

**Реализация преобразования Барроуза-Уилера с помощью языка C++**

*Ковалева Ирина Валерьевна*

*Приамурский государственный университет имени Шолом-Алейхема*

*Студент*

*Размахнина Анна Николаевна*

*Приамурский государственный университет имени Шолом-Алейхема*

*Студент*

*Лучанинов Дмитрий Васильевич*

*Приамурский государственный университет имени Шолом-Алейхема*

*Старший преподаватель кафедры информационных систем, математики и методик обучения*

**Аннотация**

В статье рассматривается алгоритм преобразования Барроуза-Уиллера. Описана реализация данного алгоритма на языке программирования C++. Приводятся результаты сжатия, подтверждающие эффективность предложенного метода.

**Ключевые слова:** C++, преобразование Барроуза-Уилера, шифрование.

**Implementation of the conversion of the Burrows-Wheeler using C++**

*Kovaleva Irina Valeryevna,*

*Sholom-Aleichem Priamursky State University*

*Student*

*Razmakhnina Anna Nikolaevna*

*Sholom-Aleichem Priamursky State University*

*Student*

*Luchaninov Dmitry Vasilyevich*

*Sholom-Aleichem Priamursky State University*

*Senior Lecturer of department of information systems, mathematics and teaching methods*

**Abstract**

The article discusses the transformation algorithm burrows-Wheeler. Describes the implementation of this algorithm in the programming language C++.

The results of compression, confirming the effectiveness of the proposed method.

**Keywords:** C++, BWT, encryption.

На сегодняшний день важнейшей задачей теории информации является сжатие данных, с возможностью их восстановления. Эффективным алгоритмом решения данной задачи выступает преобразование Барроуза-Уилера.

Данное преобразование изобретено Д.Дж.Уилером в 1983 году и впервые опубликовано в совместной статье с Барроузом [3].

Алгоритм Барроуза-Уиллера используется, как для сжатия текста без потерь, так и для дополнительной индексной информации к тексту.

BWT преобразование дает более сильное сжатие, но при этом малоприменимо на практике, за счет таких свойств описываемый алгоритм довольно популярен среди разработчиков архиваторов, и среди научных работников.

М.П.Бакулина в своей работе рассматривает эффективное сжатие данных на основе преобразования Барроуза-Уилера [1]. В.Юкин, М.Смирнов, А.Ратушняк, Д.Ватолин в своей статье делают обзор на современные методы сжатия данных, устройство архиваторов, сжатие изображений и видео [2].

Целью данного исследования является реализация преобразования Барроуза-Уилера на языке C++.

Преобразование Барроуза-Уилера состоит из трёх этапов:

1. Создается таблица циклических сдвигов исходной строки.
2. Осуществляется сортировка строк таблицы в алфавитном порядке.
3. Последний столбец таблицы преобразования является выходной строкой, а номер строки совпадает с исходной.

Рассмотрим пример работы алгоритма BWT.

Пусть нам дана исходная строка "GoodBye".

Этапы трансформации более наглядно предоставлены в таблице (табл. 1).

Таблица 1. Алгоритм преобразования Барроуза-Уилера

Трансформация			
Вход	Циклические сдвиги	Сортировка строк	Выход
GoodBye	GoodBye eGoodBy yeGoodB ByeGood dByeGoo odByeGo oodByeG	ByeGood GoodBye dByeGoo eGoodBy odByeGo oodByeG yeGoodB	deoyoGB

На первом этапе осуществляются циклические сдвиги, последний символ заданной строки переходит в её начало.

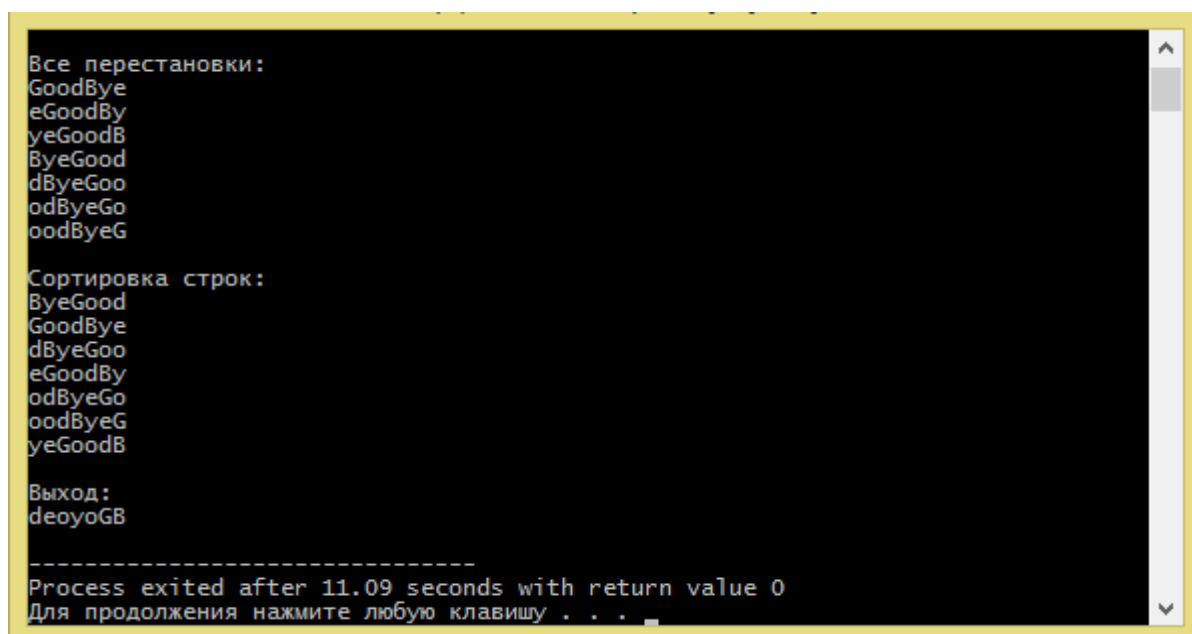
```
for (int i=0;i<k;i++) {
    St [i] =s;
    t+=s [k-1];
```

```
s.insert (0,t,0,1);  
s.erase (k,1);  
t.clear ();  
cout<<st[i]<<endl;}
```

Далее происходит сортировка полученных строк в алфавитном порядке.

```
cout <<endl;  
cout <<"Сортировка строк:"<<endl;  
sort(st,k);  
for (int i=0;i<k;i++){  
    t+=st[i][k-1];  
    cout<<st[i]<<endl;}
```

В итоге результат работы преобразования Барроуза-Уилера будет выглядеть (рис.1).



```
Все перестановки:  
GoodBye  
eGoodBy  
yeGoodB  
ByeGood  
dByeGoo  
odByeGo  
oodByeG  
  
Сортировка строк:  
ByeGood  
GoodBye  
dByeGoo  
eGoodBy  
odByeGo  
oodByeG  
yeGoodB  
  
Выход:  
deouoGB  
  
-----  
Process exited after 11.09 seconds with return value 0  
Для продолжения нажмите любую клавишу . . .
```

Рис. 1. Результат работы программы

В статье был рассмотрен алгоритм преобразования данных, позволяющий получить поток для сжатия. На эффективность данного преобразования оказывают влияние лексический порядок следования символов и направление обхода кодера (справа налево или слева направо).

На основании результата работы программы можно видеть, что данный алгоритм довольно прост в использовании и может использоваться для сжатия различного текста.

## Библиографический список

1. Бакулина М.П. Использование преобразования Барроуза - Уилера для сжатия данных // Информационные технологии. Радиоэлектроника.

- Телекоммуникации. 2012. Т. 1. № 2. С. 154-158.
2. Юкин В., Смирнов М., Ратушняк А., Ватолин Д. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. М.: Диалог-МИФИ. 2003. 381 с.
  3. Burrows M., Wheeler D.J. A block-sorting lossless data compression algorithm. Technical Report 124. Palo Alto: Digital SRC, 1994.
  4. Преобразование Барроуза-Уилера // Университет ИТМО URL: [http://neerc.ifmo.ru/wiki/index.php?title=Преобразование\\_Барроуза-Уилера](http://neerc.ifmo.ru/wiki/index.php?title=Преобразование_Барроуза-Уилера) (дата обращения: 17.12.2016).