



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» _____

Отчет по лабораторной работе №4 по курсу "Анализ алгоритмов"

Тема Реализация параллельного алгоритма Коперсмита-Винограда

Студент Якуба Д. В.

Группа ИУ7-53Б

Оценка (баллы) _____

Преподаватели Волкова Л.Л., Строганов Ю.В.

Москва — 2020 г.

Оглавление

Введение	3
1 Аналитическая часть	5
1.1 Алгоритм Копперсмита-Винограда умножения матриц . . .	5
2 Конструкторская часть	7
2.1 Блок-схема алгоритма Копперсмита-Винограда	7
2.2 Блок-схема алгоритма работы функции, запускающая в требуемом количестве потоков функцию-аргумент	11
2.3 Параллельная реализация алгоритма Копперсмита-Винограда	12
2.4 Схема параллельной реализации алгоритма Копперсмита- Винограда с разделением этапа вычисления элементов мат- рицы	18
2.5 Схема параллельной реализации алгоритма Копперсмита- Винограда без разделения этапа вычисления элементов мат- рицы	19
3 Технологическая часть	21
3.1 Требования к программному обеспечению	21
3.2 Средства реализации программного обеспечения	21
3.3 Листинг кода	21
3.4 Тестирование программного продукта	23
4 Исследовательская часть	24
4.1 Пример работы программного обеспечения	24
4.2 Технические характеристики	27
4.3 Время выполнения алгоритмов	27

Заключение	33
Литература	34

Введение

Цели лабораторной работы

1. изучение параллельных вычислений;
2. изучение последовательного и двух параллельных реализаций алгоритмов Винограда;
3. реализация последовательного и двух параллельных алгоритмов Винограда;
4. проведение сравнительного анализа временных характеристик реализованных алгоритмов на основе экспериментальных данных;
5. подготовка отчёта по лабораторной работе;
6. получение практических навыков реализации алгоритмов на ЯП Nim.

Определение

Многопоточность — это способность центрального процессора (CPU) или одного ядра в многоядерном процессоре одновременно выполнять несколько процессов или потоков, соответствующим образом поддерживаемых операционной системой. Этот подход отличается от многопроцессорности, так как многопоточность процессов и потоков совместно использует ресурсы одного или нескольких ядер: вычислительных блоков, кэш-памяти ЦПУ или буфера перевода с преобразованием (TLB).

В тех случаях, когда многопроцессорные системы включают в себя несколько полных блоков обработки, многопоточность направлена на

максимизацию использования ресурсов одного ядра, используя параллелизм на уровне потоков, а также на уровне инструкций. Поскольку эти два метода являются взаимодополняющими, их иногда объединяют в системах с несколькими многопоточными ЦП и в ЦП с несколькими многопоточными ядрами.

Многопоточная парадигма стала более популярной с конца 1990-х годов, поскольку усилия по дальнейшему использованию параллелизма на уровне инструкций застопорились. Смысл многопоточности — квазимоногозадачность на уровне одного исполняемого процесса. Значит, все потоки процесса помимо общего адресного пространства имеют и общие дескрипторы файлов. Выполняющийся процесс имеет как минимум один (главный) поток.

1 | Аналитическая часть

1.1 Алгоритм Копперсмита-Винограда умножения матриц

Пусть даны две прямоугольные матрицы

$$A_{lm} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l1} & a_{l2} & \dots & a_{lm} \end{pmatrix}, \quad B_{mn} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}, \quad (1.1)$$

тогда матрица C

$$C_{ln} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{l1} & c_{l2} & \dots & c_{ln} \end{pmatrix}, \quad (1.2)$$

где

$$c_{ij} = \sum_{r=1}^m a_{ir} b_{rj} \quad (i = \overline{1, l}; j = \overline{1, n}) \quad (1.3)$$

будет называться произведением матриц A и B .

Если посмотреть на результат умножения двух матриц, то видно, что каждый элемент в нем представляет собой скалярное произведение соответствующих строки и столбца исходных матриц. Можно заметить также, что такое умножение допускает предварительную обработку, позволяющую часть работы выполнить заранее.

Рассмотрим два вектора $V = (v_1, v_2, v_3, v_4)$ и $W = (w_1, w_2, w_3, w_4)$. Их скалярное произведение равно: $V \cdot W = v_1w_1 + v_2w_2 + v_3w_3 + v_4w_4$, что эквивалентно (1.4):

$$V \cdot W = (v_1 + w_2)(v_2 + w_1) + (v_3 + w_4)(v_4 + w_3) - v_1v_2 - v_3v_4 - w_1w_2 - w_3w_4. \quad (1.4)$$

Несмотря на то, что второе выражение требует вычисления большего количества операций, чем стандартный алгоритм: вместо четырех умножений - шесть, а вместо трех сложений - десять, выражение в правой части последнего равенства допускает предварительную обработку: его части можно вычислить заранее и запомнить для каждой строки первой матрицы и для каждого столбца второй, что позволит для каждого элемента выполнять лишь два умножения и пять сложений, складывая затем только лишь с 2 предварительно посчитанными суммами соседних элементов текущих строк и столбцов [1]. Из-за того, что операция сложения быстрее операции умножения в ЭВМ, на практике алгоритм должен работать быстрее стандартного [2].

Важно заметить, что каждый элемент матрицы C вычисляется независимо от других, а матрицы A и B не изменяются, поэтому для распараллеливания алгоритма будет достаточно распределить элементы каждой строки матрицы C между потоками.

Вывод

Был рассмотрен алгоритм умножения матриц Коперсмита-Винограда. В данной работе стоит задача реализации данного алгоритма в последовательном и параллельных видах. Независимость вычислений элементов результата даёт возможность реализовать несколько параллельных вариантов исполнения данного алгоритма.

2 | Конструкторская часть

2.1 Блок-схема алгоритма Копперсмита-Винограда

Блок-схема алгоритма Копперсмита-Винограда предоставлена на рисунках 2.1-2.3.

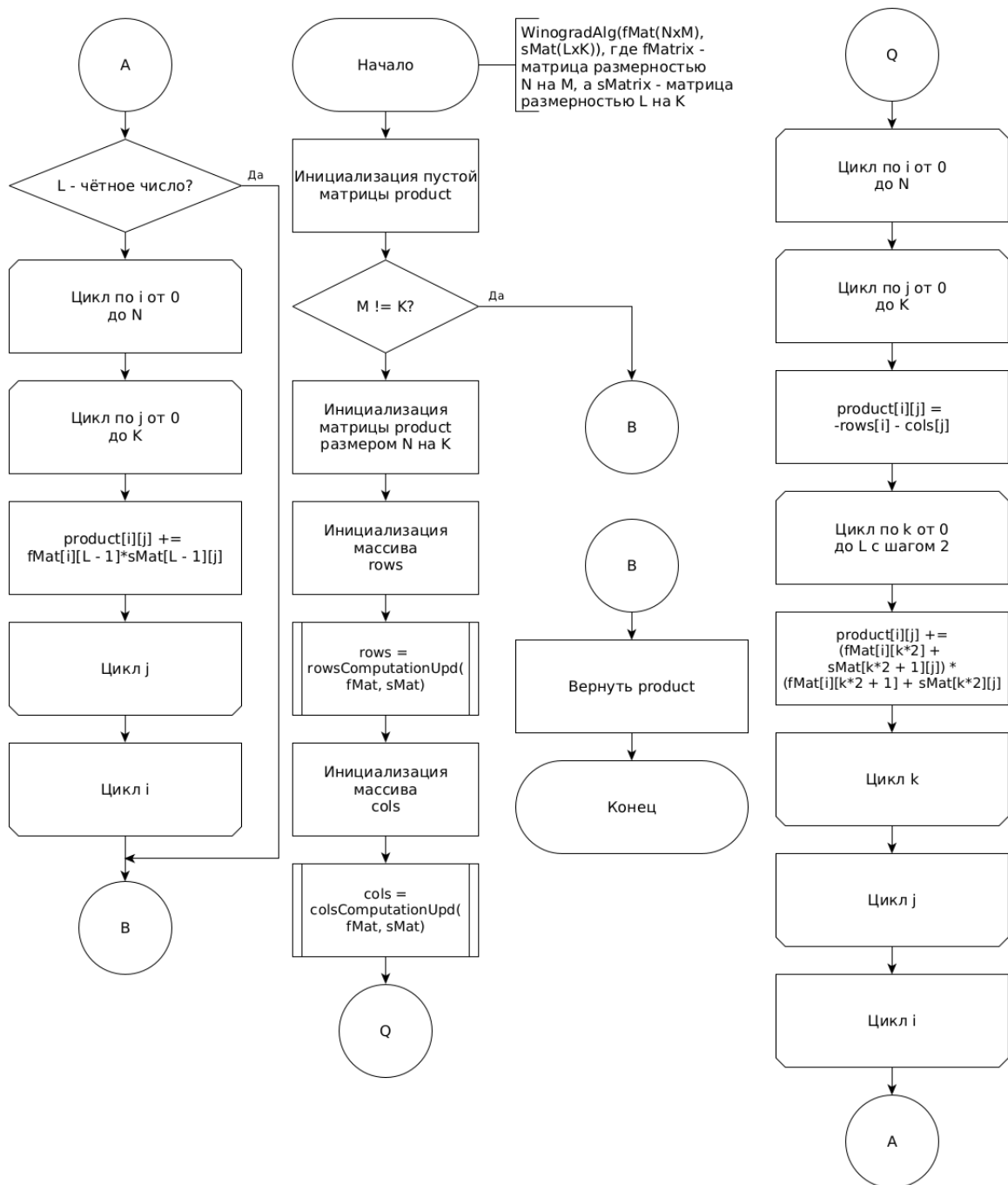


Рис. 2.1: Блок-схема алгоритма Копперсмита-Винограда.

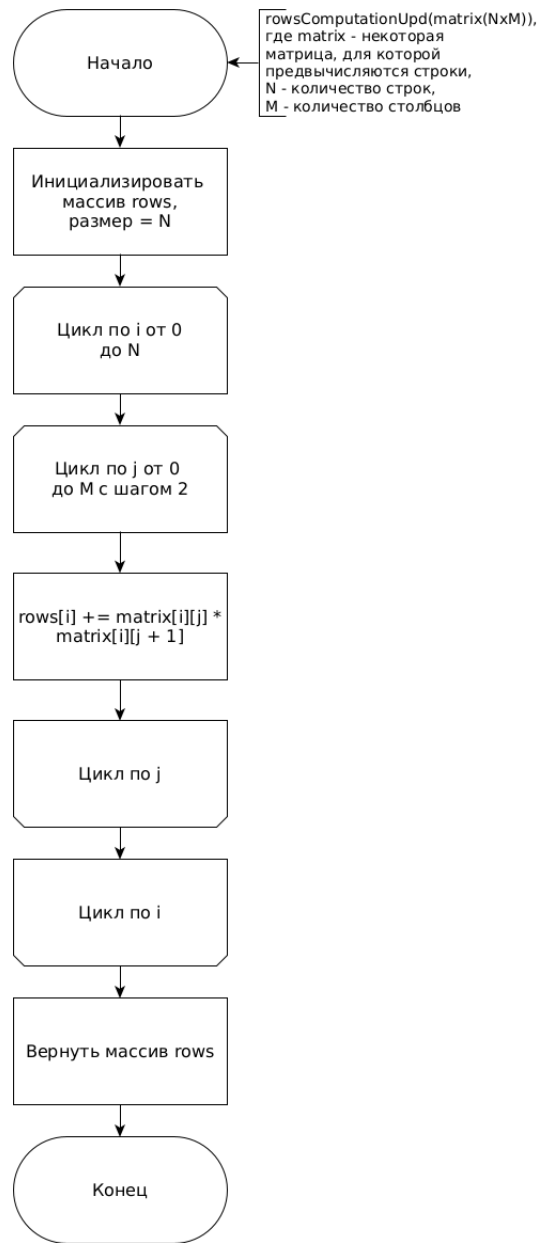


Рис. 2.2: Блок-схема предрасчёта строк для алгоритма Копперсмита-Винограда.

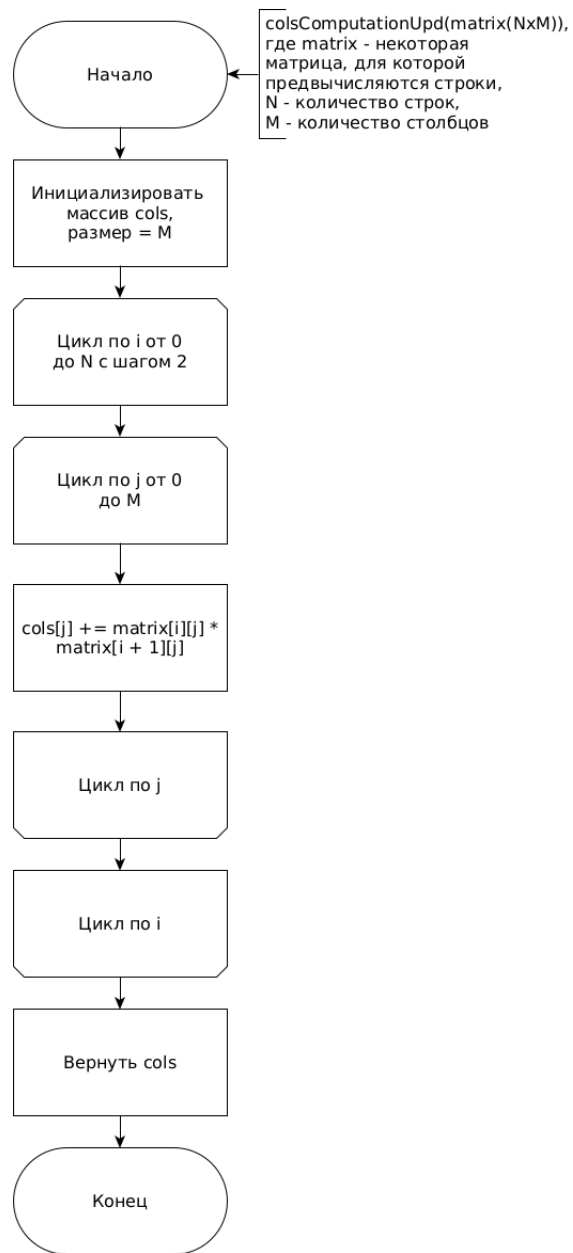


Рис. 2.3: Блок-схема предрасчёта столбцов для алгоритма Кошперсмита-Винограда.

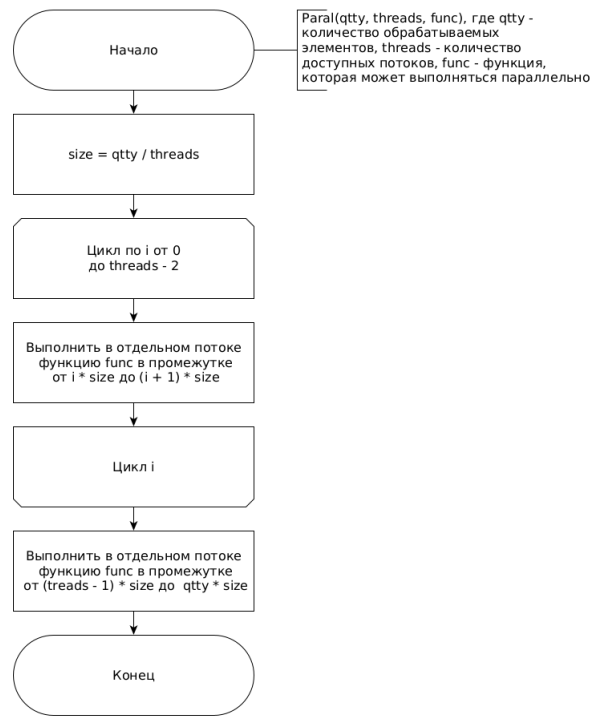


Рис. 2.4: Блок-схема алгоритма работы функции, запускающей в требуемом количестве потоков функцию-аргумент.

2.2 Блок-схема алгоритма работы функции, запускающая в требуемом количестве потоков функцию-аргумент

На рисунке 2.4 предоставлена блок-схема алгоритма работы функции, запускающей в требуемом количестве потоков функцию-аргумент.

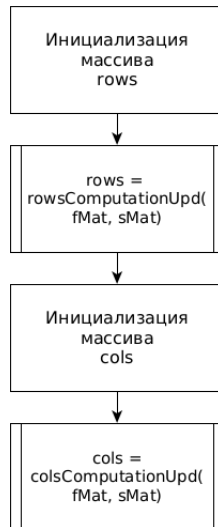


Рис. 2.5: Часть блок-схемы, в которой происходит вызов подпрограмм предрасчёта строк и столбцов.

2.3 Параллельная реализация алгоритма Копперсмита-Винограда

Для распараллеливания алгоритма требуется определить, какие из частей алгоритма могут быть выполнены вне зависимости друг от друга.

Часть блок-схемы, отвечающая за предрасчёт строк и столбцов (рисунок 2.5), однозначно может быть распараллелен. Но до завершения двух этих подпрограмм к выполнению следующих этапов приступить нельзя, так как вычисленные массивы будут использоваться далее. В данной части будет использоваться полное выделенное количество потоков последовательно - первоначально для вычисления строк, а далее - для вычисления столбцов.

Непосредственное вычисление значения каждого элемента матрицы, показанное на рисунке 2.6, может быть также поэлементно распараллелено. Однако тут важно заметить, что часть данного этапа, показанная на рисунке 2.7, и часть, показанная на рисунке 2.8, могут быть выполнены как раздельно, так и совместно при обработке каждого из элементов. Важно заметить, что, при раздельном выполнении, потребуется создавать новые потоки, что является достаточно затратным по времени действием.

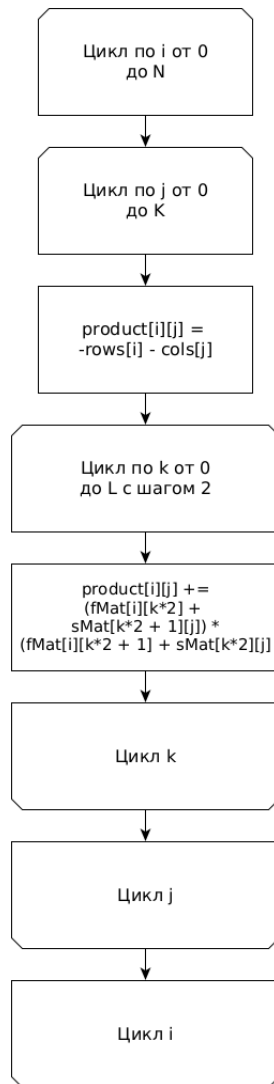


Рис. 2.6: Часть блок-схемы, в которой происходит расчёт значения каждого элемента матрицы.

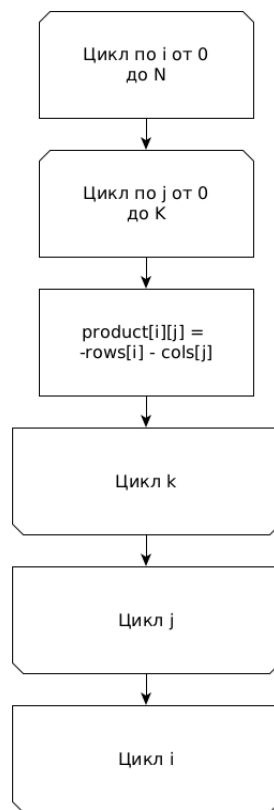


Рис. 2.7: Первая часть рассматриваемого этапа.

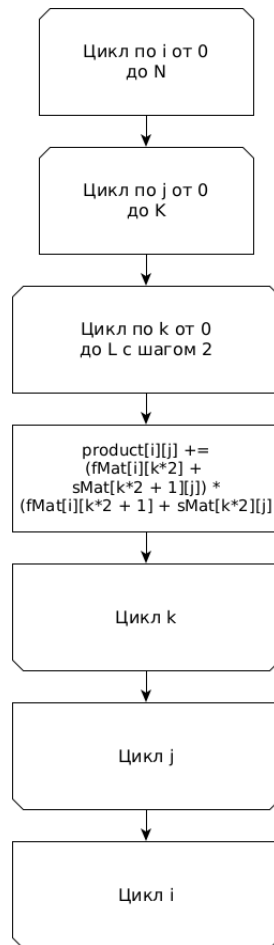


Рис. 2.8: Вторая часть рассматриваемого этапа.

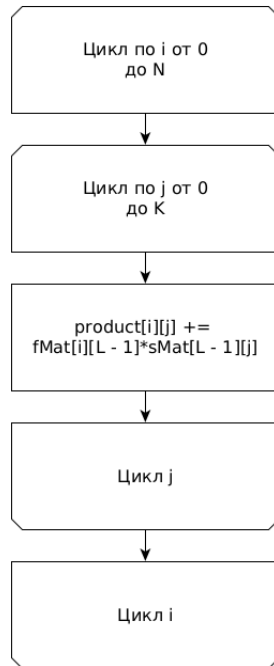


Рис. 2.9: Довычисление значений итоговой матрицы при нечётной размерности оной.

Довычисление значений в случае нечётной размерности итоговой матрицы, показанное на рисунке 2.9 в каждой итерации цикла требует обращения к матрице, что при распараллеливании приведёт к большому числу блокирований разделяемой памяти и будет неэффективно. Поэтому данный этап останется без изменений.

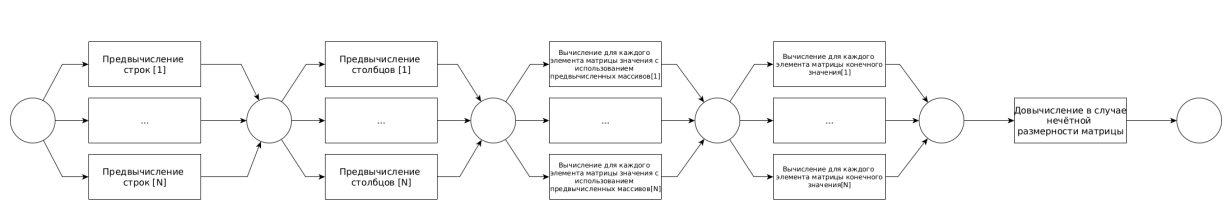


Рис. 2.10: Схема с одновременным выполнением двух частей этапа вычисления элементов матрицы.

2.4 Схема параллельной реализации алгоритма Копперсмита-Винограда с разделением этапа вычисления элементов матрицы

На рисунке 2.10 предоставлена схема с одновременным выполнением двух частей этапа вычисления элементов матрицы.

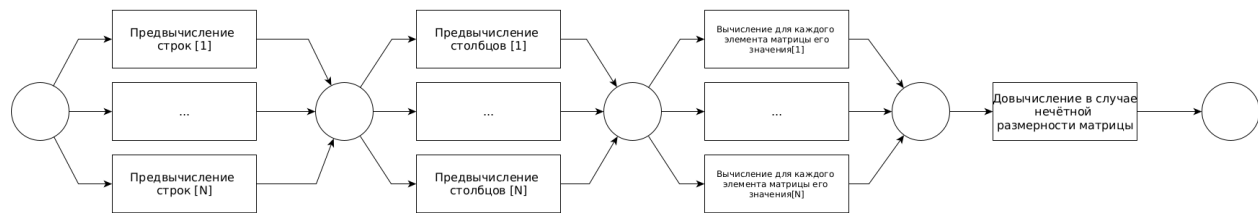


Рис. 2.11: Схема с последовательным выполнением двух частей этапа вычисления элементов матрицы.

2.5 Схема параллельной реализации алгоритма Копперсмита-Винограда без разделения этапа вычисления элементов матрицы

На рисунке 2.11 предоставлена схема с последовательным выполнением двух частей этапа вычисления элементов матрицы.

Вывод

На основе теоретических данных, полученных из аналитического раздела, была построена схема алгоритма Копперсмита-Винограда. Были предложены 2 схемы параллельной реализации рассматриваемого алгоритма.

3 | Технологическая часть

3.1 Требования к программному обеспечению

- входные данные - две матрицы размерностью $M \times N$ и $K \times L$;
- выходные данные - результат умножения двух переданных матриц.

3.2 Средства реализации программного обеспечения

При написании программного продукта был использован язык программирования Nim [3].

Данный выбор обусловлен следующими факторами:

- Компилируемость языка в C, C++, Objective C и JavaScript;
- Синтаксис языка близок к синтаксису ЯП Python;

Для тестирования производительности реализаций алгоритмов использовалась библиотека times.

При написании программного продукта использовалась среда разработки IntelliJ IDEA.

Данный выбор обусловлен тем, что данная среда разработки имеет плагин поддержки языка программирования Nim.

3.3 Листинг кода

В листингах 3.1 - 3.3 предоставлены реализации рассматриваемых алгоритмов.

Листинг 3.1: Функция реализации алгоритма сортировки пузырьком

```
1 fun bubbleSort(arr: IntArray)
2 {
3     var i = 0;
4     while (i < arr.size - 1)
5     {
6         var j = 0;
7         while (j < arr.size - 1 - i)
8         {
9             if (arr[j] > arr[j + 1])
10                 arr[j + 1] = arr[j].also { arr[j] = arr[j + 1] }
11             j++
12         }
13         i++
14     }
15 }
```

Листинг 3.2: Функция реализации алгоритма сортировки вставками

```
1 fun insertionSort(arr: IntArray)
2 {
3     var i = 1;
4     while (i < arr.size)
5     {
6         var min = arr[i]
7         var j = i;
8         while (j > 0 && arr[j - 1] > min)
9         {
10             arr[j] = arr[j - 1]
11             j--
12         }
13         arr[j] = min
14         i++
15     }
16 }
```

Листинг 3.3: Функция реализации алгоритма сортировки выбором

```
1 fun selectionSort(arr: IntArray)
2 {
3     var i = 0;
4     while (i < arr.size - 1)
5     {
6         var j = i + 1
7         var min = i
```

```

8      while (j < arr.size)
9      {
10         if (arr[j] < arr[min])
11             min = j
12         j++
13     }
14     if (min != i)
15         arr[i] = arr[min].also { arr[min] = arr[i] }
16     i++
17 }
18 }

```

3.4 Тестирование программного продукта

В таблице 3.1 приведены тесты для функций, реализующих алгоритм сортировки пузырьком, вставками и выбором. Тесты пройдены успешно.

Начальный массив	Ожидаемый результат	Полученный результат
[5, 4, 3, 2, 1]	[1, 2, 3, 4, 5]	[1, 2, 3, 4, 5]
[0, 0, 0]	[0, 0, 0]	[0, 0, 0]
[1, 2, 3]	[1, 2, 3]	[1, 2, 3]
[-8, 2, 3, -2]	[-8, -2, 2, 3]	[-8, -2, 2, 3]
[]	[]	[]

Таблица 3.1: Тестирование функций

Вывод

Спроектированные алгоритмы сортировок были реализованы и протестированы.

4 | Исследовательская часть

4.1 Пример работы программного обеспечения

Ниже на рисунках 4.1- 4.2 предоставлены примеры работы каждого из алгоритмов на введённых пользователем данных.

```
Hint: /home/trvehazzk3r/Desktop/AlgorithmAnalysis/AA_lab_04/src/main [Exec]
Rows: 3
Cols: 3
1 2 3
1 2 3
1 2 3

Got matrix:
    1    2    3
    1    2    3
    1    2    3

Rows: 3
Cols: 3
1 2 3
4 3 2
1 2 3

Got matrix:
    1    2    3
    4    3    2
    1    2    3

First parallel result:
    12    14    16
    12    14    16
    12    14    16

Second parallel result:
    12    14    16
    12    14    16
    12    14    16

Classic Winograd result:
    12    14    16
    12    14    16
    12    14    16
```

Рис. 4.1: Пример работы ПО.

```
Hint: /home/trvehazzk3r/Desktop/AlgorithmAnalysis/AA_lab_04/src/main [Exec]
Rows: 3
Cols: 3
1 2 3
2 2 2
3 3 3

Got matrix:
  1    2    3
  2    2    2
  3    3    3

Rows: 3
Cols: 3
1 2 3
3 2 1
9 6 6

Got matrix:
  1    2    3
  3    2    1
  9    6    6

First parallel result:
 34   24   23
 26   20   20
 39   30   30

Second parallel result:
 34   24   23
 26   20   20
 39   30   30

Classic Winograd result:
 34   24   23
 26   20   20
 39   30   30
```

Рис. 4.2: Пример работы ПО.

4.2 Технические характеристики

Технические характеристики ЭВМ, на котором выполнялись исследования:

- ОС: Manjaro Linux 20.1.1 Mikah
- Оперативная память: 16 Гб
- Процессор: Intel Core i7-10510U

При проведении замеров времени ноутбук был подключен к сети электропитания.

4.3 Время выполнения алгоритмов

Алгоритмы тестировались на данных, сгенерированных случайным образом.

Результаты замеров времени приведены в таблицах 4.1 - 4.5. На рисунках 4.3 и ?? приведены графики зависимостей времени работы алгоритмов от размерности обрабатываемых матрицы. В таблице КВ - последовательный алгоритм Коперсмита-Винограда, Парал1КВ - реализация схемы 2.10 параллельного алгоритма Коперсмита-Винограда, Парал2КВ - реализация схемы 2.11 параллельного алгоритма.

Таблица 4.1: Замеры времени для различных размеров массивов на 1 потоке.

Размеры матрицы	Время работы, нс		
	КВ	Парал1КВ	Парал2КВ
100	4692982	6144734	6125329
200	49620266	56025172	58077386
300	178730542	199925329	206518482
400	429056319	456832433	561695647
500	856058247	737473554	673711435

Таблица 4.2: Замеры времени для различных размеров массивов на 2 потоках.

Размеры матрицы	Время работы, нс		
	КВ	Парал1КВ	Парал2КВ
100	4250527	4315274	2910312
200	50728798	34809356	18273241
300	186202130	120880876	61885504
400	436860056	248104438	137340969
500	870833440	576761740	267199177

Таблица 4.3: Замеры времени для различных размеров массивов на 4 потоках.

Размеры матрицы	Время работы, нс		
	КВ	Парал1КВ	Парал2КВ
100	4485824	4239416	2969518
200	51678241	28600648	10260636
300	176448578	80224609	31277241
400	436787067	208705553	70302333
500	850999256	485612161	151094601

Таблица 4.4: Замеры времени для различных размеров массивов на 8 потоке.

Размеры матрицы	Время работы, нс		
	КВ	Парал1КВ	Парал2КВ
100	4403638	5304634	3712704
200	52234883	26568390	11046163
300	177331754	107137650	31637714
400	442171563	337629280	58383234
500	877219464	883596315	117636357

Вывод

При сравнении результатов замеров по времени видно, что выводы в каждом из случаев неоднозначны.

В случае уже отсортированных массивов, однозначно лучше всех показывает себя сортировка вставками. Связано это с особенностями реализации алгоритма, зависимость в текущих обстоятельствах будет, теоретически, линейной. Поэтому целесообразно указать на факт того, что сортировка пузырьком уже на 3000 элементов будет работать медленнее, чем сортировка выбором на $\approx 50.8\%$, при этом с увеличением количества элементов массива, это значение будет расти.

В случае упорядоченного в обратном порядке массива, что можно назвать наихудшим случаем, лучше всех показывает себя сортировка

Таблица 4.5: Замеры времени для различных размеров массивов на 16 потоках.

Размеры матрицы	Время работы, нс		
	КВ	Парал1КВ	Парал2КВ
100	4463337	7381390	4981564
200	50528949	32361454	17479130
300	175677695	110593792	44960321
400	433177913	342768739	68839512
500	831773652	887656714	100236284

выбором. При этом, разница между временем работы сортировки выбором и сортировки вставками на 4000 элементов массива будет равна $\approx 15\%$. Но наибольшая разница, которая заключается между временем работы сортировки выбором и сортировки пузырьком на 5000 элементов: $\approx 340\%$.

Ситуация несколько изменится при рассмотрении скорости работы алгоритмов на массивах чисел, сгенерированных случайно. Здесь можно будет увидеть преобладание алгоритма сортировки выбором над алгоритмом сортировки вставками. На значениях величины массива, равных 4000, разница в скорости обработки будет составлять $\approx 65\%$. Наибольшая выявленная разница в скорости работы алгоритма вставками и алгоритма сортировки пузырьком будет составлять $\approx 260\%$.

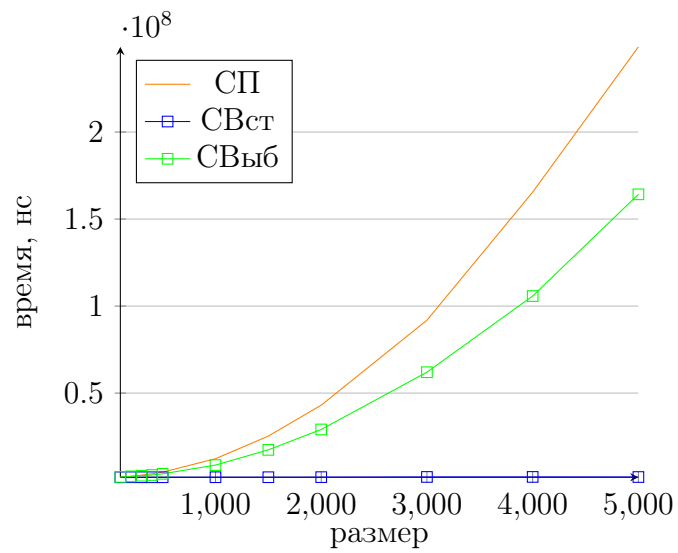


Рис. 4.3: Зависимость времени работы от размера отсортированных массивов

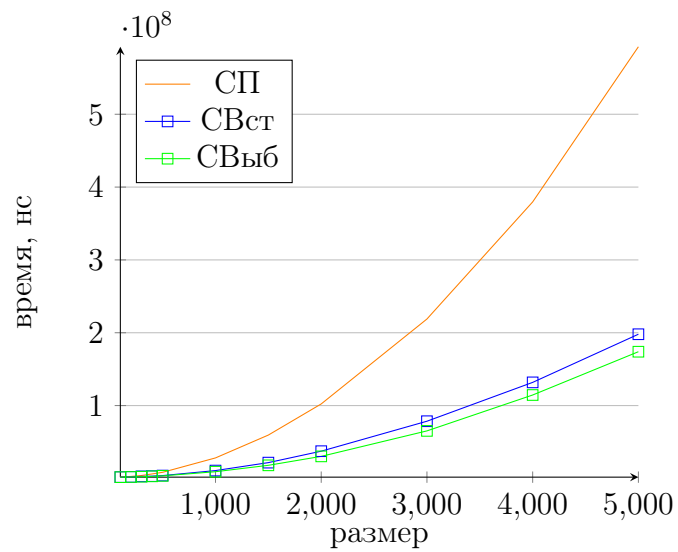


Рис. 4.4: Зависимость времени работы от размера отсортированных в обратном порядке массивов

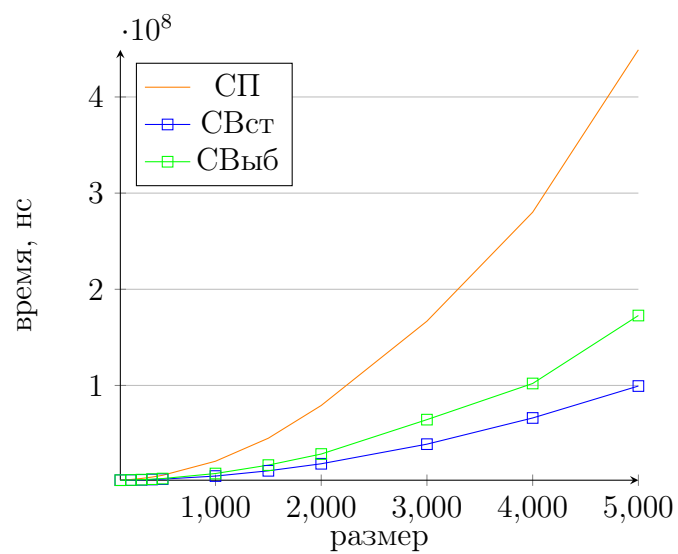


Рис. 4.5: Зависимость времени работы от размера массивов, заданных случайным образом

Заключение

В ходе выполнения лабораторной работы:

1. были изучены алгоритмы сортировки пузырьком, вставками и выбором;
2. были реализованы алгоритмы сортировки пузырьком, вставками и выбором;
3. был проведён сравнительный анализа трудоёмкости алгоритмов на основе теоретических расчетов и выбранной модели вычислений;
4. был проведён сравнительный анализ алгоритмов на основе экспериментальных данных;
5. был подготовлен отчёт по лабораторной работе;
6. были получены практические навыки реализации алгоритмов на ЯП Kotlin.

Исследования показали, что хуже всего себя показала реализация алгоритма сортировки пузырьком.

При сортировке уже отсортированного массива "пузырёк" будет работать на $\approx 66\%$ медленнее при сортировке 5000 элементов, чем алгоритм сортировки выбором. Но наилучшим вариантом в данном случае будет выбор алгоритма сортировки вставками, так как возрастание скорости его работы в данном случае линейно.

В случае отсортированного в обратном порядке массива лучше всего себя показывает сортировка вставками, она в ≈ 1.38 раз быстрее алгоритма сортировки выбором. Наибольшая разница, которая заключается между временем работы сортировки выбором и сортировки пузырьком на 5000 элементов: $\approx 340\%$.

Но в случае неупорядоченных массивов лучше всего себя показывает сортировка вставками, она в ≈ 1.74 раза быстрее, чем сортировка выбором уже на 5000 элементов. Наибольшая выявленная разница в скорости работы алгоритма вставками и алгоритма сортировки пузырьком будет составлять $\approx 260\%$

Литература

- [1] Coppersmith D., Winograd S. Matrix multiplication via arithmetic progressions // Journal of Symbolic Computation. 1990. no. 9. P. 251–280.
- [2] Погорелов Дмитрий Александрович Таразанов Артемий Михайлович Волкова Лилия Леонидовна. Оптимизация классического алгоритма Винограда для перемножения матриц // Журнал №1. 2019. Т. 49.
- [3] Nim documentation [Электронный ресурс]. Режим доступа: <https://nim-lang.org/documentation.html> (дата обращения 09.10.2020).