



Распознавание приложений по запросам блочного уровня

Демьяненко Илья

Руководитель: Платонов С. М.

QoS в RAIDIX

- **У задач есть дедлайны**
 - Не все задачи одинаково важны
- **Приоритет на уровне инициаторов**
 - Ручное управление приоритетами
- **Необходимость вовлечения администратора**

Решение: автоматическое распределение приоритетов

Исходные данные

- **Нет информации о файлах**
- **Поток запросов блочного уровня**
 - Инициатор
 - Время
 - Тип запроса
 - Название массива
 - Адрес
 - Размер
- **Реализован сборщик трейсов на Python**

Постановка задачи

- Исследовать закономерности в запросах
- Научиться распознавать приложения
- Разработать прототип
- Оценить эффективность

Выборочное распознавание

- **Ограничим класс приложений**
 - Бэкап
 - Копирование файлов
 - Бенчмарки
- **Предсказуемость**
 - Последовательные чтение/запись
 - Блоки одинакового размера
- **Искажения**
 - Файловая система
 - Транспорт

Выборочное распознавание

- **Обучение**

- Группируем запросы по последовательностям
- Ищем возможные объединения в блоки
- Запоминаем варианты разбиения

- **Распознавание**

- Ведём учёт последовательностей
- Запросы либо продолжают последовательность, либо образуют новую
- Удалось классифицировать новую последовательность — приложение обнаружено
- Последовательность давно не продолжалась — приложение завершилось

Выборочное распознавание

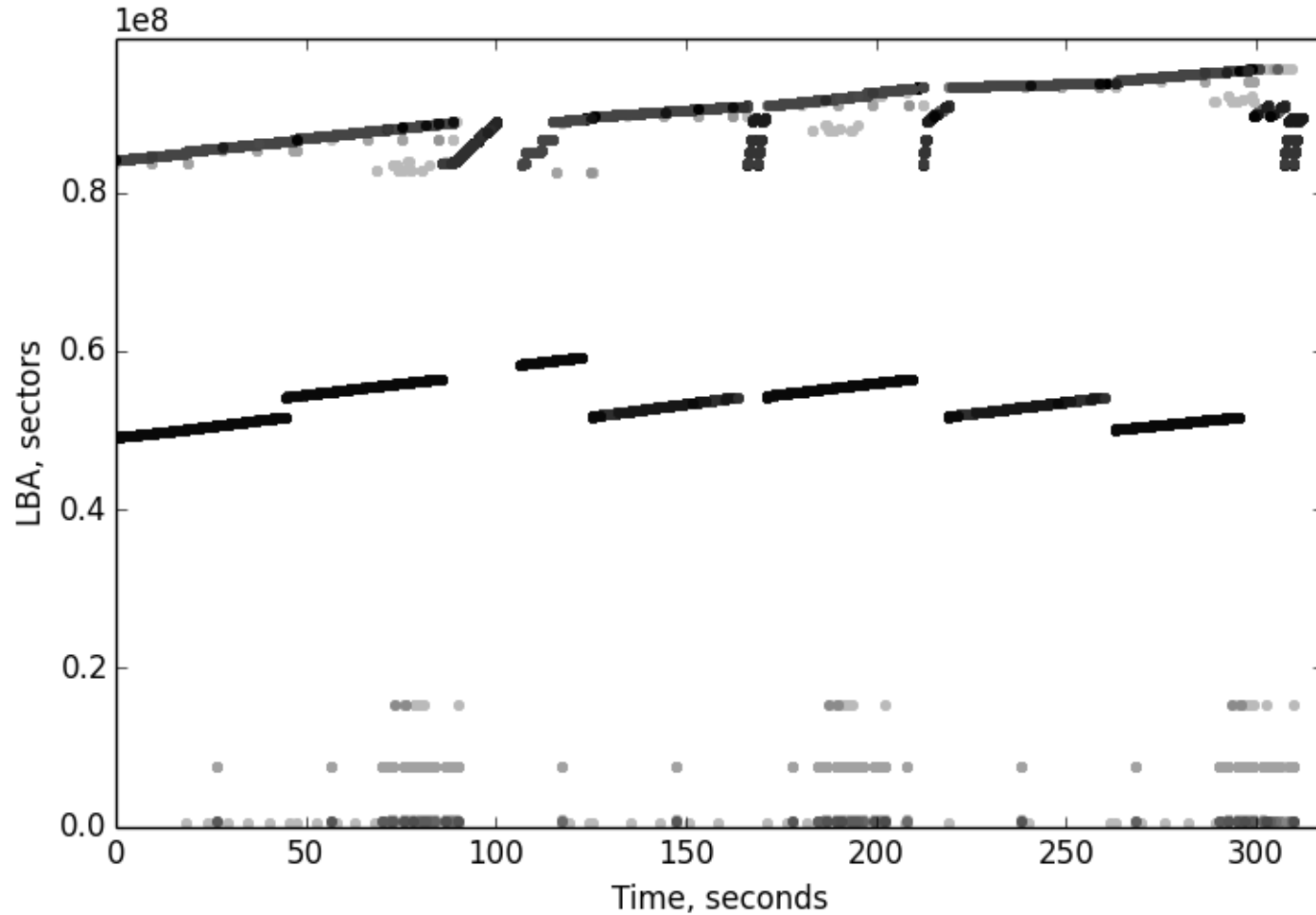
- **Преимущества**

- Детерминированный алгоритм
- Безошибочное определение

- **Недостатки**

- Узкий класс приложений
- Настоящие приложения не похожи на бенчмарки

Машинное обучение



Машинное обучение

- **Поиск последовательностей в запросах**
- **Разбиение на интервалы**
- **Сбор статистик**
 - Размер запроса
 - Частота встречаемости
 - Inter-reference distance
 - Stack distance
- **Кластеризация**
- **Неуспех**
 - Большая размерность вектора
 - Сложности с выбором метрики

Результаты

- **Сборщик трейсов**
 - Python
- **Выборочное распознавание**
 - Исследование завершилось успехом
 - Распознавание бенчмарков
 - C++
 - Модуль интегрирован в СХД
- **Машинное обучение**
 - Получена база для дальнейших исследований
 - Уменьшение размерности векторов
 - Неметрические алгоритмы