

Восстановление параметров потока обращений к устройству хранения

Докладчик: Крайчик Г.И.

Научный руководитель: Нестеров В.М.

СПбГУ, Математико-механический факультет

Кафедра системного программирования

22 апреля 2014 года

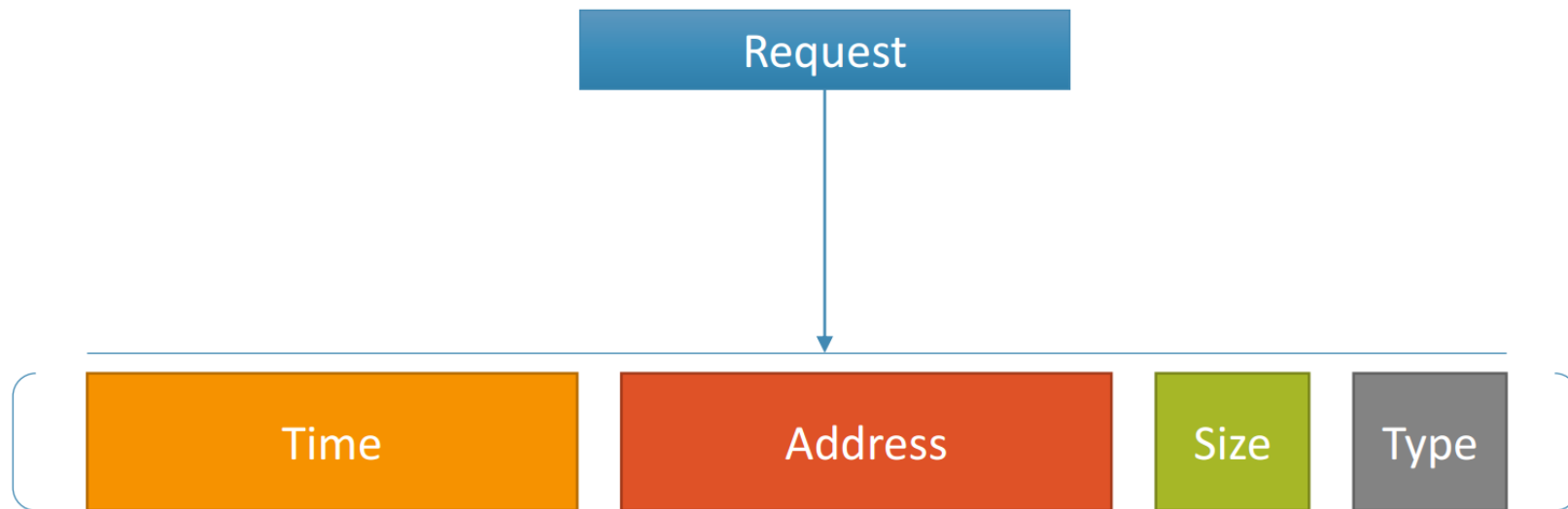


Введение

- Разработка эффективного алгоритма управления кэш-памятью
- Поток обращений зависит от набора параметров
- Целью данной работы является разработка алгоритмов восстановления параметров потока обращений



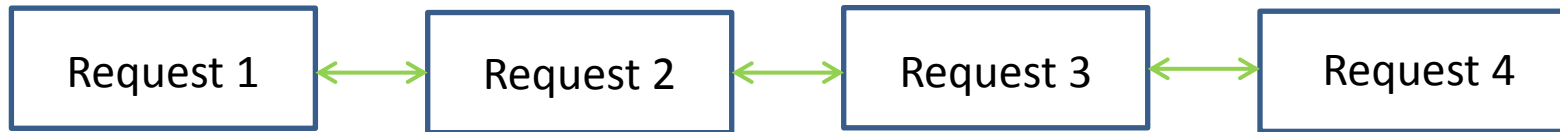
Поток обращений





Анализ времени

- Разность времен между запросами распределена экспоненциально



- Разработано 3 алгоритма
- Проведен их сравнительный анализ



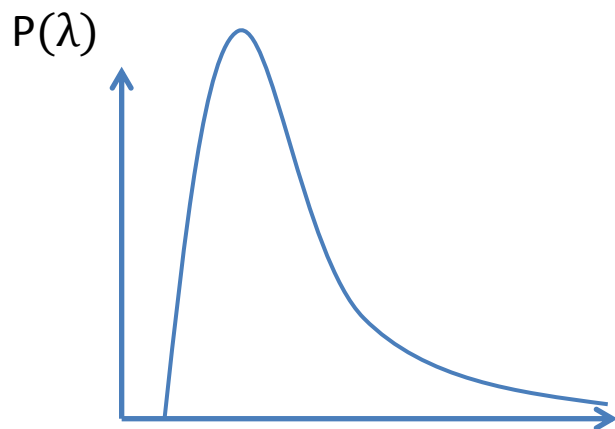
Алгоритм 1

- Уравниваем выборочное среднее последовательности $\{x_i\}_{i=1}^n$ и математическое ожидание $\frac{1}{\lambda}$
- В результате решения уравнения получим выражение $\lambda = \frac{n}{\sum_{i=1}^n x_i}$



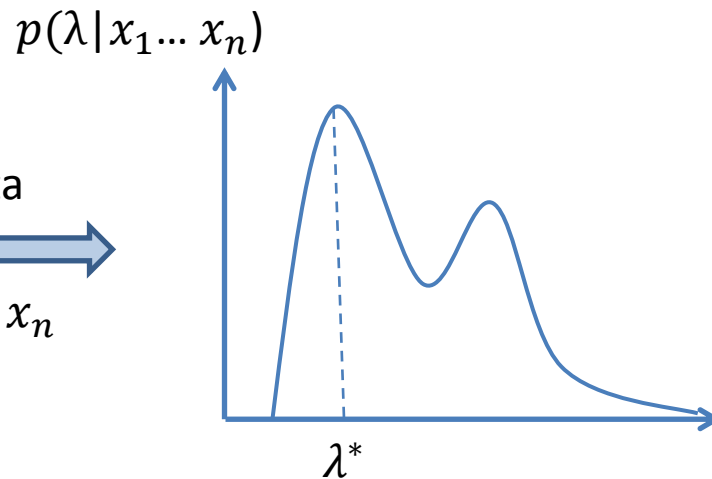
Алгоритм 2

Базовое
распределение



data
 $x_1 \dots x_n$

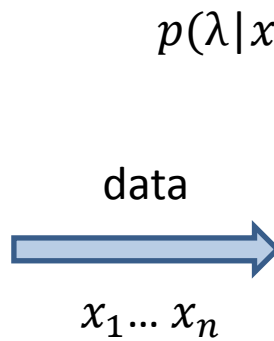
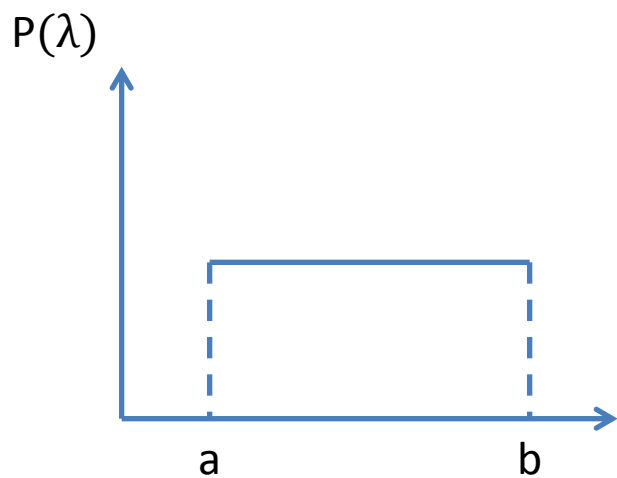
Конечное
распределение



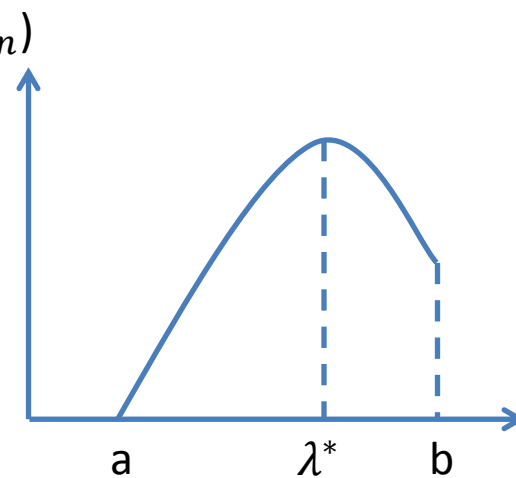


Алгоритм 2

Базовое
распределение



Конечное
распределение





Алгоритм 2

- Пусть λ равномерно распределено на $[a, b]$
- В этом случае можно получить формулу:

$$p(\lambda) = \frac{\frac{1}{b-a} \int_a^\lambda t^n e^{-t \sum_{i=1}^n x_i} dt + \frac{\lambda-a}{b-a} (\lambda^n e^{-\lambda \sum_{i=1}^n x_i})}{\int_a^b t^n e^{-t \sum_{i=1}^n x_i} dt} \chi_{[a,b]}(\lambda)$$

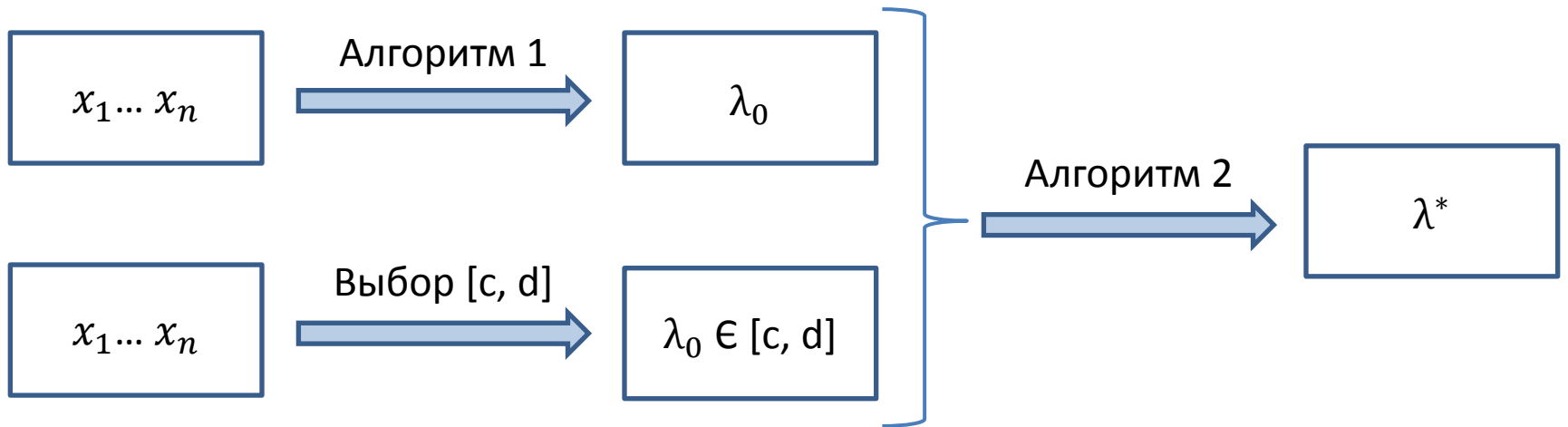
- Достигается единственный максимум в точке:

$$\lambda^* = \min \left\{ \frac{ac + n + 2 + \sqrt{(ac - n + 2)^2 + 8n}}{2c}, b \right\}$$

$$\text{где } c = x_1 + x_2 + \dots + x_n$$



Алгоритм 3





Сравнение алгоритмов

Таблица 1: Матожидание относительной погрешности.

	$a = 0.1$	$a = 0.5$	$a = 0.9$	$a = 0.95$
Алгоритм 1	0.0264	0.0254437	0.0264689	0.025155
Алгоритм 2	0.026965	0.02673426	0.028047	0.023228
Алгоритм 3	0.27016	0.02627	0.02333	0.018956

Таблица 2: Дисперсия относительной погрешности

	$a = 0.1$	$a = 0.5$	$a = 0.9$	$a = 0.95$
Алгоритм 1	$4.06 \cdot 10^{-4}$	$3.69 \cdot 10^{-4}$	$3,63 \cdot 10^{-4}$	$3.64 \cdot 10^{-4}$
Алгоритм 2	$4.22 \cdot 10^{-4}$	$3.94 \cdot 10^{-4}$	$3.88 \cdot 10^{-4}$	$1.90 \cdot 10^{-4}$
Алгоритм 3	$4.42 \cdot 10^{-4}$	$3.96 \cdot 10^{-4}$	$3.068 \cdot 10^{-4}$	$1.99 \cdot 10^{-4}$



Анализ адресов

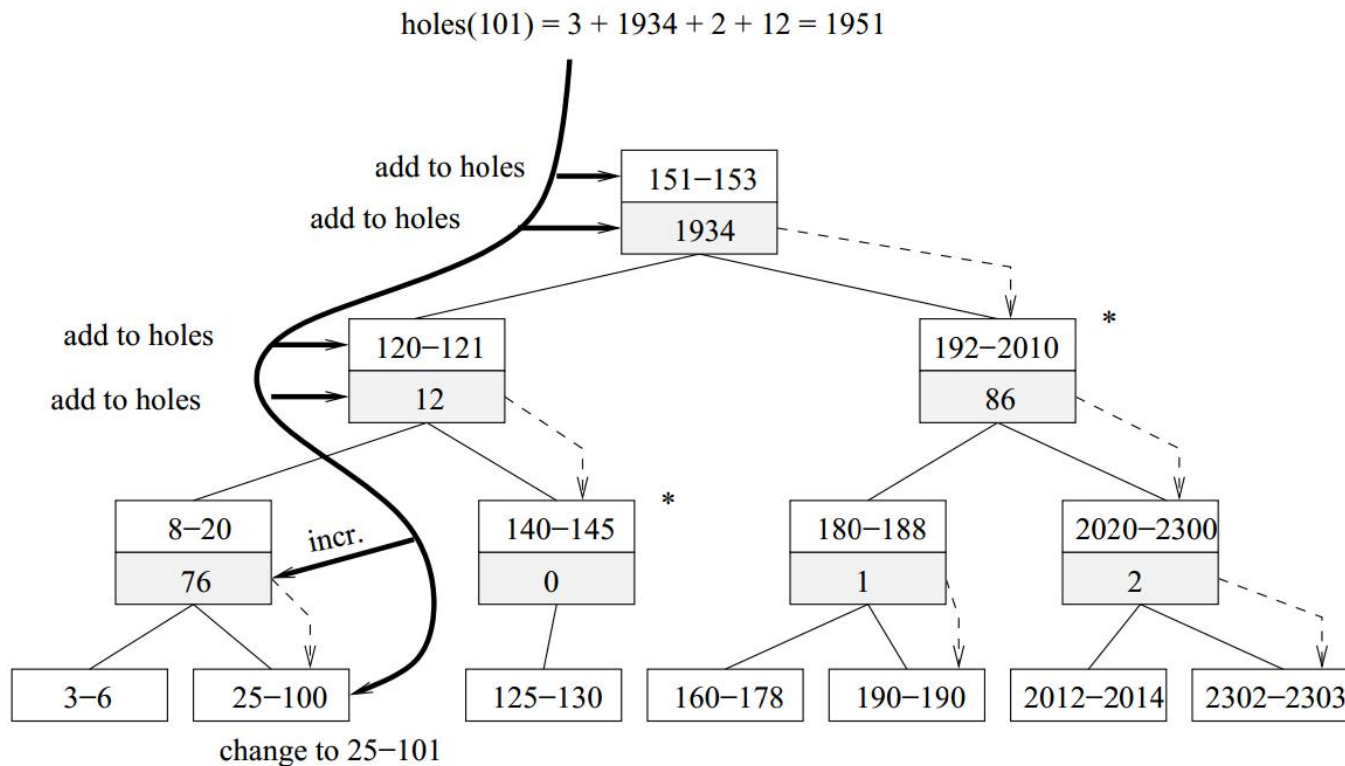


Figure 6: Updating the Tree of Holes



Анализ адресов

- Стековые расстояния распределены по закону Парето
- Применяется метод моментов

$$x_m = \frac{E_1}{E_2} \quad k = \frac{E_2}{E_2 - E_1^2}$$

- E_1, E_2 - первый и второй выборочные моменты соответственно

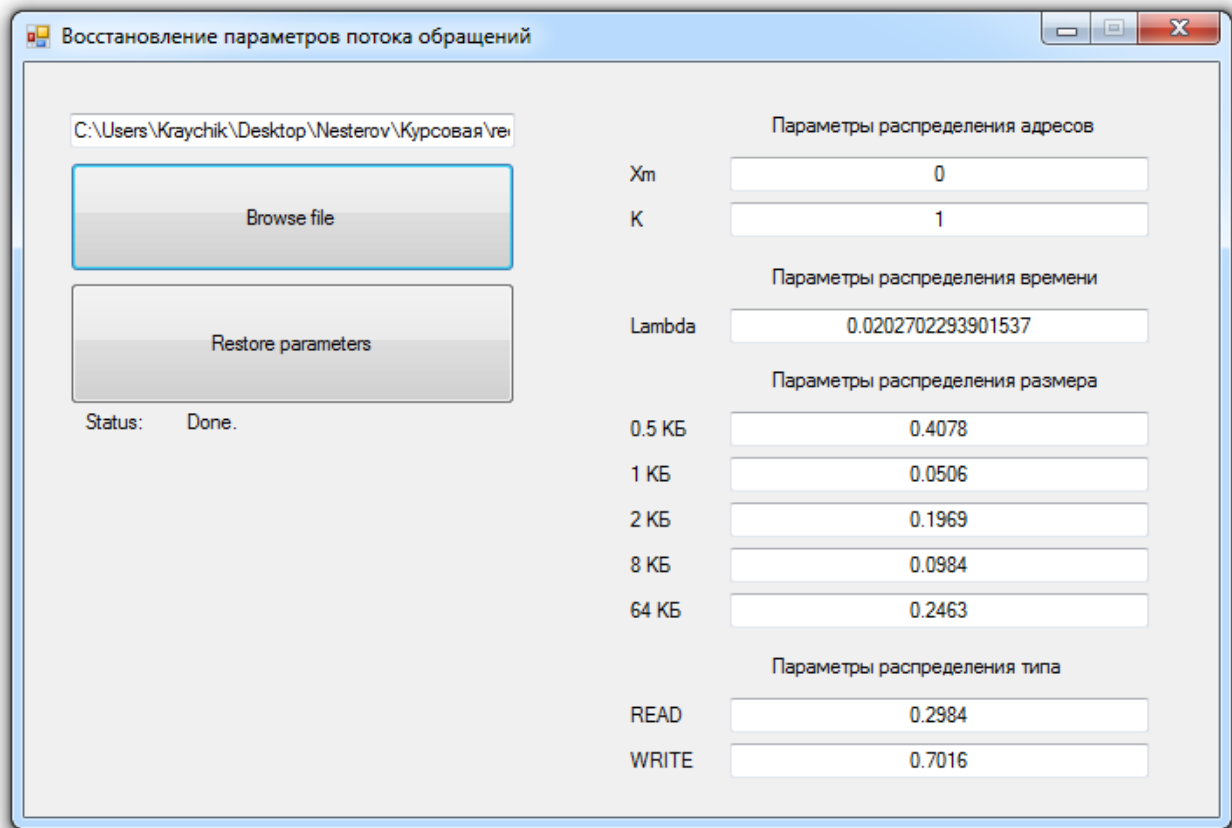


Анализ размера и типа

- Принимается конечное множество значений
- Анализируются только вероятности того, что размер/тип запроса принял соответствующее значение



Разработка приложения



C++

C#



Результаты

- Восстановлены параметры каждой из четырех компонент потока обращений
- Разработан алгоритм, вычисляющий параметры распределения времени более точно, чем метод моментов
- Разработано приложение