ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА 24.1.120.01,
созданного на базе
Федерального государственного бюджетного учреждения науки
Институт системного программирования им. В.П. Иванникова
Российской академии наук
Министерства науки и высшего образования РФ
по диссертации на соискание ученой степени доктора наук

аттестационное дело № _____

решение диссертационного совета от 17 декабря 2024 года № 2024/16

О присуждении Саргсяну Севаку Сениковичу, гражданину Республики Армения (РА), ученой степени доктора технических наук.

Диссертация «Методы оптимизации алгоритмов статического и динамического анализа программ» по специальности 2.3.5 – «математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей» принята к защите 16 сентября 2024, протокол № 2024/11 диссертационным советом 24.1.120.01, созданным на базе Федерального государственного бюджетного учреждения науки Институт системного программирования им. В.П. Иванникова Российской академии наук (ведомственная принадлежность: Министерство науки и высшего образования РФ; адрес: 109004, г. Москва, ул. А. Солженицына, дом 25), создан Приказом Минобрнауки России о советах по защите докторских и кандидатских диссертаций от 2 ноября 2012 г. № 714/нк.

Соискатель Саргсян Севак Сеникович, 1989 года рождения.

Диссертацию на соискание ученой степени кандидата физико-математических наук «Методы поиска клонов кода и семантических ошибок на основе семантического анализа программы» защитил в 2016 году в диссертационном совете, созданном на базе Федерального государственного бюджетного учреждения науки Институт системного программирования Российской академии наук.

Работает заведующим кафедрой системного программирования в государственном образовательном учреждении высшего профессионального образования Российско-Армянский (Славянский) университет (ведомственная принадлежность: Министерство науки и высшего образования РФ, Министерство образования, науки, культуры и спорта РА).

Диссертация выполнена в государственном образовательном учреждении высшего профессионального образования Российско-Армянский (Славянский) университет (ведомственная принадлежность: Министерство науки и высшего образования РФ, Министерство образования, науки, культуры и спорта РА).

Научный консультант – доктор физико-математических наук, академик РАН, Аветисян Арутюн Ишханович, Федеральное государственное бюджетное учреждение науки «Институт системного программирования им. В.П. Иванникова Российской академии наук», директор.

Официальные оппоненты:

1. Шабанов Борис Михайлович, доктор технических наук, член-корреспондент РАН, Национальный исследовательский центр «Курчатовский институт», заместитель директора по исследованиям в области информационных технологий,
2. Кознов Дмитрий Владимирович, доктор технических наук, доцент, Санкт-Петербургский государственный университет, профессор,
3. Кореньков Владимир Васильевич, доктор технических наук, старший научный сотрудник, Объединенный институт ядерных исследований, научный руководитель

дали положительные отзывы на диссертацию.

Ведущая организация Федеральное государственное учреждение "Федеральный исследовательский центр Институт прикладной математики им. М.В. Келдыша Российской академии наук" (г. Москва) в своем положительном заключении, подписанном Поляковым Сергеем Владимировичем, доктором физико-математических наук, ведущим научным сотрудником, указала, что диссертационная работа является самостоятельной и законченной научно-исследовательской работой, обладающей высокой научной

и практической значимостью, решающей важную проблему поиска ошибок и уязвимостей в больших комплексах программ путем статического и динамического анализа.

Соискатель имеет более 30 опубликованных работ, в том числе по теме диссертации опубликовано 24 работ, из них 12 работ в изданиях, входящих в список изданий, рекомендованных ВАК РФ, кроме того, 11 статей опубликовано в изданиях, индексируемых Scopus и Web of Science. Еще одна статья опубликована в IEEE Access, входящем в первый квартиль SJR. Получено 7 свидетельств о регистрации программ для ЭВМ.

Наиболее значимые работы соискателя:

1. S. Sargsyan, S. Asryan, J. Hakobyan, S. Kurmangaleev, "Combining dynamic symbolic execution, code static analysis and fuzzing", *Proceedings of the Institute for System Programming*, vol. 30, pp. 25-38, 2018.
2. S. Sargsyan, J. Hakobyan, H. Movsisyan, S. Kurmangaleev, V. Sirunyan, M. Mehrabyan, "Improving fuzzing performance by applying interval mutations", *Proceedings of the Institute for System Programming*, vol. 31, № 1, pp. 78-88, 2019.
3. S. Sargsyan, V. Vardanyan, H. Aslanyan, M. Harutunyan, M. Mehrabyan, K. Sargsyan, H. Hovahannisyan, H. Movsisyan, J. Hakobyan, S. Kurmangaleev, "GENES ISP: Code analysis platform", *Proceedings of Ivannikov Ispras Open Conference*, 2020.
4. С. Саргсян, В. Варданян, Д. Акопян, А. Агабалян, М. Меграбян, Ш. Курмангалеев, А. Герасимов, М. Ермаков, С. Вартанов, "Платформа автоматического фаззинга программного интерфейса приложений", *Труды Института системного программирования РАН*, т. 32, № 2, с. 161-173, 2020.
5. S. Sargsyan, J. Hakobyan, M. Mehrabyan, R. Mkoyan, V. Sahakyan, V. Melkonyan, M. Arutunian, A. Fahradyan, A. Avetisyan, "Advanced Grammar-Based Fuzzing", *Proceedings of Ivannikov Memorial Workshop*, 2022.
6. S. Sargsyan, S. Kurmangaleev, A. Belevantsev, A. Avetisyan, "Scalable and accurate detection of code clones", *Programming and Computer Software*, vol. 42, pp. 27-33, 2016. (Q3)
7. С. Саргсян, Ш. Курмангалеев, А. Белеванцев, А. Асланян, А. Балоян, "Масштабируемый инструмент поиска клонов кода на основе семантического анализа программ", *Труды Института системного программирования РАН*, т. 27, № 1, с. 39-50, 2015.

8. С. Саргсян, "Поиск семантических ошибок, возникающих при некорректной адаптации скопированных участков кода", *Труды Института системного программирования РАН*, т. 27, № 2, с. 93-104, 2015.
9. S. Sargsyan, "Improving Fuzzing Using Input Data Offsets Comparison Information", *Programming and Computer Software*, vol. 49, pp. 122-127, 2023. (Q3)
10. А. Асланян, Ш. Курмангалеев, В. Варданян, М. Арутюнян, С. Саргсян, "Платформенно-независимый и масштабируемый инструмент поиска клонов кода в бинарных файлах", *Труды Института системного программирования РАН*, т. 28, № 5, с. 215-226, 2016.
11. С. Асрян, С. Гайсарян, Ш. Курмангалеев, А. Агабалян, Н. Овсепян, С. Саргсян, "Обнаружение ошибок, возникающих при использовании динамической памяти после освобождения", *Труды Института системного программирования РАН*, т. 30, № 3, с. 7-20, 2018.
12. S. Asryan, S. Gaissaryan, S. Kurmangaleev, S. Sargsyan, A. Aghabalyan, N. Hovsepyan, "Dynamic Detection of Use-After-Free Bugs", *Programming and Computer Software*, vol. 45, № 7, pp. 365-371, 2019. (Q3)
13. H. Aslanyan, H. Movsisyan, M. Arutunian, S. Sargsyan, "Bin2Source: Matching Binary to Source Code", *Proceedings of Ivannikov Ispras Open Conference*, 2021.
14. Ш. Курмангалеев, В. Корчагин, В. Савченко, С. Саргсян, "Построение обфусцирующего компилятора на основе инфраструктуры LLVM", *Труды Института системного программирования РАН*, т. 23, с. 77-92, 2012.
15. H. Aslanyan, H. Movsisyan, H. Hovhannisyan, Z. Gevorgyan, R. Mkoyan, A. Avetisyan, S. Sargsyan, "Combining Static Analysis with Directed Symbolic Execution for Scalable and Accurate Memory Leak Detection", *IEEE Access*, vol. 12, pp. 80128-80137, 2024. (Q1)

Выбор официальных оппонентов и ведущей организации обосновывается их компетентностью и достижениями в данной отрасли науки, наличием публикаций в сфере исследований, соответствующей теме диссертации, и способностью определить научную и практическую ценность диссертации.

Диссертационный совет отмечает, что на основании выполненных соискателем исследований:

- Разработана и реализована платформа анализа программ, обеспечивающая: сбор артефактов для большого объема открытого ПО и информации об известных уязвимостях; единообразный подход комбинирования различных методов анализа кода в зависимости от конкретной задачи. В данную платформу интегрирован ряд разработанных диссертантом методов

статического и динамического анализа программ, которые могут использоваться в различных комбинациях в зависимости от конкретной задачи.

- Разработаны и реализованы методы поиска клонов кода для исходного и бинарного кода, основанные на поиске схожих подграфов максимального размера, что одновременно повышает масштабируемость и точность.
- Разработан и реализован метод сопоставления исходных и бинарных файлов, который сначала производит компиляцию входного исходного кода с разными уровнями оптимизации, а далее применяет поиск клонов бинарного кода для сопоставления бинарных файлов.
- Разработан и реализован метод поиска утечек памяти, который на первом этапе производит поиск утечек на специальном графовом представлении программы на языке C/C++, а затем производит направленное символьное выполнение для проверки выполнимости путей ошибок.
- Разработан и реализован метод фаззинга программ для генераций структурированных данных на основе специализированных автоматов BNF-грамматик.
- Разработан и реализован метод фаззинга интерфейсных функций с возможностью генераций цепочки вызовов функций, где возвращаемые значения одних функций используются в качестве аргументов для других. Такой подход обеспечивает возможность автоматической подготовки необходимых ресурсов для тестирования сложных сценариев использования нескольких интерфейсных функций.
- Разработан и реализован метод направленного фаззинга для выполнения конкретных инструкций или фрагментов программы, содержащией потенциальные уязвимости или дефекты. Данный метод, как правило, комбинируется со статическим анализом для определения целевых инструкций или фрагментов кода с ошибками.
- Разработан и реализован метод интеграции статического анализа с фаззингом, с помощи которого возможно выполнить обе ветви тех

условных операторов, в которых производится сравнение ячеек входного буфера с константными значениями.

- Проанализированы десятки тысяч проектов, с применением разработанной платформы и методов, суммарный объем которых превышает сотни миллионов строк исходного и соответствующего бинарного кода. В результате стало возможно найти более 90 ошибок разного типа в открытом и проприетарном ПО, включая ошибки во многих широко используемых проектах. Десятки из этих ошибок были исправлены и представлены сообществу разработчиков открытого ПО, что позволило улучшить безопасность и стабильность исследованных систем.

Теоретическая значимость исследования обоснована тем, что:

- Предложена концепция платформы анализа программ, которая позволяет путем универсальной комбинации нескольких методов, в зависимости от конкретной задачи, найти новые сложно-находимые ошибки. Эффективность предложенного метода обосновывается экспериментальными результатами.
- Разработан метод поиска клонов в исходном и бинарном коде, который по своим результатам превосходит существующие подходы.
- Предложен метод сопоставления исходных и бинарных файлов, который превосходит существующие аналоги.
- Разработан двухэтапный метод поиска утечек памяти, который показывает наилучшие результаты при сравнении с существующими методами.
- Разработаны методы фаззинга, которые позволяют увеличить покрытие кода в разных сценариях анализа.

Значение полученных соискателем результатов исследования для практики подтверждается тем, что:

- Разработанная платформа и методы внедрены в компаниях "Базальт СПО" и ООО "РусБИТех-Астра".
- Разработанная платформа используется в жизненном цикле разработки ПО в ИСП РАН и ЦППТ РАУ с 2021.

- С применением разработанной платформы и методов найдены более 90 ошибок, включая ошибки в таких широко используемых проектах, как openssl, grub2, clib, ffmpeg, что является важным вкладом в повышение уровня защищенности базового программного обеспечения.

Оценка достоверности результатов исследования выявила:

- в работе корректно применяются классические методы исследования;
- проводится анализ эффективности разработанных алгоритмов на экспериментальных данных;
- десятки найденных ошибок, с применением разработанных методов, подтверждены и исправлены в ряде открытых (open source) проектов;
- разработанные алгоритмы и полученные экспериментальные результаты соответствуют мировому уровню.

Личный вклад соискателя состоит в личном участии на всех этапах процесса разработки и реализации предложенных методов анализа. В опубликованных совместных работах постановка и исследование задач осуществлялись совместными усилиями соавторов при непосредственном участии соискателя. Выносимые на защиту результаты получены соискателем лично.

В ходе защиты диссертации были высказаны следующие критические замечания:

- Некоторые разделы работы требуют более детального описания технической реализации предложенных методов. Также не хватает описания применимости представленных методов для анализа ПО, созданного на других языках программирования, отличных от C/C++.
- В исследовании не рассмотрены нейросети для решения выбранных задач, а между тем этот подход может оказаться весьма полезным для уменьшения вычислительной сложности алгоритмов анализа кода, то есть это перспективное направление предложения представленных исследований.
- Не всегда обосновывается выбор операционных систем, программ и уязвимостей, участвующих в тестировании.

Соискатель Саргсян С.С. согласился с замечаниями, ответил на задаваемые ему в ходе заседания вопросы.

На заседании 17 декабря 2024 диссертационный совет принял решение за разработку платформы анализа программ и теоретических положений, совокупность которых можно квалифицировать как научное достижение и решение научной проблемы, имеющей важное научное и практическое значение присудить Саргсяну С.С. ученую степень доктора технических наук.

При проведении тайного голосования диссертационный совет в количестве 16 человек, из них 9 докторов наук по специальности рассматриваемой диссертации, участвовавших в заседании, из 20 человек, входящих в состав совета, проголосовали: за – 16, против – 0.

Заместитель председателя диссертационного совета,
доктор физико-математических наук Петренко А. К.

Ученый секретарь диссертационного совета,
кандидат физико-математических наук Зеленов С. В.

17 декабря 2024