

Лекция

Тема: Сервисное программное обеспечение, утилиты

Введение

Утилиты (англ. utility или tool) — компьютерная программа, расширяющая стандартные возможности оборудования и операционных систем, выполняющая узкий круг специфических задач.

Утилиты предоставляют доступ к возможностям (параметрам, настройкам, установкам), недоступным без их применения, либо делают процесс изменения некоторых параметров проще (автоматизируют его).

Утилиты зачастую входят в состав операционных систем или идут в комплекте со специализированным оборудованием.

Функции утилит

Мониторинг показателей датчиков и производительности оборудования — мониторинг температур процессора, видеоадаптера; чтение S.M.A.R.T. жёстких дисков; бенчмарки.

Управление параметрами оборудования — ограничение максимальной скорости вращения CD-привода; изменение скорости вращения кулеров.

Контроль показателей — проверка ссылочной целостности; правильности записи данных.

Расширение возможностей — форматирование и/или переразметка диска с сохранением данных, удаление без возможности восстановления.

Типы утилит

Дисковые утилиты

- Дефрагментаторы
- Проверка диска — поиск неправильно записанных либо повреждённых различным путём файлов и участков диска и их последующее удаление для эффективного использования дискового пространства.

- o CHKDSK

- o fsck

- o Scandisk

- Очистка диска — удаление временных файлов, ненужных файлов, чистка «корзины».

- o Очистка диска

- o CCleaner

- Разметка диска — деление диска на логические диски, которые могут иметь различные файловые системы и восприниматься операционной системой как несколько различных дисков.

- o PartitionMagic

- o GParted

- o fdisk

- Резервное копирование — создание резервных копий целых дисков и отдельных файлов, а также восстановление из этих копий.

- Сжатие дисков — сжатие информации на дисках для увеличения вместимости жёстких дисков.

- Менеджеры процессов

- o AnVir Task Manager

Утилиты работы с реестром

- CCleaner

- Reg Organizer

Утилиты мониторинга оборудования и бенчмарки

- SpeedFan

Тесты оборудования

Любому специалисту, так или иначе связанному с администрированием локальных вычислительных сетей (ЛВС), когда-нибудь приходится сталкиваться с необходимостью анализа их работы. Мониторинг сети необходим не только в случае возникновения каких-либо проблем, но и в качестве профилактики для обеспечения безотказной и корректной работы ЛВС.

Существует немало способов контроля работы сети — начиная от использования диагностических средств операционных систем, к примеру встроенного в операционные системы семейства Windows набора утилит мониторинга Monitoring performance, позволяющего весьма детально оценить работу отдельного компьютера, в том числе его сетевого интерфейса, и заканчивая специализированными аппаратными комплексами. Но на практике ни встроенные средства ОС, ни аппаратные средства мониторинга зачастую не удовлетворяют потребностям и возможностям системного администратора по организации полноценного мониторинга сети. В первом случае можно проанализировать лишь работу отдельного узла, а не сети в целом, вследствие чего получить целостную картину состояния сети без наличия определенного опыта и точного знания структуры конкретной ЛВС и точного понимания принципов его работы становится практически невозможно.

Второй же подход требует внушительных финансовых затрат, и не каждая организация готова пойти на это, тем более что в понимании многих руководителей расходовать деньги на нормально работающую сеть — слишком большая роскошь. В подобной ситуации своего рода панацеей становятся утилиты сетевого мониторинга, позволяющие проводить анализ сетевого трафика как отдельных узлов, так и сети в целом, а наиболее продвинутые из них анализируют и топологию сети, чтобы выявить конфликты, возникающие в результате ошибок, допущенных при построении ЛВС.

В настоящее время имеется огромное количество утилит, различающихся как по своим возможностям, так и по доступности и удобству интерфейса.

Сервисные программы (утилиты).

Самые необходимые для работы утилиты входят в состав операционной системы. Это программы, которые позволяют обслуживать диски (проверять, сжимать, дефрагментировать и т.д.), выполнять операции с файлами (архивировать и т.д.), работать в компьютерных сетях. В Windows основной набор утилит находится в папке Программы – Стандартные – Служебные, в Windows 2000 все утилиты, связанные с обслуживанием дисков объединены в программе Управление дисками, которая находится в оснастке Управление компьютером.

Обслуживание дисков.

Форматирование.

Физическое форматирование. Физическое форматирование жесткого диска делит поверхности магнитных пластин на дорожки и сектора:

- **Дорожки** — это концентрические окружности, описываемые головками чтения/записи на магнитных поверхностях. Дорожки нумеруются, начиная с нуля, по направлению к центру.

- Дорожки в свою очередь делятся на небольшие области, называемые **секторами**, содержащими фиксированное количество информации. Обычный размер сектора составляет 512 байт.

Логическое форматирование. Физически отформатированный жесткий диск должен быть отформатирован логически. Логическое форматирование заключается в размещении на диске файловой системы, при этом различные операционные системы могут использовать разные файловые системы.

Вызов команды **Форматировать** осуществляется из контекстного меню диска или из командной строки.

Тестирование диска.

При сбоях в работе компьютера, зависаниях ОС и по другим причинам системные области на диске могут быть изменены. Для выявления и исправления этих нарушений производится проверка нарушений файловой системы, проверка физической поверхности диска на наличие «сбойных секторов» и перенос данных из этих участков в безопасные.

Программа – **Проверка диска** (Scandisk).

Дефрагментация диска.

Для записи файлов на диск ОС выделяет участки памяти (кластеры), которые не обязательно расположены друг за другом. При удалении и перезаписи файлов на диске образуется много пустых мест, а файлы оказываются фрагментированными, это замедляет работу.

Приведем пример. Представим, что у нас данные находятся в полном порядке и идут друг за другом. Затем мы удалили какой-то файл и записываем на его место больший по размеру файл. Запись пойдет с первого пустого фрагмента, и системе придется записать этот файл, разбив его на две части. А если файл находится в результате такой записи в конце диска, то это приведет к потере быстродействия. Если же фрагментирован win386.swp (область подкачки, свопинга), файл, с помощью которого Windows устраняет недостаток ОЗУ, то потеря быстродействия будет еще большей из-за очень частых операций записи/чтения этого файла.

Поэтому периодически нужно устранять фрагментацию диска – дефрагментировать его. Программа – **Дефрагментация диска** (Defrag).

Сводная информация о компьютере и системе.

Производительность компьютера зависит от технических характеристик составляющих его устройств. В случаях возникновения сбоев в работе компьютера или перед его модернизацией полезно осуществить тестирование различных устройств компьютера.

Программа **Сведения о системе** дает сводную информацию об устройствах компьютера. **Системный мониторинг** позволяет в фоновом режиме отслеживать параметры функционирования компьютера, получать информацию о возникших проблемах.

Пакеты утилит.

Известны специализированные пакеты утилит, которые существенно расширяют возможности операционных сервисных программ. Например, пакет Norton Utilities от Symantec, его извечный конкурент McAfee Utilities (прежнее название Nuts&Bolts 98), а также сравнительно недавно появившийся пакет **Fix-It Utilities 3.0**, предлагаемый английской компанией Ontrack Data Systems, хорошо известной своими качественными утилитами для восстановления данных.

Norton Utilities.

Пакет Norton Utilities пользуется большой популярностью по ряду причин: быстрые и эффективные алгоритмы работы, простой и наглядный интерфейс, широкий спектр функциональных возможностей.

Рассмотрим новую версию пакета Norton Utilities 2002. При знакомстве с пакетом первое, что бросается в глаза, это широкая номенклатура поддерживаемых операционных систем. Norton Utilities 2002 успешно работает со всеми современными воплощениями Windows.

Управление всеми основными утилитами пакета сосредоточено в Norton Utilities Integrator. Утилиты распределены в 4-х основных разделах, выбираемых нажатием кнопки в левой части окна, по следующим тематическим признакам:

- Optimize Performance - оптимизация системы
- Find and Fix Problems - поиск и исправление ошибок
- System Maintenance - обслуживание системы

- Registry Management - управление системным реестром

Кроме этого, в окне Integrator'a можно создать спасательный диск (Rescue Disk), содержащий архивные копии системных файлов, системной информации, реестра, "образов" дисков и содержимого CMOS, сформированных утилитой Image. При этом создается загрузочный комплект из 3 обычных флоппи-дискет или 100 мегабайтного диска Zip, где, помимо обычных системных файлов, записываются уже известные утилиты для DOS: Diskedit, NDD, Unerase и Unformat, обеспечивающие редактирование диска на низком уровне, диагностику поверхности диска, восстановление ранее удаленных файлов. Но для комфортной работы со спасательными дискетами не помешает добавить, в дополнение к компонентам, устанавливаемым по умолчанию, драйверы для работы с CD-ROM-ом и какой-нибудь компактный файловый менеджер, например, Volcov Commander.

Optimize Performance - оптимизация системы

Speed Disk - одна из важнейших утилит в пакете, позволяющая произвести оптимизацию файловой системы путем ее дефрагментации. В процессе работы Speed Disk обрабатывает разрозненные фрагменты программ на всем дисковом пространстве, перемещая ближе к началу часто используемые (т. е. наиболее часто читаемые) файлы, а архивы и те файлы, потребность в которых не столь актуальна, - в конец диска, причем порядок размещения разных категорий файлов можно задавать самостоятельно. При использовании файла подкачки фиксированного размера он в процессе дефрагментации может быть размещен в начальных секторах дискового раздела, что немного повысит производительность ОС. Программа Speed Disk выполняет дефрагментацию диска значительно быстрее всех известных на сегодня своих аналогов, используемых в других пакетах.

Norton Optimization Wizard оптимизирует системный реестр Windows, удаляя сиротские ключи и ненужные поля, тем самым сокращая его длину. Однако, разрешая этой утилите самостоятельно копаться в реестре, не следует расслабляться и терять бдительность - иногда "ненужные" записи или поля реестра, предлагаемые мастером для удаления, могут оказаться очень даже нужными. Кроме этого, Optimization Wizard может провести проверку скорости дисков и выбрать наилучшее размещение и оптимальный фиксированный размер для файла подкачки. Утилита работает только в среде Win9x/Me, все более современные версии этой ОС подобные процедуры умеют выполнять самостоятельно.

Find and Fix Problems - поиск и исправление ошибок

Norton System Doctor осуществляет "мониторинг" системы и представляет собой панель системных индикаторов, контролирующую не только состояние значительного числа компонентов системы, но и ее производительность.

UnErase Wizard - дублер "Корзины" Windows, позволяющий восстанавливать ранее удаленные файлы. Если для корзины дополнительно установить защиту Norton Protection, то на иконке корзины в углу появляется значок щита, а удаляемые файлы будут дольше оставаться целостными при записи новой информации, и вероятность восстановления удаленных файлов будет больше.

Norton Disk Doctor используется для проверки физической и логической целостности дисков. Проверка проходит настолько быстро и эффективно, что можно рекомендовать использовать ее в сокращенном виде к ежедневному применению, а в полном (с проведением теста поверхности диска) - не реже раза в месяц. Важной особенностью утилиты является возможность отмены внесенных ранее изменений, нажав кнопку Undo. Дополнительно имеется возможность разрешить или запретить режим отображения карты диска при тесте, возможность проверки только занятого пространства и тестирование автоматизируемых дисков и ряд других, полезных и не очень, функций.

Norton WinDoctor осуществляет комплексный анализ состояния в первую очередь программного обеспечения системы: проверяет работоспособность Windows, системного реестра, важнейших библиотек и т.д. на наличие изменений, удаленных системных

файлов и т.п. Пользователю предлагается на выбор или полное, или выборочное тестирование, а также два варианта решения выявленных проблем – автоматический и ручной. В ручном режиме утилита извещает, какие меры предпринимаются, либо предлагает самостоятельно исправить проблему. Для не самых опытных пользователей или экспериментаторов большое значение имеет то, что WinDoctor ведет журнал осуществляемых изменений и исправлений, который дает возможность отката в любой момент времени.

Norton System Check - по своим функциям перекликается и расширяет возможности предыдущей утилиты и предназначен в основном для выявления и решения наиболее общих проблем системы, в число которых входят проблемы с дисками, системными файлами, системным реестром. Предлагается четыре варианта поиска неисправностей на выбор: дисковые проблемы, проблемы Windows, увеличение быстродействия и резервное копирование служебной информации на дискеты. После окончания проверки системы выводится перечень выявленных неполадок и предлагается устранить их в ручном или автоматическом режиме. Возможность отката предусмотрена.

Norton System Check и Norton WinDoctor, взаимодействуя и хорошо дополняя друг друга, совместно с системным монитором Norton System Doctor представляют из себя целый диагностически-восстановительный программный комплекс, позволяющий выявлять и оперативно устранять возникающие в процессе эксплуатации различные проблемы функционирования системы.

System Maintenance - обслуживание системы

System Information предоставляет достаточно подробную информацию о системе. Вся информация разнесена по 8 закладкам: система в целом, дисплей, принтеры, память, дисковые накопители, устройства ввода, средства мультимедиа, сетевые соединения и Интернет.

Wipe Info позволяет полностью, безвозвратно, затирать файлы или каталоги без возможности их дальнейшего восстановления и очищает свободное место на диске.

Image создает файл image.dat, содержащий "образы" загрузочных записей дисков, таблиц размещения файлов и содержимого наиболее важных каталогов. В случае утери какой-то части системной информации из-за сбоя, такой снимок поможет восстановить работоспособность всей системы. Кроме того, создается и второй файл, image.bak, в котором содержится предыдущий "образ", что страхует от возможного повреждения текущего "снимка" системы. Эта утилита может делать снимок винчестера каждый раз при запуске Windows, тем не менее, настоятельно рекомендуется запускать ее каждый раз после установки или удаления программ, а также любых других действий, влекущих за собой изменение содержимого системных областей дисков.

Norton File Compare производит сравнение текстовых файлов, выделяя одинаковые или различные блоки. После открытия разных версий файлов перед вами предстают два окна, в которых неизменный текст показывается черным, перемещенный - синим, а удаленный - красным. В дополнительном окошке показывается общее количество измененных и неизменных строк. Эту утилиту очень удобно использовать для сравнения и редактирования различных версий конфигурационных файлов или разделов системного реестра.

Registry Management - управление системным реестром

Norton Registry Editor - довольно удобный редактор системного реестра, имеющий несколько очень полезных особенностей. Во-первых, для безопасности работы можно сохранить весь реестр либо отдельные ветви дерева ключей. Во-вторых, ведя журнал изменений с помощью Registry Tracker, возможно осуществлять откат проведенных изменений в реестре. В-третьих, для удобства в реестре устанавливаются закладки по ключам, а также редактируются системные файлы *.ini, которые отображаются не как обычные текстовые файлы, а в виде деревьев, аналогично записям реестра.

Norton Registry Tracker способен фиксировать изменения ключей в системном реестре. Аналогично утилите Image, можно сконфигурировать Registry Tracker таким образом, что он будет выполнять "снимки" реестра каждый раз при запуске Windows. При отсутствии качественного деинсталлятора Registry Tracker может выполнять некоторые его функции: сохранять "образы" системного реестра перед установкой новых приложений.

Вопросы для самопроверки.

1. назначение программ-оболочек ОС. Примеры.
2. каково назначение сервисных программ.
3. в чем заключается обслуживание диска сервисными программами. Какие повреждения диска могут быть устранены, а какие нет.
4. Что представляет собой дефрагментация диска. Для чего и в каких случаях ее необходимо выполнять.
5. приведите краткие характеристики программ, входящих в Norton Utilities.

Сервисные программы, обеспечивающие резервирование и архивирование информации

Основы резервирования информации

Данные – самое ценное из того, что есть на компьютере. Если компьютер используется для создания документов, ценность труда, вложенного в их создание, обычно превышает стоимость самого компьютера.

В аварийной ситуации рабочие программы можно переустановить, если сохранился дистрибутив (а он должен храниться), восстановление программного обеспечения займет максимум несколько дней, а восстановление данных, которые, может быть, создавались годами, может быть уже невозможным.

Искажения или разрушение данных, хранящихся на внешних носителях, может происходить по следующим причинам:

- несанкционированные действия пользователей или вирусных программ
- сбой и отказы в работе аппаратно-программного обеспечения
- физическое разрушения носителя.

Устранить последствия потери данных можно только в том случае, если утерянная информация была зарезервирована.

Резервирование данных:

- фоновое – при каждой записи на диск данные параллельно записываются еще и на другой носитель. Эта функция поддерживается ОС (UNIX, Windows NT). Обеспечивает автоматическое восстановление данных в случае отказа носителя или возникновения дефектов, но не в случае некорректной работы программы или пользователя.
- Периодическое – новые данные с рабочих носителей дублируются на резервные носители через определенные интервалы времени.

Наибольшая эффективность достигается при сочетании этих способов резервирования.

Для предотвращения потери данных, хранящихся на компьютере, рекомендуется придерживаться **принципа раздельного хранения данных и программ**: все, что создается, хранится и обрабатывается на компьютере не должно храниться в тех же папках, что и сами приложения. (т.е. нельзя хранить документы, созданные редактором WORD в папке ...\\WORD, где находится сам текстовый редактор. При переустановке приложения или ОС можно потерять все, что хранилось в этих папках).

Если на одном компьютере работают несколько человек, то каждый пользователь создает для себя одну основную папку и хранит свои документы в этой папке.

Организация резервного копирования.

Виды резервирования.

- 1 Резервирование информации, хранящейся в файловых областях. Может быть выполнено двумя способами:
 - о Обычное копирование. Имеет ряд неудобств: большое количество памяти, необходимой для хранения копий, трудности в работе с копиями.
 - о Архивация. Используя специальные программы-архиваторы, позволяющие сохранять данные в более компактном (сжатом) виде.
- 2 Резервирование информации, хранящейся в нефайловых областях и энергонезависимой памяти компьютера (CMOS). Выполняется с помощью специальных программ (см. NU), которые сохраняют ее в виде файла.
- 3 Резервирование ПО и средств восстановления данных. Для восстановления ПО необходимо иметь дистрибутив установленных программ. “Ремонтный набор” должен включать в себя средства, которые позволяют:
 - о Загрузить ОС
 - о Восстановить системные данные
 - о Найти и устранить ошибки на дисках, выполнить разбиение диска на разделы.
 - о Проверить на наличие вирусов
 - о Разархивировать данные

Организационная схема резервирования данных.

Прежде чем приступить к работе со средствами архивации данных необходимо разработать организационную схему резервного копирования и затем придерживаться ее постоянно.

Необходимо принять решение о:

- 1) числе резервных копий. Обычно используют две копии. Одной недостаточно, т.к. надежность внешних носителей информации невысока (2-3% дискет могут не читаться при восстановлении данных). Большее количество копий затрудняет работу с ними;
- 2) порядке ротации (круговорота) копий. Пока одна копия хранится в удаленном месте, происходит работа с другой копией. Обычно применяют еженедельную ротацию;
- 3) обновлении копий. Архивация может быть полной или частичной при еженедельной ротации полную архивацию проводят раз в неделю, а частичную – каждый день.

Правила резервирования.

1. резервирование необходимо выполнять на внешнем носителе, удовлетворяющим требованиям:
 - носитель используется только для хранения копий.
 - носитель должен быть защищен от несанкционированного доступа
 - носитель должен быть доступен в случае необходимости.
2. Перед резервированием и в случае восстановления необходим антивирусный контроль
3. Зарезервирована должна быть вся ценная информация
4. Необходимо подготовить программные средства восстановления работоспособности системы.

Архивация данных

Архивация – это сжатие данных без потерь.

Программы-архиваторы предназначены для сжатия файлов, т.е. для создания файлов, которые занимают на диске меньше места, чем имеющиеся.

Много лет назад такие программы были очень актуальны, т.к. объем жестких дисков был незначительным (несколько десятков мегабайт) и временно не используемые программы и данные приходилось хранить на дискетах в сжатом виде. Применялись архиваторы и для переноса информации с одного компьютера на другой, если объем переносимых файлов превышал вместимость дискеты.

В настоящее время, казалось бы, ситуация другая – объемы жестких дисков у рядовых пользователей сейчас настолько велики, что многие не используют весь их потенциал, появились и широко используются компакт-диски и другие сменные носители большой емкости. Стоит ли тогда использовать архиваторы?

Стоит, и для этого есть ряд причин:

- при передаче данных по электронной почте критичным является каждый килобайт и, кроме того, при пересылке большого числа файлов проще в письмо вложить всего один файл – архивный – содержащий в сжатом виде все необходимые файлы.
- Пропускная способность локальных сетей ограничена. При пересылке больших массивов информации по сети рекомендуется использовать архиваторы не только для уменьшения объема передаваемых данных, но и с целью упрощения проверки правильности передачи данных: после передачи проще проверить корректность архива, чем проверять сотни отдельных файлов.
- Часто встает проблема, когда на носитель надо перенести сложную структуру каталогов. Решить эту проблему позволяют архиваторы, которые могут упаковать всю файловую структуру в один файл (для таких целей некоторые архиваторы содержат метод сжатия “без сжатия”)
- Многие архиваторы являются очень эффективными кодировщиками, позволяющими скрыть конфиденциальную информацию от чужих глаз, “запаковав” ее в архив и установив пароль доступа к архивному файлу.

Кроме того, архивация данных используется аппаратными средствами компьютера. Так протоколы модемной связи используют простейшие алгоритмы сжатия (например, замена одинаковых цепочек байтов на одну и количество повторений) для увеличения скорости передачи файлов при неизменной пропускной способности канала связи.

Еще один пример. Известно, что компьютерные игры, особенно трехмерные, – наиболее ресурсоемкие программы. Компания 3S предложила новую, эффективную технологию сжатия данных, в результате чего возрастает скорость их передачи, доступная память используется более эффективно. Следовательно, при одинаковой частоте и объеме видеопамати можно добиться более высокой производительности.

Программы-архиваторы

Для создания резервных копий файлов чаще всего используются программы-архиваторы. Используя методы сжатия, они создают копии файлов меньшего размера, и помещают их в один архивный файл или архив.

Различными разработчиками были созданы разные программы-архиваторы, отличающиеся между собой форматами архивных файлов, коэффициентом сжатия данных, скоростью работы, удобством интерфейса. Но ни одна из программ по всем параметрам не превосходит остальные.

Пример архиваторов. Arj, Zip, Rar, WinArj, WinRar, WinZip и др.

Архивный файл

Архивный файл (архив) – это набор из одного или нескольких файлов, помещенных в сжатом виде в один файл.

При необходимости данные из архивного файла могут быть извлечены в первоначальном виде.

Каждый архивный файл имеет **оглавление**, в котором содержится информация о каждом файле, помещенном в архив:

- имя файла
- сведения о исходном размещении
- дата и время последней модификации
- исходный размер и размер в архиве
- код циклического контроля, с помощью которого можно проверить сохранность данных в архиве.

Основные операции с архивами:

- создание нового архива
- добавление файлов в существующий архив
- обновление архива
- просмотр содержимого архива (оглавления)
- распаковка архива (извлечение файлов из архива)
- проверка целостности архива (сохранности данных)

Обычно для распаковки архивного файла нужно иметь тот архиватор, который создал этот архив (это нужно помнить, когда переносите файлы с одного компьютера на другой), но некоторые архиваторы могут создавать специальные архивные файлы, для распаковки которых уже не нужен архиватор. Такие архивные файлы называются **самораспаковывающимися архивами**.

Самораспаковывающиеся архивы представляют собой файл, который содержит и сжатые данные и программу по их распаковки, имеет расширение *exe*.

Методы сжатия.

Существуют разные алгоритмы сжатия информации. Их можно разделить на два класса: алгоритмы **сжатия без потери данных** и алгоритмы **сжатия с потерей данных**.

Если применялся метод сжатия без потери данных, и если знать каким способом был сжат файл, то можно будет “распаковать” его, т.е. вернуть в исходный вид. Иначе говоря, эти методы архивации обратимы.

Методы сжатия с потерей данных применяются для архивации графической, видео и звуковой информации, т.к. данные этих типов как правило избыточны.

Важным понятием является **коэффициент сжатия** – отношение объема исходных данных, к объему сжатых.

Замена повторяющихся блоков

В процессе сжатия исходная информация подвергается анализу на наличие повторяющихся двоичных блоков. В зависимости от типа данных блок может иметь разную длину, например, для текстовой информации длина блока может быть равна 8 битам, для графической – больше, в зависимости от того, сколько бит отводится для кодировки одной точки.

Последовательность одинаковых блоков заменяется коэффициентом повторения и одним блоком.

Перед последовательностью неповторяющихся блоков ставится коэффициент неповторения, задающий количество различных блоков.

Коэффициент повторения $0N$, где N количество блоков, коэффициент неповторения – $1M$, где M количество блоков.

Например.

Исходный текст: CCCCFBDEFFFFFFF – 14 байт

Сжатый текст: 04C14FBDE06F – 9 байт

Длина блока – 1 байт.

Замена повторяющихся групп блоков.

В процессе анализа информации учитывается не только повторение отдельных блоков, но и последовательностей. Такой способ сжатия в некоторых случаях обеспечивает более высокий коэффициент сжатия, но работает медленнее.

Коэффициент повторения одиночного блока – $0N$, где N количество блоков.

Коэффициент неповторения – $1M$, где M количество блоков.

Коэффициент повторения группы блоков – LN , где L длина группы блоков, а N количество повторов группы.

Пример:

Исходный текст: ABABABDCEFFFFFFF – 14 байт

Сжатый текст: 23AB13DCE05F – 9 байт

Для дальнейшего увеличения степени сжатия можно учитывать повторы внутри группы. Это потребует еще одного коэффициента. Увеличение максимальной длины групп, исследуемых на повторяемость, приведет к снижению быстродействия и уменьшению вероятности повтора группы. Длина группы < 32 .

Коэффициент сжатия при таком методе уплотнения существенно зависит от состава данных. Возможно, что после “уплотнения” данные будут занимать не меньше, а больше места (например в текстовых файлах, где повторяющиеся символы встречаются редко).

Методы сжатия, основанные на поиске повторяющихся блоков хорошо работают на графических и табличных данных, но плохо на текстовых.

Перекодирование.

Рассмотрим использование перекодирования на примере текстовых данных.

Представление текстов в памяти компьютера заключается в том, что каждому символу ставится в соответствие его двоичный код. В существующих сейчас системах кодировки для хранения одного символа используется 1 байт (8 бит) или 2 байта (Unicode).

Пусть в файле записано только одно слово – поле.

Объем такого файла – 4 байта.

С другой стороны, для кодировки 4 различных символов необходимо всего 2 бита. Если использовать другую кодировку, то файл будет занимать на диске всего 8 бит, а не 32, т.е. экономия будет четырехкратной.

Если в тексте используются только прописные латинские буквы, то для его хранения можно также применить другой, более экономичный принцип кодирования символов. Но такой способ не подходит в том случае, если в тексте используется большое количество различных символов. (например около 200).

Другой подход основывается на том, что в текстах, написанных на естественных языках различные буквы встречаются с разной частотой.

Итак, основная идея состоит в том, чтобы для часто повторяющихся блоков данных использовать более короткие коды. При этом можно перекодировать не только отдельные блоки (в случае с текстом – буквы), но и группы блоков (слова). Решение о том, группу блоков какой длины заменять кодом, принимается на основе анализа таблиц длин кодов.

Пример.

символ	Длина кода	Заменяется вся группа целиком, т.к. сумма длин кодов трех символов по отдельности – $4+4+3=11$, а длина кода группы – 6.
A	4	
B	4	
C	3	
ABC	6	

Коды, применяемые в данном методе имеют переменную длину. Следовательно, встает вопрос о декодировании. Например, 110111 можно трактовать 11 01 11 или 1 101 11 или 1101 11...

В этом случае используется один из следующих способов:

- Длина каждого кода указывается явным способом, тогда кодовая последовательность имеет вид: L1K1 L2K2 L3K3... где L длина кода, а K – код.
- Коды составляются так, чтобы каждый код был уникальным и не являлся началом другого кода. Примером такого кодирования является кодирование Хаффмана.

Вопросы для самопроверки.

1. Для чего применяется архивация данных.
2. Проведите классификацию методов сжатия. Каковы особенности каждого метода.
3. Какие принципы лежат в основе организационной схемы резервирования данных.
4. Что представляет собой архивный файл. Какие основные операции с архивным файлом.
5. Приведите примеры программ- архиваторов.