**ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ**

**Ордена Трудового Красного Знамени федеральное**

**государственное бюджетное образовательное**

**учреждение высшего образования**

**Московский технический университет**

**связи и информатики**

──────────────────────────────────────

[Кафедра информационной безопасности и автоматизации](http://www.kibia.ru/)

**Доклад**

по дисциплине «Сжатие и хранение данных»

на тему:

**«Особенности применения различных архиваторов»**

Выполнил: студ. гр. БПЗ1401

Груздев С. П.

Проверил:

Симонян А.Г.

Москва 2017

**Архиваторы** – программы для упаковывания и сжатия файлов. Они необходимы для компактного хранения файлов на каком-нибудь накопителе информации, в том числе и в облачных хранилищах. И для более быстрой передачи файлов ввиду их меньшего веса через Интернет.

Архиваторы отличаются друг от друга степенью и скоростью сжатия файлов, количеством поддерживаемых форматов, потреблением вычислительных ресурсов компьютера во время сжатия, а также стоимостью.

**Сжа́тие да́нных** — алгоритмическое преобразование [данных](https://ru.wikipedia.org/wiki/%D0%94%D0%B0%D0%BD%D0%BD%D1%8B%D0%B5), производимое с целью уменьшения занимаемого ими объёма.

Сжатие основано на устранении [избыточности](https://ru.wikipedia.org/wiki/%D0%98%D0%B7%D0%B1%D1%8B%D1%82%D0%BE%D1%87%D0%BD%D0%BE%D1%81%D1%82%D1%8C), содержащейся в исходных данных. Простейшим примером избыточности является повторение в [тексте](https://ru.wikipedia.org/wiki/%D0%A2%D0%B5%D0%BA%D1%81%D1%82) фрагментов. Подобная избыточность обычно устраняется заменой повторяющейся последовательности ссылкой на уже закодированный фрагмент с указанием его длины. Другой вид избыточности связан с тем, что некоторые значения в сжимаемых данных встречаются чаще других. Сокращение объёма данных достигается за счёт замены часто встречающихся данных короткими кодовыми словами, а редких — длинными ([энтропийное кодирование](https://ru.wikipedia.org/wiki/%D0%AD%D0%BD%D1%82%D1%80%D0%BE%D0%BF%D0%B8%D0%B9%D0%BD%D0%BE%D0%B5_%D0%BA%D0%BE%D0%B4%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5" \o "Энтропийное кодирование)). Сжатие данных, не обладающих свойством избыточности (например, [случайный сигнал](https://ru.wikipedia.org/wiki/%D0%A1%D0%BB%D1%83%D1%87%D0%B0%D0%B9%D0%BD%D1%8B%D0%B9_%D1%81%D0%B8%D0%B3%D0%BD%D0%B0%D0%BB) или [белый шум](https://ru.wikipedia.org/wiki/%D0%91%D0%B5%D0%BB%D1%8B%D0%B9_%D1%88%D1%83%D0%BC), зашифрованные сообщения), принципиально невозможно без потерь.

Все методы сжатия данных делятся на два основных класса:

* [Сжатие без потерь](https://ru.wikipedia.org/wiki/%D0%A1%D0%B6%D0%B0%D1%82%D0%B8%D0%B5_%D0%B1%D0%B5%D0%B7_%D0%BF%D0%BE%D1%82%D0%B5%D1%80%D1%8C)
* [Сжатие с потерями](https://ru.wikipedia.org/wiki/%D0%A1%D0%B6%D0%B0%D1%82%D0%B8%D0%B5_%D1%81_%D0%BF%D0%BE%D1%82%D0%B5%D1%80%D1%8F%D0%BC%D0%B8)

При использовании сжатия без потерь возможно полное восстановление исходных данных, сжатие с потерями позволяет восстановить данные с искажениями, обычно несущественными с точки зрения дальнейшего использования восстановленных данных. Сжатие без потерь обычно используется для передачи и хранения текстовых данных, компьютерных программ, реже — для сокращения объёма [аудио](https://ru.wikipedia.org/wiki/%D0%90%D1%83%D0%B4%D0%B8%D0%BE)- и [видеоданных](https://ru.wikipedia.org/wiki/%D0%92%D0%B8%D0%B4%D0%B5%D0%BE), [цифровых фотографий](https://ru.wikipedia.org/wiki/%D0%A6%D0%B8%D1%84%D1%80%D0%BE%D0%B2%D0%B0%D1%8F_%D1%84%D0%BE%D1%82%D0%BE%D0%B3%D1%80%D0%B0%D1%84%D0%B8%D1%8F) и т. п., в случаях, когда искажения недопустимы или нежелательны. Сжатие с потерями, обладающее значительно большей, чем сжатие без потерь, эффективностью, обычно применяется для сокращения объёма аудио- и видеоданных и цифровых фотографий в тех случаях, когда такое сокращение является приоритетным, а полное соответствие исходных и восстановленных данных не требуется.

**WinRAR**

*Программа разработана отечественным программистом Евгением Рошалом в 1995 году.*

*WinRAR хоть и является платной программой, однако покупать ее не обязательно – разработчики насильно не перекрывают доступ к работе с архиватором: по истечении пробного периода просто появляется уведомление, что неплохо было бы купить данный архиватор.*

**Достоинства WinRAR**

1. Поддержка большого количества форматов сжатия для чтения и разархивирования (более 15): RAR, ZIP, 7z, ISO и многих других.
2. Поддержка большого числа операционных систем: Windows, Linux, macOS, FreeBSD и Android.
3. Поддержка AES-шифрования файлов при использовании сжатия в формате RAR.
4. Поддержка многоядерности процессора.
5. Возможность создания самораспаковывающихся архивов.

**Недостатки WinRAR**

1. Архивирование файлов только в форматах ZIP и RAR.

**7-Zip**

*Следующим по счету архиватором идет*[*7-Zip*](http://soft-storage.ru/soft/sistema/arhivatory/7-zip)*, являющийся также очень популярным из-за своей высокой степени сжатия и бесплатности. Создан российским разработчиком Игорем Павловым и распространяется по лицензии GPU LGPL.*

**Достоинства 7-Zip**

1. Поддержка большого числа форматов для архивирования: 7z, ZIP, bzip2, gzip, tar, XZ и WIM.
2. Работает на операционных системах Windows, Linux и macOS.
3. Поддержка чтения и разархивирования файлов в более чем 20 различных форматах.
4. Высокая степень сжатия ввиду поддержки следующих методов архивирования файлов: LZMA, LZMA2, bzip2, PPmd, Deflate и Deflate64.
5. Поддержка AES-шифрования для форматов 7z и ZIP.

**Недостатки 7-Zip**

Серьезных недостатков выявлено не было.

**WinZip**

*Достаточно известный архиватор, разработанный компанией WinZip Computing. Программа является платной, однако у нее присутствует пробный период. WinZip работает на операционных системах Windows и macOS. Скачать программу можно с*[*официального сайта*](http://www.winzip.com/)*.*

**Преимущества WinZip**

1. Поддержка более 15 форматов сжатия для чтения и разархивирования.
2. Поддержка алгоритмов LZMA, PPmd, bzip2, JPEG сжатия файлов, что позволяет создать архив меньшего размера (однако увеличивается время архивирования и разархивирования).
3. Поддерживает AES-шифрование файлов.
4. Возможность загрузки архивов на облачные хранилища и на FTP. Поддерживаются: Google Drive, Dropbox, OneDrive, SugarSync, CloudMe и другие.
5. Возможность резервного копирования файлов.
6. Красивый и понятный интерфейс.
7. Поддержка многоядерности.

**Недостатки WinZip**

1. Сжатие файлов только в форматах ZIP и ZIPX.

**Сравнение степени и скорости сжатия архиваторов**

Конечно, функциональность архиватора – это очень хорошо, однако куда лучше, когда он отлично выполняет свои основные функции: архивирование и разархивирование. В данном тестировании мы посмотрим, как справятся со своими функциями описанные выше программы-архиваторы. Следует сразу отметить, что тестирование не претендует на то, чтобы быть абсолютно верным и точным, потому как оно является всего лишь частным экспериментом.

Производилось архивирование файлов, общий размер которых составляет 1 Гб (в основном – программный код игры), при различных степенях сжатия. По итогам эксперимента были отобраны наиболее подходящие конфигурации для каждого из архиваторов и сформированы две таблицы: в первой программы сравниваются по степени сжатия, во второй – по минимальному потреблению ресурсов.

*Тестирование проводилось на компьютере с процессором Intel Core i3-4170 (двухъядерный на частоте 3,7 Ггц), 8 Гб оперативной памяти и видеокартой AMD Radeon R9 270 с 1 Гб памяти.*

*Приведем также версии архиваторов, на которых выполнялось тестирование:*

* *WinRAR. Версия 5.31.*
* *7-Zip. Версия 16.02.*
* *WinZip. Версия 19.0.*
* *Hamster Free ZIP Archiver. Версия 3.0.0.8.6.*
* *IZArc. Версия 4.2.*

**Архиватор с максимальной степенью сжатия**

Зачастую именно степень сжатия является определяющим фактором для выбора той или иной программы-архиватора. Форматы сжатия могут работать по нескольким представленным алгоритмам на выбор пользователя либо могут сочетать сразу несколько алгоритмов, и каждый алгоритм применяется к конкретному файлу в зависимости от его типа данных (изображение, текстовый файл, программа и так далее).

Как видно, **лидером по степени сжатия является WinRAR**. Он мало того, что лучше всех уменьшает вес файлов, так еще и обгоняет все другие архиваторы по скорости сжатия! К тому же, следует обратить внимание на очень умеренное потребление оперативной памяти – всего-то 430 Мб при таких впечатляющих результатах. При этом при всем WinRAR использует процессор на 60-80%, что много, но не критично при таких показателях.

Программы 7-Zip, Hamster Free ZIP Archiver и IZArc используют один и тот же формат и метод сжатия (Hamster предположительно использует алгоритм LZMA), и вес архива в итоге получается одинаковый – 781 Мб. Скорость архивирования и потребление ресурсов отличаются: по оперативной памяти выигрывает Hamster, по ресурсам процессора и по скорости сжатия – 7-Zip. IZArc здесь проигрывает своим конкурентам.

В этой гонке WinZip остается не у дел: архиватор сильно проигрывает аналогичным программам в степени сжатия, да и в потреблении ресурсов сильно не выделяется. Единственное – скорость сжатия все же не такая большая.

**Архиватор с минимальным потреблением ресурсов**

Существуют и другие пользователи: те, которым в силу маломощности компьютера или по иным причинам требуется небольшое расходование вычислительных ресурсов. Конечно, можно просто заархивировать файлы без сжатия, но это не наша цель. Попробуем найти конфигурацию, при которой архиватор почти не требователен к вычислительной мощности компьютера, но при этом хотя бы немного уменьшает вес конечного архива.

Выбирались архиваторы с наилучшим соотношением ресурсы/размер/время с уклоном, разумеется, в сторону ресурсов. Если в некоторой конфигурации чуть больше «съедалось» ресурсов, но взамен «давался» намного меньший размер или время, то выбор делался в пользу такой конфигурации.

**В итоге получаем картину, представленную в таблице ниже.**

Здесь у нас, казалось бы, все претенденты неплохи. Однако давайте рассмотрим ситуацию поближе и повнимательнее. Hamster Free ZIP Archiver хоть и выигрывает у всех других архиваторов в скорости, но во многом одинаков с ними по иным показателям и проигрывает по потреблению оперативной памяти. Да, 180 мегабайт в наше время – ничто, но мало ли кому в какой ситуации может потребоваться крайне низкое расходование вычислительных ресурсов? Так что будем рассматривать и сравнивать абсолютные показатели.

Другой кандидат на выбывание – WinZip. Он намного лучше ужимает файлы, однако потребляет схожее количество памяти, что и Hamster, и в конце полностью «разбивается» о 7-Zip, проигрывая ему по трем показателям.

Остаются 3 претендента. Конечно, по потреблению ресурсов здесь безусловно побеждает IZArc. Однако WinRAR несколько лучше в плане скорости, если имеется возможность расходовать немного больше ресурсов процессора. А 7-Zip на порядок лучше уменьшает вес файлов, но в данном случае потратится намного больше времени и оперативной памяти. В конечном счете каждый пользователь выберет архив под свои нужды.

| Максимальная степень сжатия у разных архиваторов | | | | | |
| --- | --- | --- | --- | --- | --- |
| Архиватор | WinRAR | 7-Zip | WinZip | Hamster Free ZIP Archiver | IZArc |
| Оперативная память | 430 Мб | 380 Мб | 300-400 Мб | 186 Мб | 680 Мб |
| Процессор | 60-80% | ~50% | 40-60% | 35-50% | 45-55% |
| Время сжатия (мин:сек) | 01:32 | 02:31 | 02:18 | 04:43 | 04:03 |
| Размер архива | 774 Мб | 781 Мб | 828 Мб | 781 Мб | 781 Мб |
| Уменьшение веса (в %) | ↓25% | ↓24% | ↓19% | ↓24% | ↓24% |
| Формат сжатия | RAR5 | 7z | ZIPX | 7z | 7z |
| Метод сжатия | – | LZMA | Наилучший | – | LZMA |
| Степень сжатия | Макс. | Макс. | – | Макс. | Наилучшая |

| Минимальное потребление ресурсов у архиваторов | | | | | |
| --- | --- | --- | --- | --- | --- |
| Архиватор | WinRAR | 7-ZIP | WinZip | Hamster Free ZIP Archiver | IZArc |
| Оперативная память | ~10 Мб | ~70 Мб | ~160 Мб | ~180 Мб | ~6 Мб |
| Процессор | ~35% | ~27% | 20-30% | ~27% | 15-30% |
| Время сжатия (мин:сек) | 00:38 | 07:06 | 09:04 | 00:32 | 01:28 |
| Размер архива | 967 Мб | 887 Мб | 890 Мб | 969 Мб | 968 Мб |
| Уменьшение веса (в %) | ↓6% | ↓14% | ↓13% | ↓5% | ↓5% |
| Формат сжатия | ZIP | 7z | ZIPX | ZIP | ZIP |
| Метод сжатия | – | PPMd | PPMd | – | Deflate |
| Степень сжатия | Макс. | Макс. | – | Нормальная | Наилучшая |