



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Э. БАУМАНА
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)
(МГТУ им. Н.Э. БАУМАНА)

ФАКУЛЬТЕТ _____ «Информатика и системы управления»

КАФЕДРА _____ «Программное обеспечение ЭВМ и информационные технологии»

НАПРАВЛЕНИЕ ПОДГОТОВКИ _____ «09.03.04 Программная инженерия»

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №3

Название: _____ Трудоемкость сортировок

Дисциплина: _____ Анализ алгоритмов

Студент	<u>ИУ7-56Б</u>	_____	<u>Ковель А.Д.</u>
	Группа	Подпись, дата	И. О. Фамилия
Преподаватель		_____	<u>Волкова Л.Л.</u>
Преподаватель		_____	<u>Строганов Ю.В.</u>
		Подпись, дата	И. О. Фамилия

Москва, 2022 г.

Оглавление

Введение	3
1 Аналитическая часть	5
1.1 Сортировка бинарным деревом	5
1.2 Сортировка подсчетом	5
1.3 Сортировка поразрядная	6
2 Конструкторский раздел	7
2.1 Трудоемкость алгоритмов	7
2.2 Трудоемкость алгоритмов	8
2.2.1 Алгоритм сортировки бинарным деревом	8
2.2.2 Алгоритм сортировки подсчетом	8
2.2.3 Алгоритм поразрядной сортировки	8
2.3 Разработка алгоритмов	9
3 Технологический раздел	12
3.1 Требования к ПО	12
3.2 Средства реализации	12
3.3 Средства замера времени	12
3.4 Реализации алгоритма	13
3.5 Тестовые данные	15
4 Исследовательская часть	16
4.1 Технические характеристики	16
4.2 Демонстрация работы программы	16
4.3 Время выполнения алгоритмов	17
4.4 Графики функций	18
Заключение	21
Литература	22

Введение

Одной из важнейших процедур обработки структурированной информации является сортировка.

Сортировка - это процесс перегруппировки заданной последовательности (кортежа) объектов в некотором определенном порядке. Такой определенный порядок позволяет, в некоторых случаях, эффективнее и удобнее работать с заданной последовательностью. В частности, одной из целей сортировки является облегчение задачи поиска элемента в отсортированном множестве.

Алгоритмы сортировки используются практически в любой программной системе. Целью алгоритмов сортировки является упорядочение последовательности элементов данных. Поиск элемента в последовательности отсортированных данных занимает время, пропорциональное логарифму количеству элементов в последовательности, а поиск элемента в последовательности не отсортированных данных занимает время, пропорциональное количеству элементов в последовательности, то есть намного больше. Существует множество различных методов сортировки данных. Однако любой алгоритм сортировки можно разбить на три основные части:

- сравнение, определяющее упорядочность пары элементов;
- перестановка, меняющая местами пару элементов;
- собственно сортирующий алгоритм, который осуществляет сравнение и перестановку элементов данных до тех пор, пока все эти элементы не будут упорядочены.

Одной из важнейшей характеристик любого алгоритма сортировки является скорость его работы, которая определяется функциональной зависимостью среднего времени сортировки последовательностей элементов данных, определенной длины, от этой длины.

Цель лабораторной работы — реализация и исследование сортировок: бинарным деревом, поразрядной, подсчетом.

Задачи данной лабораторной:

- изучить и реализовать три алгоритма сортировки: бинарным деревом, поразрядной, подсчетом;
- провести сравнительный анализ трудоемкости алгоритмов на основе теоретических расчетов и выбранной модели вычислений;
- провести сравнительный анализ алгоритмов на основе экспериментальных данных, а именно по времени;
- описать и обосновать полученные результаты в отчете о выполненной лабораторной работе, выполненного как расчётно-пояснительная записка к работе.

1 Аналитическая часть

В этом разделе будут представлены описания алгоритмов сортировки бинарным деревом, подсчетом и поразрядная.

1.1 Сортировка бинарным деревом

Сортировка бинарным деревом [1] — деревом назовем упорядоченную структуру данных, в которой каждому элементу — предшественнику или корню (под)дерева — поставлены в соответствие по крайней мере два других элемента (преемника). Причем для каждого предшественника выполнено следующее правило: левый преемник всегда меньше, а правый преемник всегда больше или равен предшественнику. Вместо 'предшественник' и 'преемник' также употребляют термины 'родитель' и 'сын'. Все элементы дерева также называют 'узлами'.

При добавлении в дерево нового элемента его последовательно сравнивают с нижестоящими узлами, таким образом вставляя на место. Если элемент \geq корня — он идет в правое поддерево, сравниваем его уже с правым сыном, иначе — он идет в левое поддерево, сравниваем с левым, и так далее, пока есть сыновья, с которыми можно сравнить.

1.2 Сортировка подсчетом

Сортировка подсчетом [2] — Сортировка подсчетом — это алгоритм сортировки на основе целых чисел для сортировки массива, ключи которого лежат в определенном диапазоне. Он подсчитывает общее количество элементов с каждым уникальным значением ключа, а затем использует эти подсчеты для определения позиций каждого значения ключа в выходных данных.

1.3 Сортировка поразрядная

Сортировка поразрядная [3]. Массив несколько раз перебирается и элементы перегруппировываются в зависимости от того, какая цифра находится в определённом разряде. После обработки разрядов (всех или почти всех) массив оказывается упорядоченным. При этом разряды могут обрабатываться в противоположных направлениях - от младших к старшим или наоборот.

Вывод

В данной работе стоит задача реализации 3 алгоритмов сортировки, а именно: бинарным деревом, подсчетом и поразрядная. Необходимо оценить теоретическую оценку алгоритмов и проверить ее экспериментально.

2 Конструкторский раздел

В данном разделе представлены схемы реализуемых алгоритмов и их модификации.

2.1 Трудоемкость алгоритмов

Для получения функции трудоемкости алгоритма необходимо ввести модель оценки трудоемкости. Трудоемкость "элементарных" операций оценивается следующим образом:

1. Трудоемкость 1 имеют операции:

$+, -, =, <, >, <=, >=, ==, + =, - =,$
 $++, --, [], \&\&, ||, >>, <<$

2. Трудоемкость 2 имеют операции:

$*, /, \backslash, \%$

3. Трудоемкость конструкции ветвления определяется согласно формуле 2.1.

$$f_{if} = f_{\text{условие}} + \begin{cases} \min(f_{true}, f_{false}) & \text{в лучшем случае,} \\ \max(f_{true}, f_{false}) & \text{в худшем случае.} \end{cases} \quad (2.1)$$

4. Трудоемкость цикла рассчитывается по формуле 2.2.

$$f_{\text{цикл}} = f_{\text{иниц.}} + f_{\text{сравнение}} + N(f_{\text{тело}} + f_{\text{инкремент}} + f_{\text{сравнение}}); \quad (2.2)$$

5. Трудоемкость вызова функции равна 0.

2.2 Трудоемкость алгоритмов

2.2.1 Алгоритм сортировки бинарным деревом

Трудоёмкость данного алгоритма посчитаем следующим образом: сортировка – преобразование массива в бинарное дерево поиска посредством операции вставки нового элемента в бинарное дерево. Операция вставки в бинарное дерево имеет сложность $\log_2(size)$, где $size$ - количество элементов в дереве. Для преобразования массива или списка размером $size$ требуется использовать операцию вставки в бинарное дерево $size$ раз, таким образом, итоговая трудоёмкость данной сортировки будет равна (2.3):

$$f_{radix} = size \cdot \log_2(size); \quad (2.3)$$

2.2.2 Алгоритм сортировки подсчетом

Трудоёмкость алгоритма сортировки подсчётом, где $size$ — количество элементов в массиве (2.4):

$$\begin{aligned} f_{count} &= size + 10 + 2 + size \cdot 8 + 2 + 10 \cdot 6 + 3 \\ &\quad + size \cdot 14 + 1 + size \cdot 5 \\ &= 78 + 28 \cdot size = O(n); \end{aligned} \quad (2.4)$$

2.2.3 Алгоритм поразрядной сортировки

Трудоёмкость алгоритма поразрядной сортировки равна

$$f_{radix} = 1 + 2 + 5 \cdot size + 1 + 2 + m * (f_{count} + 1 + 2); \quad (2.5)$$

где m - количество разрядов в максимальном элементе, f_{count} - трудоёмкость алгоритма сортировки подсчётом.

Итоговая трудоёмкость поразрядной сортировки, использующей сортировку подсчётом в рамках одного разряда:

$$f_{radix} = 28 \cdot m \cdot size = O(n); \quad (2.6)$$

2.3 Разработка алгоритмов

На рисунке 2.3 приведена схема алгоритма сортировки бинарным деревом. На рисунке 2.2 приведена схема алгоритма сортировки подсчетом. Рисунок 2.1 демонстрирует схему алгоритма поразрядной сортировки.

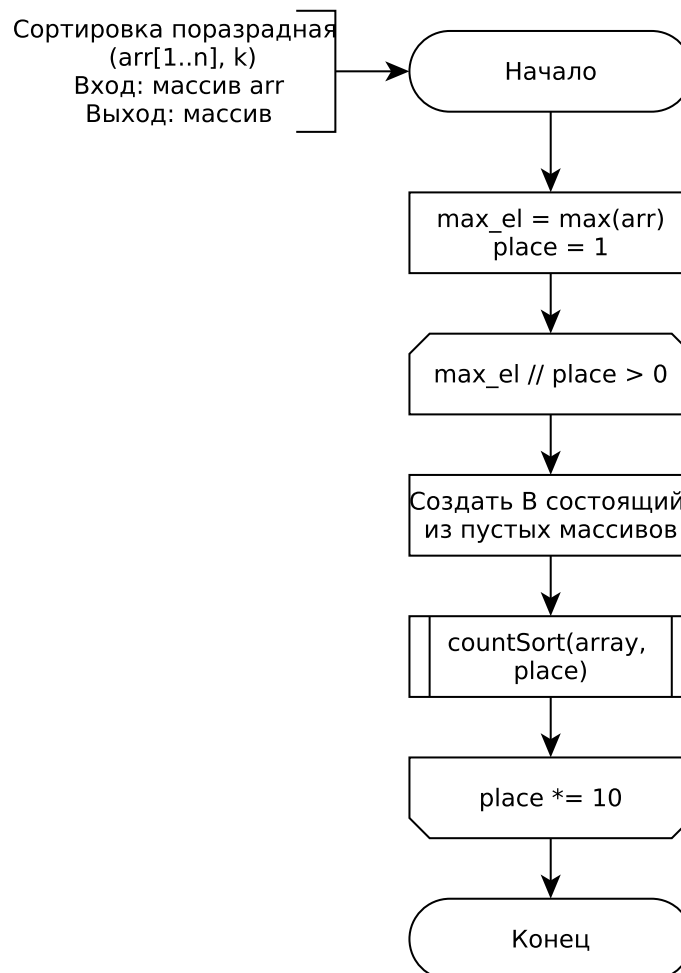


Рисунок 2.1 – Схема алгоритма поразрядной сортировки

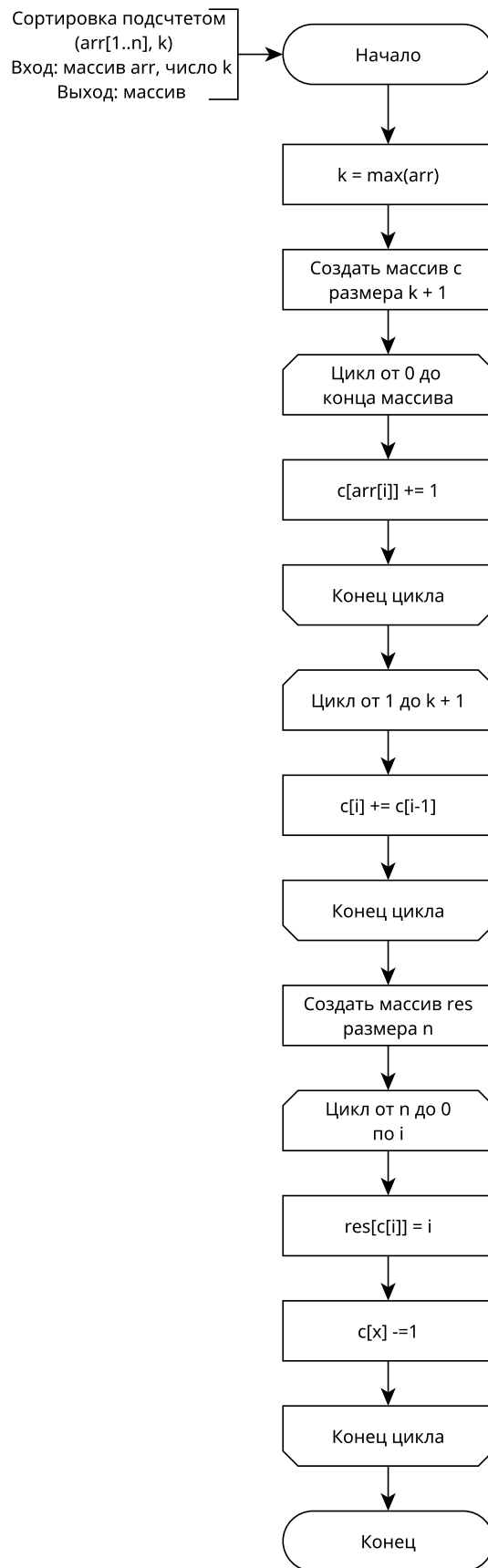


Рисунок 2.2 – Схема алгоритма сортировки подсчетом

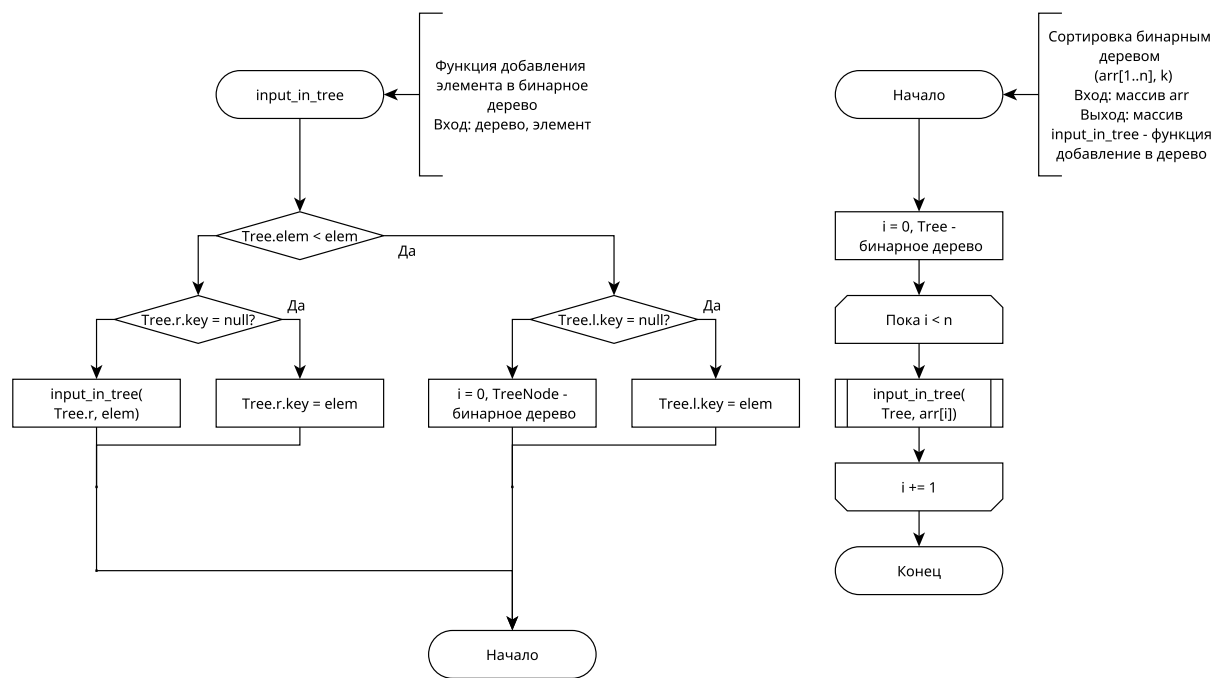


Рисунок 2.3 – Схема алгоритма сортировки бинарным деревом

Вывод

Были разработаны схемы всех трех алгоритмов сортировки. Также для каждого из них были рассчитаны трудоёмкости по введённой модели вычислений с учётом лучших и худших случаев.

3 Технологический раздел

В данном разделе будут приведены требования к программному обеспечению, средства реализации и листинга кода.

3.1 Требования к ПО

Программное обеспечение должно удовлетворять следующим требованиям:

- на вход подается массив целых чисел в диапазоне от 0 до 10000;
- возвращается отсортированный по возрастанию массив;
- в программе возможно измерение процессорного времени.

3.2 Средства реализации

Для реализации ПО был выбран язык программирования Python[4].

В данном языке есть все требующиеся инструменты для данной лабораторной работы.

В качестве среды разработки была выбрана среда VS Code[5], запуск происходил через команду `python main.py`.

3.3 Средства замера времени

Алгоритмы тестировались при помощи функции `process_time` библиотеки `time` 3.1. Данная команда возвращает значения процессорного времени типа `int` в наносекундах.

Замеры времени для каждого алгоритма проводились 100 раз.

Листинг 3.1 – Пример теста эффективности

```
1 def test_simple_mult(A, B):
2     # Start the stopwatch / counter
3     t1_start = process_time()
4     for i in range(N_TEST):
5         simple_mult(A, M, B, N, M)
6     # Stop the stopwatch / counter
7     t1_stop = process_time()
```

3.4 Реализации алгоритма

На листинг 3.2 демонстрируется алгоритм сортировки подсчетом.

Листинг 3.2 – Алгоритм сортировки подсчетом

```
1 def counting_sort(alist, largest):
2     c = [0]*(largest + 1)
3     for i in range(len(alist)):
4         c[alist[i]] += 1
5     c[0] -= 1
6     for i in range(1, largest + 1):
7         c[i] += c[i - 1]
8     result = [0]*len(alist)
9     for x in reversed(alist):
10        result[c[x]] = x
11        c[x] -= 1
12    return result
```

На листинге 3.3 демонстрируется алгоритм сортировки бинарным деревом.

Листинг 3.3 – Алгоритм сортировки бинарным деревом

```
1 def binary_sort(alist):
2     tree = TreeNode()
3     for i in alist:
4         tree.input_in_tree(i)
5     arr = []
6     tree.pre_order(arr)
7
8     return arr
```

На листинге 3.4 демонстрируется класс бинарного дерева.

Листинг 3.4 – Класс бинарного дерева

```
1  class TreeNode:
2  def __init__(self, value=None):
3      self.value = value
4      self.left = None
5      self.right = None
6
7  def input_in_tree(self, elem):
8      if self.value is None:
9          self.value = elem
10         return
11
12     if elem < self.value:
13         if self.left is None:
14             self.left = TreeNode(elem)
15         else:
16             self.left.input_in_tree(elem)
17
18     elif elem >= self.value:
19         if self.right is None:
20             self.right = TreeNode(elem)
21         else:
22             self.right.input_in_tree(elem)
23
24     def pre_order(self, arr):
25         if self.value:
26             if self.left:
27                 self.left.pre_order(arr)
28             if self.value:
29                 arr.append(self.value)
30
31         if self.right:
32             self.right.pre_order(arr)
```

На листинге 3.5 демонстрируется алгоритм поразрядной сортировки.

Листинг 3.5 – Алгоритм поразрядной сортировки

```
1  def radixSort(array):
2      max_element = max(array)
3
4      place = 1
5      while max_element // place > 0:
6          countingSort(array, place)
7          place *= 10
8      return array
```

3.5 Тестовые данные

В таблице 3.1 приведены тесты для функций, реализующих алгоритмы сортировки. Тесты для всех сортировок пройдены *успешно*.

Таблица 3.1 – Функциональные тесты

Входной массив	Ожидаемый результат	Результат
[1, 2, 3, 4, 5, 6, 7]	[1, 2, 3, 4, 5, 6, 7]	[1, 2, 3, 4, 5, 6, 7]
[7, 6, 5, 4, 3, 2, 1]	[1, 2, 3, 4, 5, 6, 7]	[1, 2, 3, 4, 5, 6, 7]
[9, 7, 5, 1, 4]	[1, 4, 5, 7, 9]	[1, 4, 5, 7, 9]
[69]	[69]	[69]
[]	[]	[]

Вывод

Было написано и протестировано программное обеспечение для решения поставленной задачи.

4 Исследовательская часть

4.1 Технические характеристики

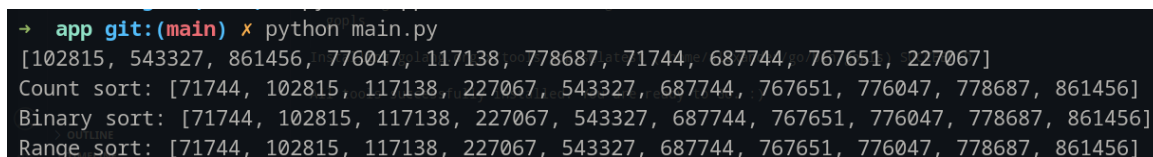
Тестирование выполнялось на устройстве со следующими техническими характеристиками:

- Операционная система Pop!_OS 22.04 LTS [6] Linux [7];
- Оперативная память 16 Гб;
- Процессор AMD® Ryzen 7 2700 eight-core processor × 16 [8].

Во время тестирования устройство было подключено к блоку питания и не нагружено никакими приложениями, кроме встроенных приложений окружения, окружением и системой тестирования.

4.2 Демонстрация работы программы

На рисунке 4.1 представлен результат работы программы.



```
→ app git:(main) x python main.py
[102815, 543327, 861456, 776047, 117138, 778687, 71744, 687744, 767651, 227067]
Count sort: [71744, 102815, 117138, 227067, 543327, 687744, 767651, 776047, 778687, 861456]
Binary sort: [71744, 102815, 117138, 227067, 543327, 687744, 767651, 776047, 778687, 861456]
Range sort: [71744, 102815, 117138, 227067, 543327, 687744, 767651, 776047, 778687, 861456]
```

Рисунок 4.1 – Пример работы программы

4.3 Время выполнения алгоритмов

Результаты замеров времени работы реализаций алгоритмов сортировки на различных входных данных (в мс) приведены в таблицах 4.1, 4.2 и 4.3.

Таблица 4.1 – Результаты замеров реализаций сортировок, входными данными являлись отсортированные по возрастанию значений массивы.

Размер	Подсчетом	Поразрядная	Бинарным деревом
100	0.1662	0.0714	0.8650
200	0.5113	0.2058	3.3357
300	1.1026	0.3131	7.6464
400	2.0140	0.4364	13.6841
500	3.3046	0.5591	21.5524
600	5.0567	0.6798	31.3052
700	6.6944	0.7852	43.0406
800	8.5163	0.8766	56.4318

Таблица 4.2 – Результаты замеров реализаций сортировок, входными данными являлись отсортированные по убыванию значений массивы.

Размер	Подсчетом	Поразрядная	Бинарным деревом
100	0.1606	0.1048	0.7138
200	0.5005	0.2008	2.7633
300	1.0747	0.3110	6.3060
400	1.9383	0.4312	11.3831
500	3.1148	0.5427	18.0577
600	4.6409	0.6693	26.0260
700	6.7969	0.8317	36.7397
800	8.7922	0.9583	47.2628

Таблица 4.3 – Результаты замеров реализаций сортировок, входными данными являлись заполненные числами со случайными значениями массивы.

Размер	Подсчетом	Поразрядная	Бинарным деревом
100	0.2734	0.1043	0.1560
200	0.8321	0.2090	0.3756
300	1.6837	0.3142	0.6025
400	2.8938	0.4281	0.9785
500	4.4438	0.5419	1.1784
600	6.4153	0.6704	1.5523
700	8.6692	0.7678	1.9018
800	11.3752	0.8992	2.2986

4.4 Графики функций

На графике 4.3 представлено время работы сортировок, входными данными являлись заполненные числами со случайными значениями массивы.

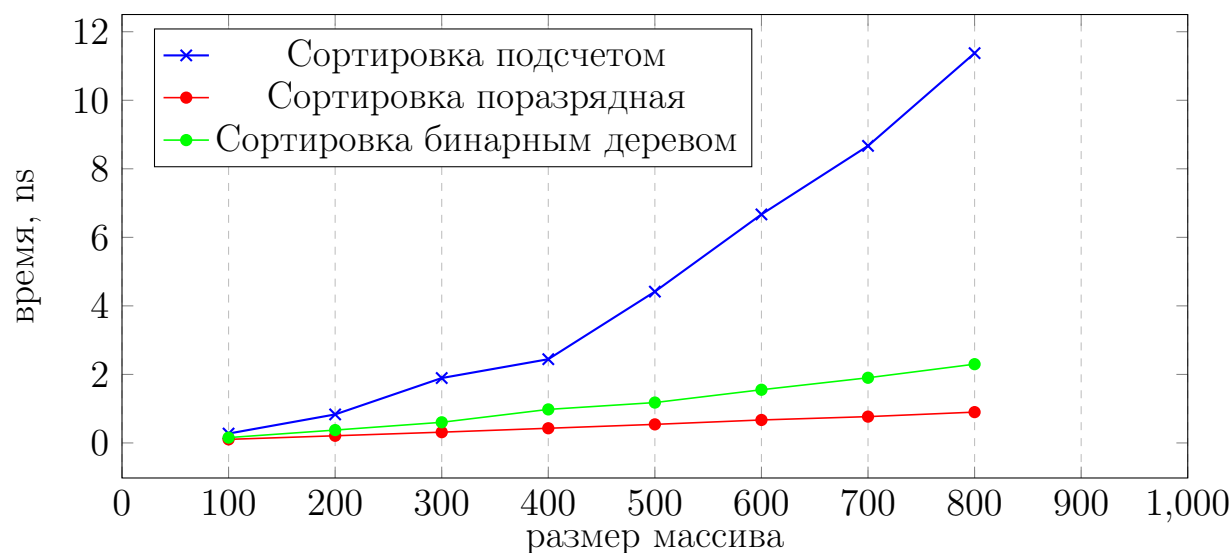


Рисунок 4.2 – Результаты замеров реализаций сортировок, входными данными являлись заполненные числами со случайными значениями массивы.

На графике 4.3 представлено время работы сортировок, входными дан-

ными являлись отсортированные по возрастанию значений массивы.

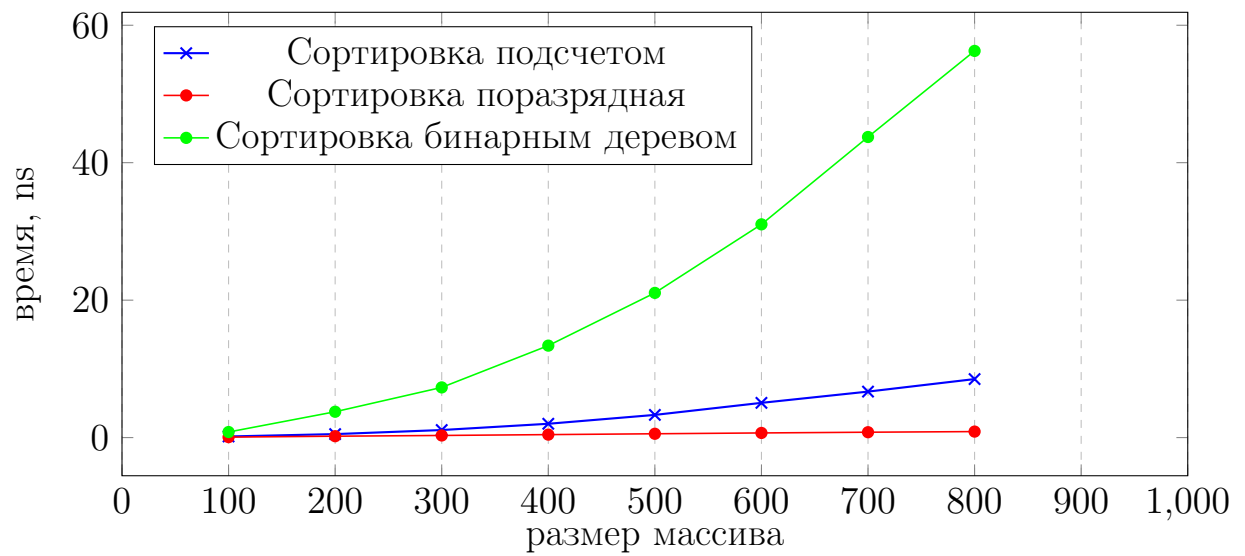


Рисунок 4.3 – Результаты замеров реализаций сортировок, входными данными являлись отсортированные по возрастанию значений массивы.

На графике 4.3 представлено время работы сортировок, входными данными являлись отсортированные по убыванию значений массивы.

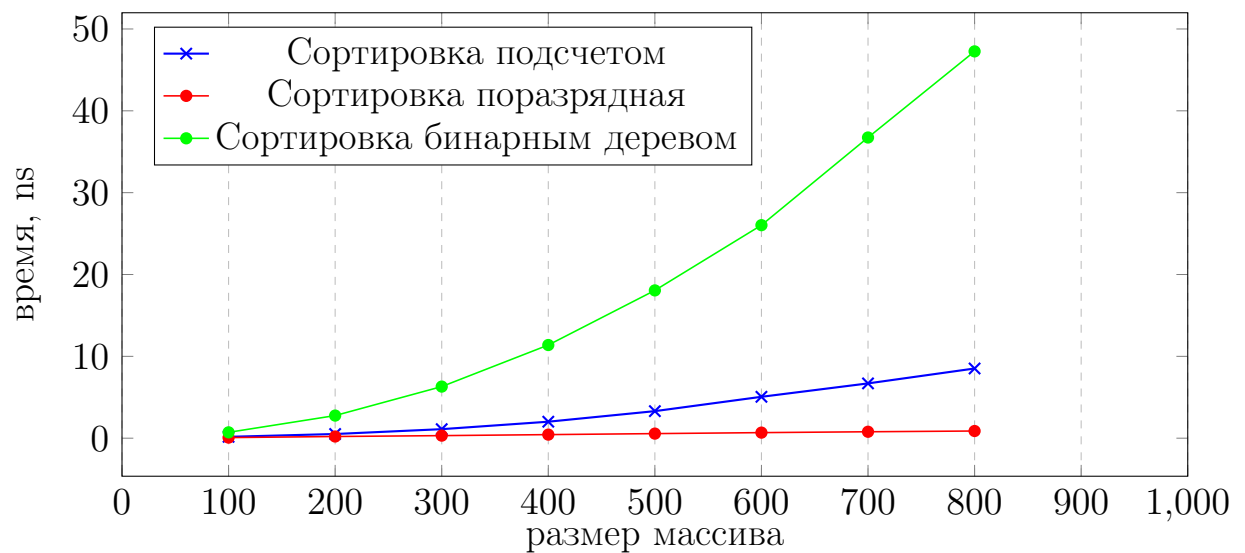


Рисунок 4.4 – Результаты замеров реализаций сортировок, входными данными являлись отсортированные по убыванию значений массивы.

Вывод

В данном разделе были сравнены алгоритмы по времени. Сортировка бинарным деревом на отсортированных массивах и сортировка подсчетом работает на случайном массиве дольше всех.

Самая быстрая сортировка поразрядная на любых данных.

Теоритические результаты оценки трудоемкости и полученные практическим образом результаты замеров совпадают.

Заключение

В ходе выполнения лабораторной работы были решены следующие задачи:

- были изучены и реализованы 3 алгоритма сортировки массива: бинарным деревом, поразрядная, подсчетом;
- был произведен анализ трудоёмкости алгоритмов на основе теоретических расчетов и выбранной модели вычислений;
- был сделан сравнительный анализ алгоритмов на основе экспериментальных данных;
- подготовлен отчет о лабораторной работе.

Сортировка бинарным деревом на отсортированных массивах и сортировка подсчетом работает на случайном массиве дольше всех. Самая быстрая сортировка поразрядная на любых данных.

Поставленная цель достигнута.

Литература

- [1] Сортировка бинарным деревом [Электронный ресурс]. Режим доступа: <http://algotlist.ru/sort/faq/q7.php> (дата обращения: 24.10.2022).
- [2] Сортировка подсчетом [Электронный ресурс]. Режим доступа: <https://www.techiedelight.com/ru/counting-sort-algorithm-implementation/> (дата обращения: 24.10.2022).
- [3] Сортировка по разрядам [Электронный ресурс]. Режим доступа: <http://algotlab.valemak.com/radix> (дата обращения: 24.10.2022).
- [4] Python Документация[Электронный ресурс]. Режим доступа: <https://docs.python.org/3/> (дата обращения: 24.09.2022).
- [5] Vscode Документация[Электронный ресурс]. Режим доступа: <https://code.visualstudio.com/docs> (дата обращения: 24.09.2022).
- [6] Pop OS 22.04 LTS [Электронный ресурс]. Режим доступа: <https://pop.system76.com> (дата обращения: 04.09.2022).
- [7] Linux – Документация [Электронный ресурс]. Режим доступа: <https://docs.kernel.org> (дата обращения: 24.09.2022).
- [8] Процессор AMD® Ryzen 7 2700 eight-core processor × 16 [Электронный ресурс]. Режим доступа: <https://www.amd.com/en/products/cpu/amd-ryzen-7-2700> (дата обращения: 04.09.2022).