

Этап 1: Создание инфраструктуры – созданы классы по заданию.

Этап 2: Реализация HyperLogLog, оценка точности.

$B = 10$, т.к. тогда есть 1024 регистра, что обеспечивает хороший баланс между точностью и памятью.

Формула оценки:

$$E = a_m * m^2 / \sum(2^{-M[i]})$$

Размеры потоков: 10,000, 50,000, 100,000 элементов

Количество экспериментов: 10 независимых потоков для каждого размера

Шаг разбиения: 10%.

Этап 3: Анализ результатов

Средняя ошибка: 2-4%, укладывается в теоретическую границу

Дисперсия алгоритма:

Стандартное отклонение составляет 3-4% от истинного значения

Анализ параметров:

$B = 10$ ($m = 1024$) – оптимальное значение с точностью и скоростью.