О Т З Ы В

официального оппонента, старшего научного сотрудника
Федерального государственного учреждения
«Федеральный исследовательский центр
Институт прикладной математики им. М.В. Келдыша
Российской академии наук»,
к.ф.-м.н. Климова Юрия Андреевича
на диссертацию
Андрианова Павла Сергеевича
«Анализ корректности синхронизации
компонентов ядра операционных систем»,
представленную на соискание ученой степени кандидата физико-математических наук по специальности 05.13.11 – математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

Актуальность темы

Обеспечение надежного функционирования многопоточного программного обеспечения (ПО), очевидно, является важной и актуальной задачей. Количество ядер в современных процессорах исчисляется несколькими десятками и рост количества ядер продолжается, поэтому для достижения высокой производительности практически любое ПО обязано становиться многопоточным. В ядрах современных операционных систем одновременно может выполняться множество запросов от различных устройств и пользовательских программ. А от корректности ядра операционной системы зависит надежность функционирования всего ПО. Поэтому обнаружение ошибок в многопоточном системном ПО (в том числе и в ядрах операционных систем) является актуальной задачей.

В многопоточном ПО, в дополнении к ошибкам, которые свойственны всем видам программ, могут встречаться ошибки, связанные с параллельной

работой нескольких потоков, например состояния «гонки» и состояния взаимной блокировки. Ошибки корректной синхронизации являются достаточно «неудобными» для программиста: они могут проявляться только в определенных или редких ситуациях, поэтому традиционные методы отладки недостаточно эффективны. В тоже время для критически важного ПО (каким являются, например, ядра операционных систем) нам хочется не столько исправить все обнаруженные ошибки, сколько доказать отсутствие ошибок (в необходимых предположениях).

Для решения подобных задач используются различные методы анализа и верификации программ. Однако существующие инструменты не способны обеспечить надежное решение указанной задачи: на объемном коде простые и быстрые методы не обеспечивают приемлемое качество проверки кода, а более точные методы требуют слишком много ресурсов.

В результате автором диссертации были поставлены следующие актуальные задачи:

- Разработать метод поиска состояний гонки на основе подхода с раздельным анализом потоков.
- Разработать алгоритм построения модели окружения потока, который будет обладать конфигурируемостью и масштабироваться на большие объемы исходного кода, и доказать его корректность.
- Модернизировать алгоритм адаптивного статического анализа с целью обеспечения возможности выполнять верификацию программного обеспечения с раздельным анализом потоков и доказать его корректность.
- Реализовать разработанные алгоритмы.
- Провести эксперименты и сравнить результаты с другими инструментами статической верификации.

Степень обоснованности научных положений, выводов и рекомендаций

Диссертационная работа Андрианова П.С. состоит из введения, четырех глав, заключения, списка литературы и двух приложений.

Во **введении** обосновывается актуальность темы, определяются цели и задачи работы.

В **первой главе** рассматриваются существующие методы анализа и верификации многопоточных программ. Автор выделяет и подробно рассматривает три основные группы методов: методы динамического анализа, методы статического анализа и методы статической верификации многопоточных программ. Основным выводом обзора является то, что хотя существует множество методов и инструментов, но ни один из них не является достаточным для качественного и надежного анализа больших объемов критически важного программного обеспечения.

Во **второй главе** формально описывается предложенный метод верификации многопоточных программ с раздельным рассмотрением потоков. Основой описания является предложенное автором расширение классического адаптивного статического анализа, основанное на рассмотрении частичных абстрактных состояний, содержащих лишь часть информации о конкретных состояниях программы. Формулируется и доказывается теорема о корректности предложенного подхода. Далее на основе разработанного формализма приводятся описания различных вариантов анализа и доказательства того, что они удовлетворяют условиям теоремы о корректности. Это один из самых интересных разделов диссертации. Разнообразие рассмотренных и разработанных методов производит впечатление глубоко проработанной теории: алгоритм раздельного анализа потоков и построения модели окружения, формализующих влияние потоков друга на друга; оператор проекции, выдающий вид одного потока для другого, и его применение для анализа; композиция нескольких видов анализа; варианты анализа множества активных потоков с учетом и без зависимостей по созданию потоков; формальная модель синтаксической достижимости точек программы; анализы отслеживания множества используемых примитивов синхронизации и присваивания явных значений; и другие компоненты метода. В завершение

формулируется основной метод поиска состояния гонки многопоточных программ.

В **третьей главе** описана программная реализация предложенного метода в существующем открытом проекте CPAchecker. Описывается как общая инфраструктура CPAchecker, так и особенности реализаций и оптимизаций различных вариантов анализа, описанных во второй главе. Обсуждаются особенности принятых решений, множество отдельных оптимизаций, повышающих эффективность и масштабируемость инструмента. Здесь представлен поучительный живой опыт разработки и реализации систем такого рода.

В **четвертой главе** приведены результаты экспериментального исследования разработанного инструмента. Приведено как сравнение с ведущими инструментами в области поиска состояний гонки и верификации многопоточных программ, так и сравнение результатов работы разработанного инструмента при использовании различных конфигураций. Сравнение проводилось как на наборе небольших тестовых задач, так и на задачах на основе драйверов ядра Linux и на основе ядер закрытых операционных систем реального времени. Показано, что разработанный инструмент на тестовом наборе демонстрирует результаты, сравнимые с результатами других инструментов. Но эти инструменты не справляются со сложными реальными задачами, основанными на драйверах ядра Linux, в отличие от разработанного автором инструмента. Особый интерес представляет подробное обсуждение результативности отдельных компонентов метода и их вариаций. Приведены результаты экспериментов и выводы как для вариантов с хорошими результатами, так и те, которые не дали выгоды.

В **Заключении** перечислены основные результаты диссертационной работы.

В **Приложении А** приведены доказательства теорем из второй главы.

В **Приложении Б** приведены описания ошибок в ядре Linux, на которых исследовалась работа разработанного инструмента.

Таким образом, диссертация содержит обоснование актуальности темы, формулировку задачи и предлагает решение поставленной задачи. Доказана теорема о корректности предложенного метода. Показано, что предлагаемые методы применимы на практике в современных инструментах верификации программ.

Оценка научной новизны и достоверности

В качестве основных научных результатов, полученных автором работы, можно указать следующие:

- Новый метод поиска состояний гонки на основе раздельного анализа потоков, использующий средства абстракции состояний и переходов для управления точностью и ресурсоемкостью верификации.
- Новый алгоритм построения окружения потока, который позволяет гибко настраивать уровень абстракции над взаимодействием потоков, и доказательство его корректности.
- Новый алгоритм, который является обобщением существующего алгоритма статической верификации программ при помощи метода адаптивного статического анализа (Configurable Program Analysis, CPA), расширяющий типовой набор CPA-анализаторов средствами верификации многопоточных программ с раздельным анализом потоков, и доказательство его корректности.

Достоверность результатов работы подтверждается теоретическим доказательством корректности предложенного метода, а также успешным применением разработанного метода в инструменте статического анализа CPAchecker и экспериментальной оценкой разработанного инструмента.

Основные результаты диссертации опубликованы в 8 печатных работах и доложены в ряде научных конференций и семинаров (в том числе международных).

Замечания по диссертационной работе

По тексту работы имеются следующие замечания.

- Диссертационная работа направлена на развитие метода статистического анализа для многопоточных программ, в первую очередь ядра операционной системы. Но в тексте нет явного описания, чем код ядра операционной системы отличается от кода других многопоточных программ, какие особенности кода ядра позволили решить поставленную задачу, а какие особенности кода других программ мешают успешному применению предложенного метода. Другими словами, недостаточно четко описана область применимости разработанного метода.
- В рассматриваемом во второй главе модельном языке присутствует операция создания потока thread_create, но отсутствует операция ожидания завершения потока. В работе не сказано является ли это принципиальным ограничением метода. Представляется интересным расширить модельный язык, добавив в него все широко используемые конструкции реальных языков.
- В работе приводятся как количественные, так и качественные результаты экспериментальной оценки предложенного метода. Но в тексте нет примеров исходного кода анализируемых программ. Было бы интересно явно проанализировать подобные программы, однако их размер может быть достаточно велик.
- В диссертационной работе имеется небольшое число опечаток и ошибок пунктуации.

Эти замечания, впрочем, не ставят под сомнение ценность проделанной диссертантом работы, а скорее являются пожеланиями по дальнейшему обобщению и углублению полученных им результатов. Отмеченные недостатки не затрагивают сущность исследований, не снижают их качество и не влияют на главные теоретические и практические результаты диссертации.

Заключение

Диссертационная работа Андрианова П.С. является законченной научно-квалификационной работой, выполненной автором самостоятельно на высоком научном уровне. Полученные автором результаты достоверны, выводы и заключения обоснованы. Результаты представляют высокий научный интерес и подтверждаются экспериментальными данными.

Автореферат соответствует основному содержанию диссертации.

Представленная диссертационная работа отвечает требованиям «Положения о порядке присуждения ученых степеней», предъявляемым к кандидатским диссертациям по специальности 05.13.11 (математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей), а ее автор, Андрианов Павел Сергеевич, заслуживает присуждения ему ученой степени кандидата физико-математических наук по специальности 05.13.11.

Официальный оппонент,
старший научный сотрудник ИПМ им. М.В. Келдыша РАН, к.ф.-м.н.

Ю.А. Климов

Подпись старшего научного сотрудника Климова Юрия Андреевича заверяю.
Ученый секретарь ИПМ им. М.В. Келдыша РАН, к.ф.-м.н.

А.А. Давыдов