот вектор в СОНДУ, находится решение, из которого извлекается информация о количественном соотношении выхода веществ. Найденные про-

центные доли веществ, которые являются теоретическим решением, сравниваются с данными, полученными в процессе эксперимента. Цель ОА заключается в подборе таких скоростных коэффициентов, при которых разность между теоретическими и экспериментальными данными будет удовлетворять некоторой наперед заданной величине.

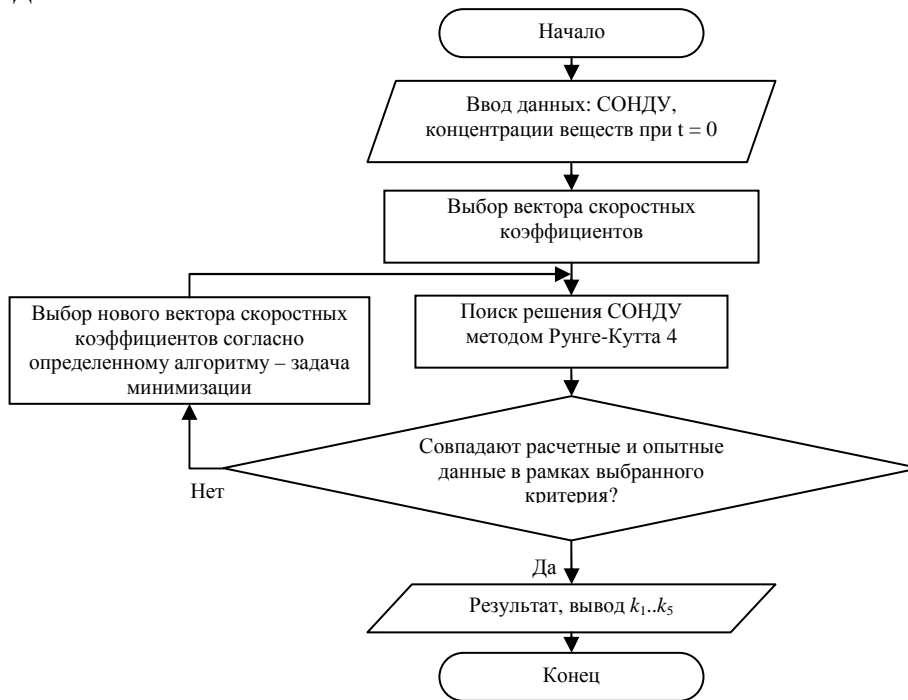


Рисунок 5 – Блок-схема программы расчета скоростных коэффициентов

Блок решения системы обыкновенных нелинейных дифференциальных уравнений в текущей версии реализован с помощью алгоритмов Эйлера, Кутты-Мерсона и Рунге-Кутты. В качестве оптимизационного алгоритма используется параболический спуск. В случае заикливания оптимизационного алгоритма, вызванного попаданием в локальный минимум, производится перезапуск программы с выбором другой начальной точки. В текущей версии, несмотря на наличие различных алгоритмов решения первого блока, переключение между ними происходит в диалоговом режиме, и в случае отказа одного из блоков программа останавливается в ожидании действий оператора.

Функционирование рассматриваемого алгоритма зависит от выбора первоначального вектора k , т.е. от выбора начальной точки: $\{k_1, k_2, k_3, k_4\}$. Поэтому вероятность достижения цели программы рассчитывается в зависимости от выбора начальной точки. Достижением цели считалось событие, при котором метод оптимизации завершил работу в пределах допуска $\xi < \xi_0$, и заданного времени $T < T_0$. Примем, что все элементы вектора k распределены по нормальному закону. В данном примере среднее квадратическое отклонение для всех элементов принято равным 1, а математическое ожидание (МО) элементов было расположено в интервале от 1 до 11 с шагом 2.

Имеется множество способов решения СОНДУ и различные способы многомерной минимизации. Для увеличения эффективности программы необходимо использовать различные методы, реализующие описанные блоки. Применяя определенные блоки, получим различные показатели T и ξ .

Рассмотрено нарушение формальных свойств алгоритма, реализующего решение обратной задачи химической кинетики. Первое формальное свойство алгоритма для этой задачи выполняется строго – алгоритм является детерминированным, последующий шаг алгоритма однозначно определяется текущим состоянием. Однако может быть нарушено свойство конечности, приводящее к недостижению цели программы. Во-первых, для алгоритмов оптимизации характерно «попадание» в локальные минимумы, в которых значение целевой функции (5) окажется недопустимым большим. Во-вторых, для алгоритмов решения системы обыкновенных дифференциальных уравнений используются численные методы, время решения которых зависит от точности используемых методов, что, в конечном счете, сказывается на времени вычисления. Отсутствие результата за заданный интервал времени также можно считать недостижением цели. Более того, рассматриваемую задачу можно считать некорректной, так как она не всегда устойчива к изменению небольших возмущений исходных данных.

Функционально вероятность достижения цели программы зависит от первоначальной точки, т.е. $P_{\text{дц}}(k_1, k_2, k_3, k_4)$. Метод оценки эффективности основывается на моделировании работы программы, когда на вход поступают данные, распределенные по определенному закону распределения.

В результате моделирования для конкретного вектора k , состоящего из математических ожиданий, было определено значение вероятности достижения цели. Для наглядности зафиксируем математические ожидания k_3 и k_4 равными 11 и 1, соответственно. Изменение вероятности достижения цели программой $P_{\text{дц}}$ в зависимости от МО k_1 и k_2 можно представить графически (рис. 6).

Области, в которых достижение цели наиболее вероятно, обозначены темным цветом. Для обеспечения высокой эффективности программы необходимо выбирать МО k_1 и k_2 из интервалов [1; 6] и [2; 5], соответственно.

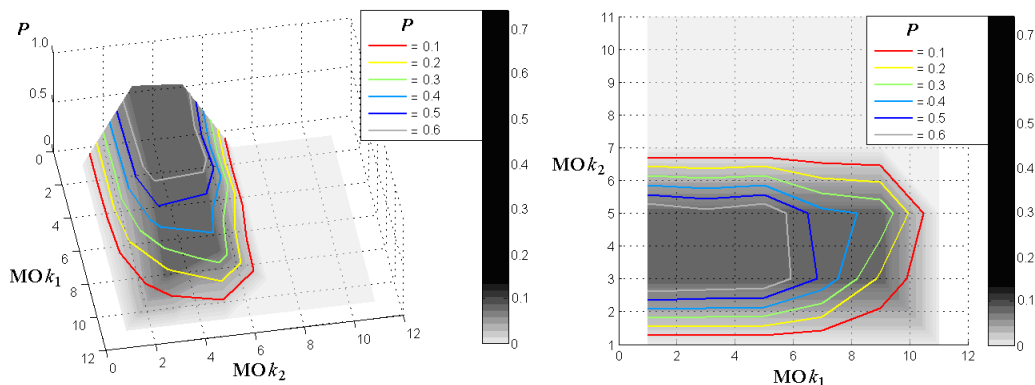


Рисунок 6 – Вид функции $P_{\text{дц}}(\text{МО } k_1, \text{МО } k_2)$ при $\text{МО } k_3 = 11$ и $\text{МО } k_4 = 1$

В результате моделирования работы вычислительных методов ode23s (модуль A1), ode45 (модуль A2), ode113 (модуль A3), аналогично примеру с моделированием задачи Лотки–Вольтера, получены графики вероятности безотказной работы P в зависимости от МО k_1 и k_2 , при этом k_3 и k_4 фиксированы.

Для лучшего восприятия результатов моделирования рассматривается двумерный вариант. Предположим, что k_2 постоянно. На рис. 7 представлены графики эффективности программы в зависимости от МО k_1 .

Согласно результатам моделирования, каждый ВМ имеет области, в которых вероятность достижения цели P менее 0.3 (на рис. 7 таким ВМ является ode23s). Но при этом значение $P_{\text{дц}}$ общей системы, мажоранта, не опускается ниже уровня 0.7. В рассматриваемой системе работа одних ВМ оказалась более эффективной в определенных областях, в результате существенно повысилась эффективность общей системы.

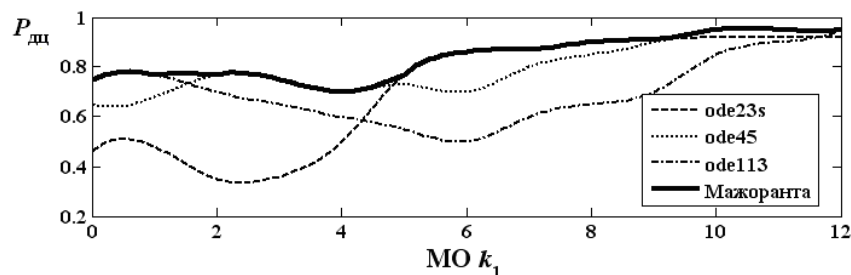


Рисунок 7 – Графики $P_{\text{дц}}$ в зависимости от МО k_1

Исходя из полученных результатов, делается вывод о том, что разработанный метод, основанный на моделировании работы программы в зависимости от законов распределения исходных данных, может рассматриваться не только в качестве инструмента повышения эффективности, но и в качестве инструмента анализа задачи на устойчивость к малым возмущениям. Метод позволяет определить области, в которых рассматриваемая задача наиболее устойчива. По виду полученной зависимости на рис. 7 можно сделать вывод о том, корректна ли задача в плане устойчивости. В случае плавного графика этой зависимости (сплошная линия) – задача является более «грубой» (малочувствительной). А для зависимости, представленной методом ode23s, имеется область, для которой небольшое изменение исходных данных приводит к резкому изменению вероятности достижения цели программы. Таким образом, с помощью предлагаемого метода можно также оценить чувствительность вероятности достижения цели программы в зависимости от параметров закона распределения исходных данных. В зависимости от МО входной переменной можно получить совокупную эффективность программы на основе комплексирования вычислительных методов, которая будет выше эффективности блока A1, A2 или A3, взятых в отдельности. Выбирая более эффективный ВМ в зависимости от исходных данных, программа работает более эффективно при более широком диапазоне их изменения.

Метод управления эффективностью ПО на основе комплексирования программных модулей реализован в виде отдельного блока для информационно-аналитической системы решения обратных задач химической кинетики Института нефтехимии и катализа РАН, что позволило:

- 1) уменьшить общее количество остановов при решении задачи;
- 2) определить области распределения исходных данных с заданной вероятностью достижения цели, при которых программа выполняется наиболее эффективно;
- 3) увеличить точность определения скоростных коэффициентов при разработке кинетических моделей частных реакций каталитического гидроалюминирования олефинов.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

1. Предложен метод оценки эффективности прикладного ПО, основанный на статистических данных, полученных в ходе моделирования работы ПО, отличающийся тем, что используются знания законов распределения исходных данных для получения статистических зависимостей, характеризующих вероятность достижения цели программой.

2. Разработаны алгоритм и методика проектирования программ с заданным уровнем эффективности, основанные на предложенном методе оценки и позволяющие на основе выбранных метрик определить параметры законов распределения исходных данных, при которых программа будет обладать заданным уровнем эффективности.

3. Разработаны метод и алгоритм управления эффективностью ПО на основе комплексирования альтернативных программных модулей в зависимости от законов распределения исходных данных, отличающиеся использованием метода оценки эффективности ПО в зависимости от законов распределения исходных данных при реализации алгоритма комплексирования, и позволяющие повысить эффективность прикладного ПО, состоящего из различных программных модулей.

4. Разработано ПО для оценки эффективности исследуемых программ и реализующее разработанный метод управления. Для задачи моделирования кинетических параметров химической кинетики получена интервальная оценка погрешности определения вероятности достижения цели.

5. Разработанный программный модуль по расчету кинетических параметров каталитических реакций внедрен в практику работу лабораторий ИНК РАН и применяется для построения кинетических моделей промышленно значимых реакций. Метод управления эффективностью ПО на основе комплексирования позволил повысить эффективность ПО, увеличив вероятность достижения цели с 0.3 до уровня 0.7.

ПУБЛИКАЦИИ ПО МАТЕРИАЛАМ ДИССЕРТАЦИИ

В рецензируемых журналах из списка ВАК

1. Реализация N -вариантности для обратных задач кинетики механизмов цикло- и гидроалюминирования олефинов / И.И. Губайдуллин, И.М. Губайдуллин // Обозрение прикладной и промышленной математики. М.: ТВП, 2006. Т.13, №4. С. 629–631.
2. Метод экспериментальной оценки надежности программного обеспечения / И.И. Губайдуллин, А.И. Фрид // Информационные технологии. М.: Новые технологии, 2007. №5. С.165–173.
3. Оценка надежности одного типа программ с определением погрешности полученной оценки / И.И. Губайдуллин, А.И. Фрид // Вестник УГАТУ, 2010. Т.14, №4 (39). С. 145–151.

В других изданиях

4. Система управления базой данных сложных химических реакций в присутствии циркониевого катализатора $(C_2H_5)_2ZrCl_2$ / И.И. Губайдуллин, И.М. Губайдуллин // Химреактор-17: сб. науч. тр. 17-й междунар. конф. по химическим реакторам, 15–16 мая 2006, Афины-Крит. М.: ЗАО «Калвис», 2006. С. 636–639.
5. Методика оценки надежности программного обеспечения в зависимости от распределения исходных данных / И.И. Губайдуллин, А.И. Фрид // Естественные и технические науки. М.: «Спутник+», 2007. № 3. С. 194–197.
6. Об одном подходе к оценке надежности программного обеспечения / И.И. Губайдуллин, А.И. Фрид // Труды междунар. конф. комп. наук и информ. техн. (CSIT'2007). Красноустьинск–Уфа: УГАТУ, 2007. Т.2. С. 153–157 (Статья на англ. яз).
7. Об одном методе оценки надежности программного обеспечения / И.И. Губайдуллин, А.И. Фрид // Королевские чтения. Самара: Издательство Самарского государственного аэрокосмического университет, 2007. С. 284.
8. Оценка алгоритмической надежности программного обеспечения в одной задаче химической промышленности / И.И. Губайдуллин // Актуальные проблемы в науке и технике. Т. 1: сб. ст. рег. зимней шк.-сем. аспирантов и молодых ученых, 20–23 февраля 2008. Уфа: Технология, 2008. С. 287–293.
9. Оценка надежности программы решения одной задачи химической кинетики / И.И. Губайдуллин, А.И. Фрид // Труды междунар. конф. комп. наук и информ. техн. (CSIT'2008). Анталия: УГАТУ, 2008. Т.1. С. 184–187 (Статья на англ. яз.).
10. Оценка точности расчета вероятности отказа одной программы / И.И. Губайдуллин, А.И. Фрид // Мавлютовские чтения : Всерос. молодежн. конф : сб. тр. Т. 3. Уфа: УГАТУ, 2008. С.122–123.
11. Управление надежностью программного обеспечения при известных законах распределения исходных данных / И.И. Губайдуллин // Актуальные проблемы в науке и технике. Т. 1: сб. ст. рег. зимней шк.-сем. аспирантов и молодых ученых, 19–21 февраля 2009. Уфа : Технология, 2009. С. 167–173.
12. Надежность системы вычислительных методов / И.И. Губайдуллин, А.И. Фрид // Труды междунар. конф. комп. наук и информ. техн. (CSIT'2010). Санкт-Петербург: УГАТУ, 2010. Т.1. С. 158–162 (Статья на англ. яз.).
13. Управление надежностью программного обеспечения при известных законах распределения исходных данных / И.И. Губайдуллин // Актуальные проблемы в науке и технике. Т. 1: сб. ст. рег. зимней шк.-сем. аспирантов и молодых ученых. Уфа: Технология, 2010. С. 136–141.

ГУБАЙДУЛЛИН Ильдaр Ирекович

КОМПЛЕКСИРОВАНИЕ МОДУЛЕЙ
ПРИКЛАДНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ
НА ОСНОВЕ МЕТРИЧЕСКИХ ОЦЕНОК ЭФФЕКТИВНОСТИ

Специальность 05.13.11 – Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Подписано к печати 19.11.10. Формат 60x84 1/16.
Бумага офсетная. Печать плоская. Гарнитура Times New Roman.
Усл. печ. л. 1,0. Усл. кр.-отг. 1,0. Уч.- изд. л. 1,0.
Тираж 100 экз. Заказ № 473.

ГОУ ВПО Уфимский государственный
авиационный технический университет
Центр оперативной полиграфии
450000, Уфа-центр, ул. К.Маркса, 12