ОТЗЫВ НАУЧНОГО РУКОВОДИТЕЛЯ

на диссертацию Дудиной Ирины Александровны «Поиск ошибок переполнения буфера в исходном коде программ с помощью символьного выполнения», представленную на соискание ученой степени кандидата физико-математических наук по специальности 05.13.11 — математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей

Статический анализ программ повсеместно используется для поиска ошибок в их исходном коде, что в свою очередь влечёт как повышение их уровня безопасности, так и качества. Пожалуй, самым известным видом критических ошибок является переполнение буфера из-за своей распространённости и разрушительных последствий для программ на языках C и C++. Эти обстоятельства определили цель диссертационной работы И. А. Дудиной — построить методы поиска переполнений буфера, которые годятся для проверки современных программных проектов размером в 10^8 и более строк кода, то есть показывают приемлемое время анализа (6-8 часов) с высоким процентом истинных срабатываний (не менее 60%) и, по возможности, с пропуском небольшого числа ошибок.

Актуальность поставленной задачи следует из отсутствия открытых методов поиска переполнений, которые бы соответствовали указанным критериям; подходы, применяемые в закрытых коммерческих анализаторах, известны лишь в самых общих чертах. Инструмент Svace, разрабатываемый в Институте системного программирования им. В. П. Иванникова РАН, до работы И. А. Дудиной не поддерживал поиск переполнений буфера, чувствительный к путям выполнения, и тем самым пропускал существенное число ошибок.

В качестве базового подхода для выполняемого статического анализа диссертантка использовала статическое выполнение с объединением состояний. Сделанный выбор обоснован как тщательно выполненным обзором существующих методов анализа, приложимых к поиску переполнений, так и анализом известных тестовых наборов для инструментов поиска ошибок и уязвимостей из базы CVE. В рамках выбранного подхода И. А. Дудиной выполнено системное построение решения поставленной задачи. Во-первых, формализовано определение ошибки доступа к буферу и на основе него предложен способ поиска таких ошибок, заключающийся в построении достаточных условий ошибки специального вида, который удобен для самого анализа и для проверки SMT-решателями. Во-вторых, описан модельный язык и для него разработан внутрипроцедурный алгоритм построения указанных условий на основе символьного выполнения. При этом полностью обоснована не только корректность выполняемого анализа, но и сформулированы и доказаны теоремы об ограничениях на входные программы, при выполнении которых предложенные абстракции анализа будут точными, то есть совместность построенных формул означает существование конкретного выполнения анализируемой функции, на котором происходит

ошибка. В-третьих, разработанные алгоритмы расширены для межпроцедурного анализа, и аналогично доказана их корректность. Наконец, не оставлены без внимания важные типы переполнений — ошибки при доступе к буферам динамического размера, переполнения при работе со строками и с данными из недоверенных источников. Для поиска всех этих типов переполнений диссертанткой описаны вариации базового алгоритма анализа.

Все предложенные методы были реализованы в анализаторе Svace в семи различных детекторах переполнения буфера. Результаты тестирования на пакете Juliet Test Suite показали покрытие в среднем около 50%, при этом ложных срабатываний при поиске переполнений правой границы буфера не было, а при поиске левой границы их доля составила всего 0,02%. Анализ не найденных детекторами ошибок показывает, что существенная доля пропусков связана с ограничениями ядра анализатора (в частности, с недостаточно точной обработкой вызовов виртуальных функций и коллекций C++), а не с ограничениями собственно предложенного диссертанткой подхода. Покрытие на этом же тестовом наборе лучшего из доступных открытых анализаторов, инструмента Infer, составляет около 10%.

Результаты тестирования на больших программных продуктах (ОС Android и Tizen) показали количество истинных срабатываний в среднем около 70%. Доля детекторов переполнения в общем времени анализа, как правило, составляет менее 5%. Показанные результаты позволили с успехом использовать разработанные детекторы в составе анализатора Svace в промышленном цикле разработки программ в компании «Самсунг» и у ряда российских заказчиков.

Считаю, что диссертационная работа соответствует всем требованиям, предъявляемым ВАК при Минобрнауки РФ к работам на соискание ученой степени кандидата физико-математических наук по специальности 05.13.11 — математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей, а её автор, Дудина Ирина Александровна, безусловно заслуживает присуждения ей ученой степени кандидата физико-математических наук по указанной специальности.

Научный руководитель:
в.н.с. ИСП РАН, д. ф.-м. н. А. А. Белеванцев

15 февраля 2019 г.

Подпись Белеванцева А. А. удостоверяю:

Директор ИСП РАН А. И. Аветисян