Математические модели и актуальные алгоритмические проблемы, связанные с оптимизацией нагрузки и защите информации в распределенных вычислительных системах

- Н.Н. Кузюрин

Алгоритмы оптимизации управления потоком задач в распределенных вычислительных системах

- Однородный кластер
- Grid
- Cloud

История

- Теория расписаний
- Задача о m машинах (Грэхем, 1966)
- NP-трудна
- Приближенные алгоритмы

Распределение задач на группе кластеров

- Задача хорошо исследована для одного кластера
- Для группы не рассматривалась
- Мы начали ее исследование в 2003 г.
- Естественные эвристики не гарантируют константной точности
- (возможен растущий дисбаланс)

Двухуровневое планирование

- Искусственные ограничения: маленькие задачи не размещаются на большие кластеры

- Онлайновые алгоритмы с гарантией точности константной (IEEE, 2004)

Система моделирования: анализ открытых workload данных

- Моделирование системы Sharcnet
  - Изначально большой дисбаланс
  - Эксперименты показали значительное уменьшение времени ожидания
- Анализ других Grid-систем
  - Grid500
  - DAS2

Задача упаковки прямоугольников в несколько полос

- В одну полосу Strip Packing (1980) - классика
- Мы начали исследование в 2003 г.

- Предложены онлайновые алгоритмы гарантирующие точность 2e
- Получена нижняя оценка точности e

Исследования за рубежом

- Нашу модель начали исследовать во Франции, Германии, Китае, Японии с 2009 г.

- Multiple Strip Packing (все полосы имеют одинаковую ширину)

Разные производительности

- Related machines
- Задача трудна даже для случая однопроцессорных задач

- Наша модель и результаты обобщены на случай параллельных задач и кластеров с различной производительностью с теми же оценками точности

Математика

- Предложен новый онлайновый алгоритм упаковки, значительно превосходящий результат Коффмана-Шора (1993 г.) по точности

Сверхбольшие графы

- Модель для сетей в разных областях: биологии, компьютерных сетях, социальных сетях и др.
- Классическая теория случайных графов (Эрдеш-Реньи) плохо описывает реальные сети
- 2000 г. – power law distribution
- Размер графов до миллиарда вершин

Алгоритм порождения

- Много генераторов без обоснования свойств получаемых графов
- Нами предложен распределенный алгоритм порождения случайных графов, имеющих порядка миллиарда вершин
- Вычисления на облаке Amazon и верификация свойств графов
- Complex Networks, Italy, 2014

Обфускация программ

Обфускация программ –преобразование O программы π, к виду $O(\pi)$ при котором сохраняются функциональные возможности исходной программы и $O(\pi)$ становится **виртуальным «черным ящиком»**

Потенциальные применения обфускации

- защита интеллектуальной собственности на программное обеспечение,
- маскировка «водяных знаков» и «отпечатков пальцев»,
- защита мобильных агентов в компьютерных сетях,
- защита облачных вычислений в виртуальных сетях,
- защита проектных решений микроэлектронных схем,

- Построение анонимных каналов, в частности для протоколов голосования
- обеспечение конфиденциальности запросов к базам данных
- маскировка вирусов,
- сокрытие уязвимостей и недекларированных возможностей программ.

Современное состояние исследований проблемы обфускации

1. Проблема обфускации стимулировала развитие методов декомпиляции, обратной инженерии, статического и динамического анализа программ.
2. Построено много коммерческих обфускаторов программ. Стойкость этих обфускаторов не исследована.
3. Много вариантов определений стойкости обфускации.
4. Доказано существование функций, не допускающих обфускации.
5. Доказано, что некоторые программы (перешифрования, проверки пароля) можно обфускировать.

Вклад ИСП РАН

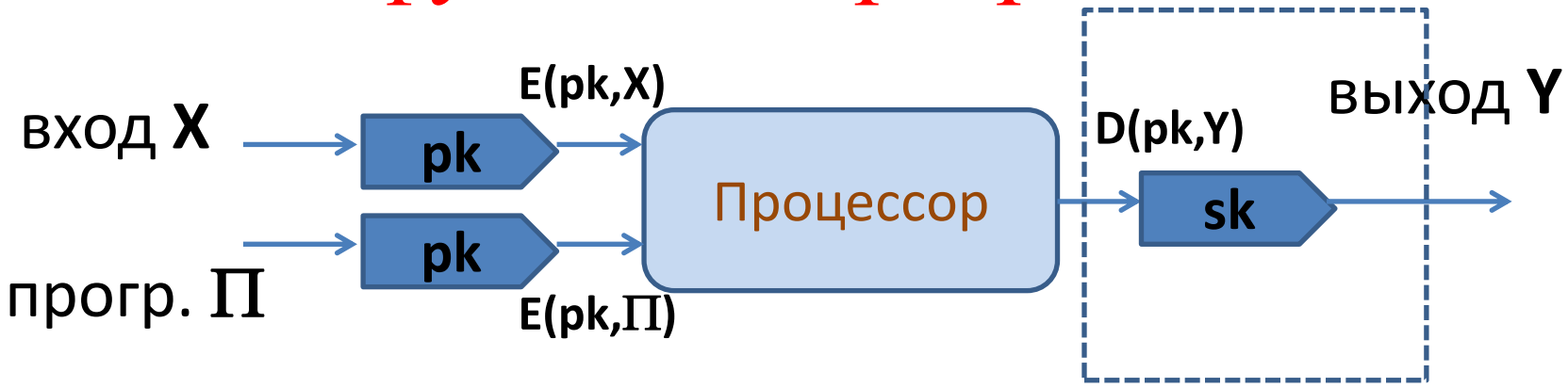
1. Разработаны новые определения стойкости обфускации.
2. Разработаны методы информационной защиты проектных решений микроэлектронных схем.
3. Разработана экспериментальная система обфускации С программ (Poirot) для противодействия алгоритмам статического анализа программ.
4. Защищена 1 диссертация (всего в мире - 10)
5. 15 публикаций (всего в мире - 110~120)

Новые подходы к проблеме обфускации программ

В основу положен метод гомоморфного шифрования (K Gentry, 2009).

Выбирается гомоморфная система шифрования с открытым ключом (*KG,E,D*). При помощи открытого ключа pk шифруются входные данные X и программа П, а при помощи секретного ключа расшифровывается результат Y.

Новые подходы к проблеме обфускации программ



В этой схеме нужно обфускировать лишь процедуру дешифрования. Таким образом задача обфускации произвольной программы сводится к задаче обфускации процедур дешифрования в гомоморфных криптосистемах.

Новые надежды

- Положительные результаты о стойкой обфускации (реанимация интереса)
- Результаты о схемах гомоморфного шифрования - чисто теоретические

- Кажется, что они (хотя бы теоретически) решают проблему защиты данных в облачных вычислениях

Новые подходы к защите информации в облачных вычислениях

- Простая модель облачных вычислений всего лишь с двумя пользователями
- Стойкая защита – невозможна!

  (Ван Дейк и др., 2010)

- Альтернативный подход (ИСП РАН)
- Использование пороговых гомоморфных криптосистем с открытым ключом

Направления исследований

- Теория рекурсии
- Теория автоматов
- Теория верификации и проблемы эквивалентности
- Комбинаторика и теория графов
- Эффективные алгоритмы
- Алгоритмические методы в алгебре (базисы Гребнера)