

Предисловие

Очередной, 20-й том Трудов Института системного программирования РАН содержит 14 статей, посвященных различным аспектам системного программирования.

В статье А.Ю.Тихонова и А.И. Аветисяна «Развитие taint-анализа для решения задачи поиска некоторых типов закладок» задача анализа программного обеспечения (ПО) рассматривается как задача получения некоторых свойств исследуемого ПО. Этими свойствами могут быть и описания реализуемых алгоритмов, и информация о структурах файлов, данных в памяти или пакетов сетевых протоколов, и информация об имеющихся ошибках. В статье рассматривается подход к анализу программного обеспечения, представленного в виде исполняемого кода при отсутствии исходных текстов, с целью выявления некоторых видов программных закладок, которые, в соответствии с ГОСТ Р 51275-2006 являются преднамеренно внесенными в ПО функциональными объектами, при определенных условиях инициирующими реализацию недеklarированных возможностей ПО.

Статья Д. В. Силакова «Использование аппаратной виртуализации в контексте информационной безопасности» посвящена возможностям использования аппаратной виртуализации для решения различных задач информационной безопасности. Предлагается обзор подходов к повышению безопасности программных систем, основанных на использовании виртуализации. Приводится обзор возможных сценариев использования виртуализации злоумышленниками. Указываются области применения и ограничения существующих решений и дальнейшие перспективы развития рассматриваемой области.

Кирилл Батузов и Алексей Меркулов представили статью «Оптимизация динамической двоичной трансляции». Двоичная трансляция – это процесс получения по заданной программе Р программы Q, удовлетворяющей заданным требованиям, если обе программы записаны в виде машинных кодов. Если двоичная трансляция производится во время выполнения программы, то она называется динамической двоичной трансляцией. В статье рассматриваются возможности применения различных оптимизаций во время динамической двоичной трансляции.

В статье «Декларативный интерфейс поиска дефектов по синтаксическим деревьям: язык KAST» ее автор С. В. Сыромятников отмечает, что во многих случаях дефекты программного кода могут быть выявлены путём анализа соответствующего синтаксического дерева. Рассматриваются преимущества и недостатки данного подхода в сравнении с более сложными видами статического анализа, и обосновывается необходимость предоставления пользователю интерфейса для написания собственных обнаружителей дефектов. Обсуждаются различные подходы к реализации подобного

интерфейса. Описывается новый декларативный язык, позволяющий пользователю описывать дефекты кода в виде шаблонов для синтаксических деревьев, и рассматриваются некоторые аспекты работы анализатора этого языка.

В статье Л. В. Дворянского «О моделировании счётчиков с бесконечным числом значений в обыкновенных сетях Петри» проводится анализ моделирования в обыкновенных сетях Петри счётчиков с бесконечным числом состояний. Обосновывается выбор отношения эквивалентности симуляции готовности в качестве отношения реализации для моделирования счётчиков. Демонстрируется, что в сетях Петри невозможно промоделировать счётчики с бесконечным числом состояний. Представлена минимальная модель счётчика с конечным числом значений.

В.В. Липаев в своей статье «Кросс-система программирования ЯУЗА-6 для специализированных ЭВМ реального времени» рассматривает особенности и проблемы эффективного создания в 70-е годы сложных комплексов программ оборонного применения, базирующихся на специализированных ЭВМ. В статье представлены методы, требования и реализация адаптируемых кросс-систем автоматизации программирования и тестирования для различных типов сложных комплексов программ объектных ЭВМ. Изложены результаты многолетнего применения на БЭСМ-6 в ряде предприятий кросс-системы автоматизации программирования и отладки ЯУЗА-6.

Тематике систем реального времени посвящена и статья С.В. Зеленова «Планирование строго периодических задач в системах реального времени». Как указывает автор, одним из важнейших аспектов функционирования систем реального времени является планирование задач. Классические алгоритмы планирования периодических задач работают лишь в случае, когда время запуска каждой задачи может варьироваться внутри разных периодов ее выполнения. Однако в настоящее время имеется потребность в составлении таких расписаний, в которых время между соседними запусками одной периодической задачи было бы фиксировано и равнялось бы длине периода. Предлагается алгоритм построения расписаний, близких к оптимальным в смысле минимизации общего количество прерываний задач, с учетом этого дополнительного требования.

В статье Н. В. Пакулина и А. Н. Тугаенко «Тестирование протоколов электронной почты Интернета с использованием моделей» рассматриваются вопросы тестирования почтовых протоколов с использованием формальных моделей протоколов: предложен метод моделирования почтовых протоколов, рассмотрены особенности почтовых протоколов в контексте тестирования, представлены результаты тестирования популярных почтовых серверов с открытым кодом. В качестве примера представлена разработка тестовых наборов для протоколов SMTP и POP3 на языке JavaTESK – расширении языка Java, реализующем тестирование с применением формальных методов. Тестовые наборы состоят из двух частей: независимого тестирования

соответствия протоколов спецификациям и совместного теста, имитирующего работу почтовых протоколов в сети.

Тема тестирования на основе применения формальных методов продолжается в статье А.С. Камкина и М.М. Чупилко «Механизмы поддержки функционального тестирования моделей аппаратуры на разных уровнях абстракции». На разных этапах проектирования аппаратуры используются разные представления целевой системы. Соответственно, в зависимости от зрелости проекта применяются разные методы верификации, в частности, разные методы построения эталонных моделей, используемых для оценки корректности проектируемой аппаратуры. Различие в уровнях абстракции усложняет повторное использование тестовых систем, созданных в начале проектирования, для верификации аналогичных компонентов, но на более поздних этапах. В статье предлагается подход к построению эталонных моделей аппаратуры и тестовых оракулов на их основе, упрощающий повторное использование тестовых систем и снижающий затраты на верификацию.

В статье В.С. Мутилина, Е.М. Новикова, А.В. Страха, А.В. Хорошилова и П.Е. Шведа «Архитектура Linux Driver Verification» статье исследуются требования к построению архитектуры открытой системы верификации, которая могла бы предоставить площадку для экспериментов с различными методами статического анализа кода на реальном программном обеспечении и в то же время являлась полноценной системой верификации, готовой к индустриальному применению. По результатам обсуждения требований предлагается архитектура такой системы верификации и детально рассматриваются ее компоненты. В заключении описывается имеющийся практический опыт работы с предложенной архитектурой и предлагаются пути дальнейшего развития.

В статье С.Д. Кузнецова «Транзакционные параллельные СУБД: новая волна» отмечается, что в области систем управления данными без совместного использования ресурсов образовались два фронта: "NoSQL", где отрицаются основные принципы, свойственные СУБД, и "один размер непригоден для всех", где упор делается на специализацию систем при сохранении важнейших свойств СУБД. Особенно интересным является противостояние этих фронтов в области "транзакционных" систем управления данными. Опираясь на "теорему" CAP Эрика Брюера, представители лагеря NoSQL отказываются от обеспечения в своих системах традиционных свойств ACID. В статье обсуждается суть "теоремы" Брюера, и обосновывается, что она не имеет отношения к свойствам ACID. Рассматриваются современные исследовательские работы, обеспечивающие классические ACID-транзакции в параллельных средах без общих ресурсов, а также наиболее здравые подходы, в которых из чисто прагматических соображений свойства ACID частично ослабляются.

Е. А. Костычев, В. А. Омельченко и С.В. Зеленев в статье «Нацеленная генерация данных для тестирования приложений над базами данных» справедливо утверждают, что широко распространенным типом программного обеспечения являются приложения, обрабатывающие большие массивы данных. В частности, такие приложения решают задачи интеграции данных в области интеграции корпоративных приложений. Специфика функционального тестирования подобных приложений заключается в большом количестве возможных комбинаций входных данных. При использовании существующих подходов гарантированное покрытие функциональности тестируемого приложения может быть достигнуто только полным перебором возможных комбинаций исходных данных. При помощи подхода генерации данных, предлагаемого в статье, возможно достижение покрытия функциональности приложения с близким к оптимальному объемом данных.

В статье А.В. Коршунова «Извлечение ключевых терминов из сообщений микроблогов с помощью Википедии» описывается способ извлечения ключевых терминов из сообщений микроблогов с использованием информации, полученной путём анализа структуры и содержимого интернет-энциклопедии Википедия. Работа алгоритма основана на расчёте для каждого термина его «информативности», т.е. оценки вероятности того, что он может быть выбран ключевым в тексте. В ходе тестирования разработанный алгоритм показал удовлетворительные результаты в условиях поставленной задачи, существенно опережая аналоги. В качестве демонстрации возможного применения разработанного алгоритма был реализован прототип системы контекстной рекламы. Сформулированы также варианты использования информации, полученной путём анализа сообщений Twitter, для реализации различных вспомогательных сервисов.

Завершает том статья И. С. Таранова «Использование префиксного дерева для хранения и поиска строк во внешней памяти». В этой статье представлена структура данных для поиска и эффективного хранения во внешней памяти массивов текстовых строк, реализованная для поддержки индексов в XML СУБД Sedna. Описываются алгоритмы для вставки, удаления и поиска строк переменной длины в префиксных деревьях, хранимых на дисках. Приводятся результаты сравнения реализации новой структуры с существующей реализацией В-дерева. Показано, что в некоторых случаях предложенная структура данных занимает в несколько раз меньше места во внешней памяти при той же скорости поиска.

Академик РАН В.П. Иванников