ОТЗЫВ

ОФИЦИАЛЬНОГО ОППОНЕНТА

на диссертационную работу Каушана Вадима Владимировича «Поиск ошибок выхода за границы буфера в бинарном коде программ», представленную к защите на соискание учёной степени кандидата технических наук по специальности 05.13.11 — «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей»

Диссертационная работа В.В. Каушана посвящена разработке метода поиска ошибок выхода за границы буфера в программах. Проблема поиска ошибок выхода за границы буфера **актуальна**, поскольку данный класс ошибок является одним из наиболее распространённых и опасных. **Новизна метода**, предложенного диссертантом, связана с использованием анализа длин буферов, при котором происходит абстрагирование от реальных размеров буферов, что позволяет подбирать такие параметры выполнения программы, которые приводят к проявлению ошибки. Этот метод позволяет находить ошибки, даже если они не проявляются во время анализа программы. Кроме того, в отличие от многих существующих методов, данный метод не требует наличия исходных кодов анализируемой программы, а также отладочной информации, что существенно расширяет область его применения.

Диссертационная работа содержит 92 страницы и состоит из введения, четырёх глав, заключения, списка литературы (55 наименований) и одного приложения.

Во введении обосновывается актуальность работы, формулируется цель, основные научные результаты, обосновывается практическая ценность, а также приводятся сведения об апробации работы.

В **первой главе** приводится обзор существующих подходов к решаемой задаче. Рассматриваются подходы на основе статического анализа, а также на основе инструментирования программы и динамического символьного выполнения. Отмечаются недостатки существующих подходов. Относительно высокий уровень ложных срабатываний у методов статического анализа требует ощутимых трудозатрат на этапе проверки ошибок для больших проектов; к проверке предупреждений необходимо привлекать специалистов высокой квалификации. К недостаткам методов динамического анализа относят их неспособность обнаруживать ошибки, не проявившиеся при выполнении на наборе тестов. Формулируются требования к перспективному методу поиска ошибок: наличие подтверждающего примера, абстрагирование от конкретной

операционной системы (ОС) и процессорной архитектуры, возможность анализировать не только отдельные приложения, но и код ядра ОС, а также взаимодействие нескольких процессов, обнаружение ошибок, которые не реализовались во время запуска программы.

Вторая глава посвящена предлагаемому методу поиска ошибок выхода за границы буфера. Метод использует динамический анализ трассы машинных команд, но совмещает его с абстрагированием от конкретных длин входных буферов, что позволяет анализировать обобщенное состояние программы вместо одного конкретного.

Обрабатываемые в анализируемой программе буферы аннотируются: каждому буферу приписывается его длина, которая затем распространяется вместе с данными буфера. Для описания операционной семантики машинных кодов трассы программы используется промежуточное представление Pivot, в которое отображаются различные аппаратные архитектуры, что позволяет унифицировать алгоритмы анализа и символьной интерпретации. В рамках символьной интерпретации трассы могут интерпретироваться как на уровне отдельных инструкций, так и на уровне целых вызовов функции. Ошибки доступа к памяти обнаруживаться при выполнении отдельных инструкций и при вызове функций работы со строками. Границы буферов, с которыми работает программа, определяются с помощью анализа карты динамической памяти, а также анализа статического представления анализируемой программы. Также в главе описываются вспомогательные методы, позволяющие увеличить точность анализа: метод предварительного расширения покрытия кода и метод восстановления функций работы со строками.

Третья глава посвящена описанию программной реализации предложенных методов. Метод символьной интерпретации трасс реализован на базе среды анализа бинарного кода, разрабатываемой в ИСП РАН. Анализируемые трассы получаются с помощью эмулятора QEMU. Следует отметить вклад автора диссертационной работы в описание машинной семантики машинных инструкций различных процессорных архитектур в рамках промежуточного представления Pivot, конкретные примеры которых приведены **в приложении** к работе. Метод предварительного расширения покрытия кода, позволяющий более полно анализировать программу, реализован на базе платформы динамического символьного выполнения S2E.

В четвёртой главе приведены результаты применения предложенных методов на различных примерах. Для оценки возможностей обнаружения выхода за границы буфера использованы приложения для ОС Windows и Linux, а также встроенное программное

обеспечение для сетевого маршрутизатора. Методы продемонстрировали свою эффективность, позволив обнаружить ошибки во всех приведенных программах, в части из которых были подтверждены ранее известные серьезные уязвимости. На нескольких примерах детально рассмотрены различные ситуации, для которых удалось обнаружить ошибку выхода за границы буфера.

В заключении перечисляются основные результаты работы, приводятся особенности разработанного метода, а также обозначаются направления дальнейшего развития. Перечислим основные результаты, полученные в ходе диссертационного исследования:

1. Разработан метод поиска ошибок выхода за границы буфера на основе символьной интерпретации трасс выполнения с использованием абстрактной длины буферов. Для работы метода не требуется наличие исходных кодов и отладочной информации. Кроме того, метод позволяет абстрагироваться от используемой процессорной архитектуры, а также операционной системы, для которых предназначена исследуемая программа.
2. Разработаны методы, повышающие точность поиска ошибок: метод выявления функций работы со строками, которые подверглись встраиванию в процессе компиляции, а также метод предварительного расширения покрытия кода.

Результаты, выносимые на защиту, опубликованы в четырёх публикациях в рецензируемых журналах, рекомендованных ВАК, один из которых индексируется Scopus, а также в двух публикациях в трудах научной конференции.

В диссертационной работе следует отметить ряд **недостатков**:

1. При описании результатов подробно проанализированы только 4 примера, но явно не сказано, что нарушение границ буфера было найдено во всех приведенных примерах программ.

2. В работе сказано, что приведенные методы не зависят от процессорной архитектуры, однако результаты представлены только для платформы Intel x86, x86-64.

3. Трудно сопоставить время, необходимое на обнаружение ошибки нарушения границ буфера с размером программы и объемом трассы, необходимой для ее обнаружения.

4. Не всегда понятны обозначения (например, промежуточное представление Pivot, примеры которого имеются в приложении А), некоторые понятия вводятся отсылками к опубликованным работам, что затрудняет восприятие сатериала.

Перечисленные недостатки не влияют на общую положительную оценку работы.

Диссертационная работа В.В. Каушана является законченным научным исследованием. Основное содержание диссертации отражено в опубликованных диссертантом статьях, доложено на научных конференциях. Практическая ценность подтверждается результатами применения разработанных методов.

Автореферат полно и правильно отражает содержание диссертационной работы.

Вывод

Диссертационная работа соответствует всем требованиям ВАК, предъявляемых к диссертациям на соискание учёной степени кандидата технических наук, а её автор, Каушан Вадим Владимирович, заслуживает присуждения ему учёной степени кандидата технических наук по специальности 05.13.11 — «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Официальный оппонент:

Кандидат технических наук, старший научный сотрудник, начальник отделения «Системы программирования» публичного акционерного общества «ИНЭУМ им. И.С. Брука»

В.Ю. Волконский

"29" января 2018 г.

Подпись кандидата технических наук Волконского В.Ю. заверяю, зам. генерального директора ПАО «ИНЭУМ им. И.С. Брука» по науке

В.И. Перекатов