

11 класс

Урок по теме «Самораспаковывающиеся и многотомные архивы»

Цель урока:

- Рассмотреть практические ситуации, связанные с необходимостью создания самораспаковывающихся и многотомных архивов; научить создавать такие архивы.
- *Развивать* сообразительность учащихся.

Программное обеспечение:

Архиватор WinRAR

Ход урока

I. Вводная беседа.

Вопросы для обсуждения:

1. Актуальна ли проблема уплотнения информации на сегодняшнем этапе развития компьютерной техники? Если актуальна, то почему?
2. Почему данные и программы можно сжать (упаковать)?
3. Как называются программы для работы с уплотнённой информацией?
4. Назовите методы сжатия информации.
5. На доске запись: A:, C:, D:.. Что это означает?
6. На доске запись: D:\Мои документы\NetWalk.rar. Что это означает?

II. Проверка знаний учащихся.

Включает две части: теоретическую (устный ответ у доски) и практическую (создание простого архива).

1. Вопросы для теоретической части:

1. Методы уплотнения информации: метод RLE, метод Хаффмана.
(Можно кинуть жребий кому из двоих учащихся отвечать у доски метод RLE, кому метод Хаффмана.)

2. Задания для практической части:

1. В папке D:\Мои документы\Ученик\Архив\ найти файл *Happy.zip* и распаковать его в эту же папку.
2. Упаковать полученные из архива файлы в формат *Happy.rar*.
3. Сравнить степень сжатия информации.
4. Очистить папку \Архив, но не удалять.

III. Объяснение нового материала.

До сих пор мы учились создавать простые (обычные) архивы. Но бывают ситуации, когда необходимо использовать другие архивы. Рассмотрим их.

Самораспаковывающийся (SFX, от англ. **Self eXtracting**) архив – это архив, к которому присоединён исполняемый модуль. Этот модуль позволяет извлекать файлы простым запуском архива как обычной программы. Таким образом, для извлечения содержимого SFX-архива не требуется дополнительных внешних программ. Вместе с тем, WinRAR может работать с SFX-архивом точно так же, как и с любым другим, поэтому если вы не хотите запускать SFX-архив (например, когда не можете гарантировать, что в нём нет вирусов), то для просмотра или извлечения его содержимого можно использовать WinRAR.

- SFX-архивы, как и любые другие исполняемые файлы, обычно имеют расширение .exe.
- SFX-архивы удобны в тех случаях, когда нужно передать кому-то архив, но вы не уверены, что у адресата есть соответствующий архиватор для его распаковки. Обычно в виде SFX-файлов распространяются различные фирменные продукты, которые могут быть установлены на любом компьютере.

Непрерывный (solid) архив – это архив, упакованный специальным способом, при котором все сжимаемые файлы рассматриваются как один последовательный поток данных. Непрерывная архивация в программе WinRAR поддерживается только в формате RAR, для формата ZIP такого типа упаковки не существует. Метод сжатия для архивов RAR – обычный или непрерывный – выбирается пользователем.

Непрерывная архивация значительно увеличивает степень сжатия, особенно при добавлении в архив существенного количества небольших файлов с похожим содержанием. Однако следует иметь в виду, что у непрерывной архивации есть и некоторые недостатки, например:

- работа с непрерывными архивами происходит медленнее;
- если в непрерывном архиве какой-либо файл окажется повреждённым, то не удастся извлечь и все файлы, следующие после него.

Многотомные архивы – это архивы, состоящие из нескольких частей, – *томов*.

Тома поддерживаются только в формате RAR, но не в ZIP. Обычно тома используются для сохранения большого архива на нескольких дискетах или других сменных носителях.

По умолчанию тома RAR получают имена вида «имя_архива.partN.rar», где N – номер тома. Если по каким-либо причинам эта схема именования томов вас не устраивает, то можно включить старую схему, основанную на расширениях файлов, при которой первый том многотомного архива получает расширение .rar, а расширения последующих томов нумеруются как .r00, .r01, .r02 и так далее до .r99.

- Уже созданные многотомные архивы не допускают изменения, т.е. в них нельзя добавлять, обновлять или удалять файлы.
- Для распаковки томов необходимо начинать извлечение с первого тома. Если тома находятся на несменном носителе (например, на жёстком диске), то сначала нужно переписать все тома в одну папку.

Тома также могут быть непрерывными и самораспаковывающимися.

Практическая работа с ЭВМ. Создание самораспаковывающегося и многотомного архивов.

Задания:

1. Создать самораспаковывающийся архив Harry.exe. Проанализировать свойства архива. Убедиться, что для его распаковки не нужен архиватор.

2. Найти папку D:\Мои документы\Ученик\NetWalk и скопировать её в D:\Мои документы\Ученик\Архив\.
3. Создать многотомный непрерывный архив папки NetWalk на двух дискетах.

IV. Подведение итогов урока и домашнее задание.

§ 1.3 (практикум), кроме пункта «Файловые менеджеры»;
закреплять навыки работать с архивами при использовании Интернета.

Материал для учащихся

Метод RLE

Возьмём, например, файл с чёрно-белым рисунком, который может состоять из байтов всего двух видов: 0 – чёрная точка, 255 – белая и заменим повторяющиеся последовательности парой чисел. Первое указывает цвет точки, а второе – количество точек. Например:

■■■■■■■■■■ = 0 0 0 0 0 0 0 0 = 0 8 (коэффициент сжатия 4:1)

■■■■□□□□□ = 0 0 0 255 255 255 255 255 = 0 3 255 5 (сжатие 2:1)

Как видно, степень уплотнения зависит от состава данных. Бывают данные, которые при уплотнении таким методом станут не меньше, а больше, чем исходные. Методы RLE хорошо работают на графических и табличных данных, но плохо уплотняют текстовые данные.

Метод Хаффмана

Суть этого метода состоит в следующем. На кодирование любого символа требуется один байт (8 бит). Однако это нерационально, потому что одни символы встречаются чаще других. Например, буквы «А» и «О» встречается чаще таких букв, как «Ы» и «Э». По аналогии с азбукой Морзе архиваторы вместо 8-битовой кодировки каждого символа ASCII используют *коды переменной длины*, основанные на частотном анализе сжимаемого текста: чем чаще встречается символ, тем короче его код.

Рассмотрим пример, когда экран монитора плотно забит «текстом» (25 строк по 80 колонок = 2000 символов), содержащим всего четыре разные буквы А, В, С и D, которые встречаются соответственно 1000, 500, 300 и 200 раз. При 8-разрядной кодировке такого документа потребуется 2000 байт или 16000 бит. Если кодировать каждую букву двухразрядным двоичным кодом (А – 00, В – 01, С – 10, D – 11), то потребуется всего 4000 бит:

$$2 \cdot 1000 + 2 \cdot 500 + 2 \cdot 300 + 2 \cdot 200 = 4000.$$

Однако возможна и более эффективная кодировка:

$$A - 0, \quad B - 10, \quad C - 110, \quad D - 111.$$

С её помощью тот же текст можно закодировать ещё меньшим числом двоичных разрядов:

$$1 \cdot 1000 + 2 \cdot 500 + 3 \cdot 300 + 3 \cdot 200 = 3500.$$

При таком методе упаковки данных и программ в состав каждого уплотнённого файла должна входить таблица, по которой он кодировался (у каждого файла она своя), иначе файл не разуплотнить. К маленькому файлу прикладывать таблицу не выгодно – она займёт места больше, чем сам файл, но чем длиннее файл, тем выгод-

нее применение метода Хаффмана. Метод Хаффмана хорошо работает с текстовыми данными и оцифрованным звуком.