**Государственное бюджетное профессиональное образовательное учреждение**

**«Урюпинский агропромышленный техникум»"**

Учебно-методическое пособие

Дисциплина "Информатика и ИКТ"

**«Компьютерные сети»**

Составитель: Шмыкова Галина Алексеевна

Урюпинск 2023г.

**СОДЕРЖАНИЕ**

Цель изучения темы …………………………………………...…............3

1. Компьютерные сети.……………………………………………...........4

1.1. Основные понятия и сведения............................................................4

1.2. Классификация компьютерных сетей................................................5

1.3. Основные характеристики современных компьютерных сетей…..7

1.4. Сетевое оборудование.........................................................................9

2. Локальная вычислительная сеть (ЛВС)...............................................10

2.1 Одноранговая ЛВС...............................................................................11

2.2. ЛВС с выделенным сервером.............................................................12

3. Глобальные сети. Интернет...................................................................16

3.1. История развития Интернета..............................................................16

3.2. Протоколы передачи информации.....................................................18

3.3. Адресация в Интернете.......................................................................20

3.4. Система адресации URL.....................................................................22

4. Поисковые системы...............................................................................23

4.1 Краткая история развития поисковых систем...................................23

4.2. Основные характеристики поисковой системы…………………...24

4.3. Состав и принципы работы поисковой системы……………….....26

Практическое задание...............................................................................31

Контрольные вопросы по разделу.………………………………..…....34

Список использованной литературы………...…………………....…....35



*Цель изучения темы:* ознакомление с принципом работы локальных и глобальных сетей.

*Студент должен знать:*

* что такое локальная сеть, назначение и виды ЛВС;
* принцип работы и особенности каждого вида локальной сети;
* что такое сервер в локальной сети и его функции;
* администратор сети, его назначение и функции;
* способы доступа к ресурсам сети;
* что такое глобальная сеть, история развития Интернета;
* адресация в сети Интернет;
* основные сетевые протоколы;
* принцип работы поисковых систем.

*Студент должен уметь:*

* входить в локальную сеть;
* выполнять операции с файлами и папками в локальной сети;
* открывать и закрывать доступ к папкам, изменять права доступа к сетевым ресурсам;
* просматривать ресурсы сети Интернет с помощью популярных браузеров;
* осуществлять поиск в сети Интернет.

1.**Компьютерные сети**

**1.1. Основные понятия и сведения**

***Компьютерная сеть*** (Computer NetWork, net - сеть, work - работа) представляет собой сложный комплекс взаимосвязанных и согласованно функционирующих программных и аппаратных компонентов.

***Основным назначением компьютерной сети*** является:

-совместное использование информации;

совместное использование оборудования и ПО;

централизованное администрирование и обслуживание.

***Основные компоненты компьютерной сети:***

- компьютеры (аппаратный слой);

- коммуникационное оборудование;

- сетевые операционные системы;

- сетевые приложения.

Основными компонентами сети являются рабочие станции, серверы, передающие среды (кабели) и сетевое оборудование.

***Рабочими станциями*** называются компьютеры сети, на которых пользователями сети реализуются прикладные задачи.

Сетевое приложение представляет собой распределенную программу, т. е. программу, которая состоит из нескольких взаимодействующих частей, каждая из которых выполняется на отдельном компьютере сети.

***Сервер*** – специальная программа, предназначенная для обслуживания запросов на доступ к ресурсам данного компьютера от других компьютеров сети. Модуль сервера постоянно находится в режиме ожидания запросов, поступающих по сети.

***Клиент*** - специальная программа, предназначенная для составления и посылки запросов на доступ к удаленным ресурсам, а также получения и отображения информации на компьютере пользователя.

***Сетевая служба*** - пара модулей «клиент - сервер», обеспечивающих совместный доступ пользователей к определенному типу ресурсов. Обычно сетевая операционная система поддерживает несколько видов сетевых служб для своих пользователей - файловую службу, службу печати, службу электронной почты, службу удаленного доступа и т.п. (Примеры сетевых служб – WWW, FTP, UseNet).

Термины «клиент» и «сервер» используются не только для обозначения программных модулей, но и компьютеров, подключенных к сети. Если компьютер предоставляет свои ресурсы другим компьютерам сети, то он называется сервером, а если он их потребляет - клиентом. Иногда один и тот же компьютер может одновременно играть роли и сервера, и клиента.

**1.2.Классификация компьютерных сетей**

Классифицируя сети по территориальному признаку, различают локальные (LAN), глобальные (WAN) и городские (MAN) сети.

LAN - сосредоточены на территории не более 1-2 км; построены с использованием дорогих высококачественных линий связи, которые позволяют, применяя простые методы передачи данных, достигать высоких скоростей обмена данными порядка 100 Мбит/с. Предоставляемые услуги отличаются широким разнообразием и обычно предусматривают реализацию в режиме online.

WAN - объединяют компьютеры, рассредоточенные на расстоянии сотен и тысяч километров. Часто используются уже существующие не очень качественные линии связи. Более низкие, чем в локальных сетях, скорости передачи данных (десятки килобит в секунду) ограничивают набор предоставляемых услуг передачей файлов, преимущественно не в оперативном, а в фоновом режиме, с использованием электронной почты. Для устойчивой передачи дискретных данных применяются более сложные методы и оборудование, чем в локальных сетях.

MAN - занимают промежуточное положение между локальными и глобальными сетями. При достаточно больших расстояниях между узлами (десятки километров) они обладают качественными линиями связи и высокими скоростями обмена, иногда даже более высокими, чем в классических локальных сетях. Как и в случае локальных сетей, при построении MAN уже существующие линии связи не используются, а прокладываются заново.

В зависимости от масштаба производственного подразделения, в пределах которого действует сеть, различают сети отделов, сети кампусов и корпоративные сети.

Сети отделов используются небольшой группой сотрудников в основном с целью разделения дорогостоящих периферийных устройств, приложений и данных; имеют один-два файловых сервера и не более тридцати пользователей; обычно не разделяются на подсети; создаются на основе какой-либо одной сетевой технологии; могут работать на базе одноранговых сетевых ОС.

Сети кампусов объединяют сети отделов в пределах отдельного здания или одной территории площадью в несколько квадратных километров, при этом глобальные соединения не используются. На уровне сети кампуса возникают проблемы интеграции и управления неоднородным аппаратным и программным обеспечением.

Корпоративные сети объединяют большое количество компьютеров на всех территориях отдельного предприятия. Для корпоративной сети характерны:

1. м*асштабность* - тысячи пользовательских компьютеров, сотни серверов, огромные объемы хранимых и передаваемых по линиям связи данных, множество разнообразных приложений;
2. *высокая степень гетерогенности* - типы компьютеров, коммуникационного оборудования, операционных систем и приложений различны;
3. *использование глобальных связей* - сети филиалов соединяются с помощью телекоммуникационных средств, в том числе телефонных каналов, радиоканалов, спутниковой связи.

**1.3. Основные характеристики современных компьютерных сетей**

Качество работы сети характеризуют следующие свойства: производительность, надежность, совместимость, управляемость, защищенность, расширяемость и масштабируемость.

Существуют два основных подхода к обеспечению качества работы сети. Первый - состоит в том, что сеть гарантирует пользователю соблюдение некоторой числовой величины показателя качества обслуживания. Например, сети frame relay и АТМ могут гарантировать пользователю заданный уровень пропускной способности. При втором подходе (best effort) сеть старается по возможности более качественно обслужить пользователя, но ничего при этом не гарантирует.

К основным характеристикам производительности сети относятся: время реакции, которое определяется как время между возникновением запроса к какому-либо сетевому сервису и получением ответа на него; пропускная способность, которая отражает объем данных, переданных сетью в единицу времени, и задержка передачи, которая равна интервалу между моментом поступления пакета на вход какого-либо сетевого устройства и моментом его появления на выходе этого устройства.

Для оценки надежности сетей используются различные характеристики, в том числе: коэффициент готовности, означающий долю времени, в течение которого система может быть использована; безопасность, то есть способность системы защитить данные от несанкционированного доступа; отказоустойчивость - способность системы работать в условиях отказа некоторых ее элементов.

Расширяемость означает возможность сравнительно легкого добавления отдельных элементов сети (пользователей, компьютеров, приложений, сервисов), наращивания длины сегментов сети и замены существующей аппаратуры более мощной.

Масштабируемость означает, что сеть позволяет наращивать количество узлов и протяженность связей в очень широких пределах, при этом производительность сети не ухудшается.

Прозрачность - свойство сети скрывать от пользователя детали своего внутреннего устройства, упрощая тем самым его работу в сети.

Управляемость сети подразумевает возможность централизованно контролировать состояние основных элементов сети, выявлять и разрешать проблемы, возникающие при работе сети, выполнять анализ производительности и планировать развитие сети.

Совместимость означает, что сеть способна включать в себя самое разнообразное программное и аппаратное обеспечение.

**1.4. Сетевое оборудование**

*Выделяют следующие виды сетевого оборудования:*

*1. Сетевые карты* – контроллеры, которые предназначены для приема сигналов из сети и передачи сигналов в сеть.

*1. Коммутаторы (Switch)* – управляемые программным обеспечением центральные устройства кабельной системы, сокращающие сетевой трафик за счет того, что пришедший пакет анализируется для выяснения адреса его получателя и соответственно передается только ему. Построенные с помощью коммутаторов сети могут охватывать несколько сотен машин и иметь протяженность в несколько километров.

*2. Концентраторы* – центральные устройства сети или кабельной системы, пересылающие пакеты из своего порта на все порты системы. Их разделяют на пассивные и активные. Пассивные – пропускают сигнал без усиления, а активные – усиливают и передают полученные сигналы.

3. *Маршрутизаторы (Router)* - стандартные устройства сети, работающие на сетевом уровне и позволяющие переадресовывать и маршрутизировать пакеты из одной сети в другую, а также фильтровать широковещательные сообщения.

*4. Мосты (Bridge)* - устройства сети, которое соединяют два отдельных сегмента, ограниченных своей физической длиной, и передают трафик между ними. Мосты также усиливают и конвертируют сигналы для кабеля другого типа. Это позволяет расширить максимальный размер сети, одновременно не нарушая ограничений на максимальную длину кабеля, количество подключенных устройств или количество повторителей на сетевой сегмент.

*5. Шлюзы (Gateway)* - программно-аппаратные комплексы, соединяющие разнородные сети или сетевые устройства. Шлюзы позволяет решать проблемы различия протоколов или систем адресации.

*6. Мультиплексоры* – специальные офисные устройства, которые поддерживают множество цифровых абонентских линий.

7. Повторители – устройства, которые усиливают входящий сигнал на расстояние одного сегмента сети.

8. Межсетевые экраны – некие сетевые устройства, которые контролируют поступающую и исходящую информацию сети, тем самым обеспечивая ее защиту.

**2. Локальная вычислительная сеть (ЛВС)**

*Локальная**вычислительная сеть* – это компьютеры, находящиеся в пределах одного или нескольких зданий и объединённые с помощью сетевого оборудования в единую компьютерную информационную систему.

Правильно построенная ЛВС, отвечающая современным стандартам безопасности, позволяет получать доступ к необходимой информации, обеспечивает защиту от несанкционированного доступа к данным, обеспечивая в вашем офисе стабильное информационное взаимодействие. Локально-вычислительной сети ЛВС обеспечивает следующие преимущества.

*Преимущества использования ЛВС:*

1. распределение данных (Data Sharing). Данные в ЛВС хранятся на сервере и могут быть доступны для чтения и записи на рабочих станциях пользователей;
2. совместное использование элементов сети, доступ к локальным сетевым устройствам (принтеры, сканеры, факсы и другие внешние устройства);
3. возможность быстрого доступа к необходимой информации;
4. распределение программ (Software Sharing). Все пользователи ЛВС могут совместно иметь доступ к программам, поддерживающим сетевой режим;
5. надежное хранение и резервирование данных;
6. защиту информации;
7. использование ресурсов современных технологий (доступ в Интернет, системы электронного документооборота и проч.).

ЛВС должны соответствовать следующим требованиям:

1. Локальная сеть должна быть эффективной (сочетание минимальных затрат на её построение и эксплуатацию при высоком качестве работы).
2. Длительный срок эксплуатации, оправдывающий капиталовложения.
3. Модульность и масштабируемость, возможность изменения конфигурации и наращивания без замены всей существующей сети.
4. Открытость сети, возможность подключать дополнительное оборудование при необходимости, не меняя технические и программные параметры сети.
5. Гибкость сети, при неисправностях того или иного компьютера или прочего оборудования, сеть продолжает функционировать.

ЛВС бывают одноранговыми и с выделенным сервером.

**2.1 Одноранговая ЛВС**

*Одноранговая ЛВС* *(рис.1)* – это простейшая компьютерная сеть. Она создаётся в пределах одного помещения или небольшого офиса. Как правило, число машин в такой сети не превышает 10. Все компьютеры, входящие в одноранговую сеть равноправны. Одноранговая сеть очень проста в эксплуатации и не требует администратора (специалиста по обслуживанию сети). Главная цель создания такой сети – совместное использование внешних устройств (например, принтера). Это позволяет сэкономить деньги и не приобретать лишнее оборудование. Одну из машин можно использовать как файл-сервер для хранения общих файлов или резервных копий документов с других машин.

Доступ к устройствам в одноранговой сети осуществляется, как правило, *на уровне ресурсов*, т.е. все пользователи имеют одинаковые права доступа к сетевым ресурсам.



Рис. 1. Одноранговая сеть

Таким образом, одноранговая сеть – это очень экономичное (дешёвое) решение многих проблем малого офиса, позволяющее простейшим способом организовать совместную работу небольшого коллектива людей.

**2.2. ЛВС с выделенным сервером**

*Л**ВС с выделенным сервером (рис.2)* – это более сложная сеть. Она очень хорошо масштабируется до больших размеров. Такие сети используются на крупных и средних предприятиях. На её базе создаётся сложная распределённая информационная система в пределах предприятия (фирмы).

Рис. 2. ЛВС с выделенным сервером Топология локальных сетей

Под топологией (компоновкой, конфигурацией, структурой) компьютерной сети обычно понимается физическое расположение компьютеров сети друг относительно друга и способ соединения их линиями связи. Важно отметить, что понятие топологии = относится, прежде всего, к локальным сетям, в которых структуру связей можно легко проследить. В глобальных сетях структура связей обычно скрыта от пользователей и не слишком важна, так как каждый сеанс связи может производиться по собственному пути.

Топология определяет требования к оборудованию, тип используемого кабеля, допустимые и наиболее удобные методы управления обменом, надежность работы, возможности расширения сети. И хотя выбирать топологию пользователю сети приходится нечасто, знать об особенностях основных топологий, их достоинствах и недостатках надо.

Существует три базовые топологии сети локальных сетей:

**Шина** (bus) — все компьютеры параллельно подключаются к одной линии связи. Информация от каждого компьютера одновременно передается всем остальным компьютерам (рис. 3).


Рис.3. Сетевая топология «Шина»

В топологии шина отсутствует явно выраженный центральный абонент, через который передается вся информация, это увеличивает ее надежность (ведь при отказе центра перестает функционировать вся управляемая им система). Добавление новых абонентов в шину довольно просто и обычно возможно даже во время работы сети. В большинстве случаев при использовании шины требуется минимальное количество соединительного кабеля по сравнению с другими топологиями.

**Звезда** (star) — к одному центральному компьютеру присоединяются остальные периферийные компьютеры, причем каждый из них использует отдельную линию связи. Информация от периферийного компьютера передается только центральному компьютеру, от центрального — одному или нескольким периферийным (рис. 4).


Рис. 4. Сетевая топология «Звезда»

Звезда — это единственная топология сети с явно выделенным центром, к которому подключаются все остальные абоненты.

В отличие от шины, в звезде на каждой линии связи находятся только два абонента: центральный и один из периферийных. Чаще всего для их соединения используется две линии связи, каждая из которых передает информацию в одном направлении, то есть на каждой линии связи имеется только один приемник и один передатчик.

Если говорить об устойчивости звезды к отказам компьютеров, то выход из строя периферийного компьютера или его сетевого оборудования никак не отражается на функционировании оставшейся части сети, зато любой отказ центрального компьютера делает сеть полностью неработоспособной. В связи с этим должны приниматься специальные меры по повышению надежности центрального компьютера и его сетевой аппаратуры.

Недостатком данной топологии является значительно больший, чем при других топологиях, расход кабеля. Например, если компьютеры расположены в одну линию, то при выборе топологии звезда понадобится в несколько раз больше кабеля, чем при топологии шина.

**Кольцо** (ring) — компьютеры последовательно объединены в кольцо. Передача информации в кольце всегда производится только в одном направлении. Каждый из компьютеров передает информацию только одному компьютеру, следующему в цепочке за ним, а получает информацию только от предыдущего в цепочке компьютера (рис. 5).


Рис. 5. Сетевая топология «Кольцо»

Четко выделенного центра при кольцевой топологии нет, все компьютеры могут быть одинаковыми и равноправными. Однако довольно часто в кольце выделяется специальный абонент, который управляет обменом или контролирует его. Понятно, что наличие такого единственного управляющего абонента снижает надежность сети, так как выход его из строя сразу же парализует весь обмен.

**3. Глобальные сети. Интернет**

**3.1. История развития Интернета**

*Глобальная сеть* — это множество компьютеров, расположенных на большом расстоянии друг от друга и имеющих возможность связываться между собой при помощи сервера по каналам связи.

Первоначально глобальные сети решали задачу доступа удаленных ЭВМ и терминалов к мощным ЭВМ, которые назывались host-компьютер (часто используют термин сервер). Такие подключения осуществлялись через коммутируемые или некоммутируемые каналы телефонных сетей или через спутниковые выделенные сети передачи данных

Слово *Интернет* (*Internet*) происходит от словосочетания Interconnected networks (связанные сети), это глобальное сообщество малых и больших сетей.

Internet является старейшей глобальной сетью. Internet предоставляет различные способы взаимодействия удаленных компьютеров и совместного использования распределенных услуг и информационных ресурсов.

Зарождением Интернета принято считать момент появления первой компьютерной сети, родиной которой в середине 60-х годов двадцатого века стала Америка.

В то время еще не существовало персональных компьютеров, и крупные американские университеты могли себе позволить 1–2 больших компьютера. Когда появилась идея соединить между собой компьютеры разных университетов, был создан проект, который получил название ARPANET. К концу 1969 года были соединены компьютеры четырех университетов и появилась первая компьютерная сеть.

В середине 70-х годов для ARPANET были разработаны новые стандарты передачи данных, которые позволяли объединять сети произвольной архитектуры, тогда же было придумано слово «Интернет».

Настоящий расцвет Интернета начался в 1992 году, когда была изобретена новая служба, получившая странное название «Всемирная паутина» (World Wide Web, или WWW, или просто «веб»). WWW позволял любому пользователю Интернета публиковать свои текстовые и графические материалы в привлекательной форме, связывая их с публикациями других авторов и предоставляя удобную систему навигации. Постепенно Интернет начал выходить за рамки академических институтов и стал превращаться из средства переписки и обмена файлами в гигантское хранилище информации. К 1992 году Интернет насчитывал более миллиона соединенных компьютеров.

В Россию Интернет впервые проник в начале 90-х годов. Ряд университетов и исследовательских институтов приступили в это время к построению своих компьютерных сетей и обзавелись зарубежными каналами связи.

**3.2. Протоколы передачи информации**

Для того чтобы при обмене данными компьютеры, объединенные в сеть, действовали согласованно, разработан ряд стандартов и правил, называемых *протоколами*.

Основополагающим протоколом сети Internet является **протокол TCP/IP**. (Transmission Control Protocol/Internet Protocol – Протокол управления передачей/Межсетевой протокол) - протокол управления передачей. Он определяет, каким образом информация должна быть разбита на пакеты и отправлена по каналам связи. TCP располагает пакеты в нужном порядке, а также проверяет каждый пакет на наличие ошибок при передаче.

Протоколы, входящие в семейство TCP/IP разделяются на уровни.

*Физический уровень* описывает среду передачи данных (будь то кабель, оптоволокно или радиоканал), физические характеристики такой среды и принцип передачи данных (разделение каналов, модуляцию, амплитуду сигналов, частоту сигналов, способ синхронизации передачи, время ожидания ответа и максимальное расстояние).

*Канальный уровень* описывает, каким образом передаются пакеты данных через физический уровень, включая кодирование (т.е. специальные последовательности битов, определяющих начало и конец пакета данных). Примеры протоколов канального уровня — Ethernet, IEEE 802.11 Wireless Ethernet, SLIP, Token Ring, ATM и MPLS.

*Сетевой уровень* изначально разработан для передачи данных из одной сети (подсети) в другую.

Протоколы *транспортного уровня* могут решать проблему гарантированной доставки сообщений («дошло ли сообщение до адресата?»), а также гарантировать правильную последовательность прихода данных. Транспортные протоколы определяют, для какого именно приложения предназначены эти данные.

На *прикладном уровне* работает большинство сетевых приложений. Эти программы имеют свои собственные протоколы обмена информацией.

Основные протоколы:

Протокол телеконференций – *News Net Transfer Protocol* (**NNTP)**

Протокол получения электронных писем - *Post Office Protocol 3*(**POP3)**

Протокол отправки электронных писем – *Simple Mail Transfer Protocol* (**SMTP)**

Протокол передачи файлов – *File Transfer Protocol* (**FTP)**

Протокол передачи гипертекста – *Hyper Text Transfer Protocol* (**HTTP)**

**3.3. Адресация в Интернете**

Каждый компьютер, подключенный к сети Internet, имеет свой уникальный **IP-адрес**.

**IP-адрес** — это уникальный номер, однозначно идентифицирующий компьютер в Internet. IP-адрес представляет собой четыре числа (октета), разделенные точками, например, 1**94.67.67.97** (после последнего числа точка не ставится).

Каждое число может быть в интервале от 0 до 255, что соответствует информационному объему в 1 байт или 8 бит. Таким образом, IP-адрес – это 4 байта или 32 бита.

Но, для человека такая система адресации сложна, так же как нам сложно помнить, набирать и диктовать одиннадцатизначные телефонные номера, поэтому в 1984 году Полом Мокапетрисом была разработана *надстройка над IP-адресацией*, называемая системой *DNS (domain name system, система доменных имен)*.

*DNS - система доменных имен*

Cистема запросов в сети Интернет, позволяющая получать информацию о соответствии адресов и имен по сети называется **доменной системой имен** - DNS (Domain Name System). Имена доменов отделяются друг от друга точками.

Домены *первого уровня* разделяются на *тематические* и *географические (табл.1)*.

Таблица 1. Тематические домены 1-ого уровня

|  |  |
| --- | --- |
| COM | Commercial (для коммерческих организаций) |
| NET | Networks (Интернет, телекоммуникационные сети) |
| ORG | Organizations (некоммерческие организации либо организации, не попадающие в другие категории) |
| INFO | Information (открытый для всех домен) |
| BIZ | Business Organizations (аналог com) |
| NAME | Personal (для частных лиц) |
| INT | International Organizations (международные организации) |
| EDU | Educational (образовательные проекты США) |
| MIL | US Dept of Defense (департамент безопасности США) |
| GOV | US Government (правительство США) |
| MUSEUM | Museums (музеи) |
| AERO | Air-transport industry (воздушно-транспортная индустрия) |
| COOP | Cooperatives (кооперативы) |

Территориальные домены первого уровня, в отличие от тематических, всегда двухбуквенные.

**3.4. Система адресации URL**

URL был изобретён Тимом Бернерсом-Ли в 1990 году в стенах Европейского совета по ядерным исследованиям (фр. Conseil Européen pour la Recherche Nucléaire, CERN) в Женеве, Швейцария. URL стал фундаментальной инновацией в Интернете. Изначально URL предназначался для обозначения мест расположения ресурсов (чаще всего файлов) во Всемирной паутине. Сейчас URL применяется для обозначения адресов почти всех ресурсов Интернета.

Стандарт URL закреплён в документе RFC 1738, прежняя версия была определена в RFC 1630. Сейчас URL позиционируется как часть более общей системы идентификации ресурсов URI, сам термин URL постепенно уступает место более широкому термину URI. Стандарт URL регулируется организацией IETF и её подразделениями.

Чтобы найти в Интернете какой-либо документ, достаточно знать *ссылку* на него - так называемый у*ниверсальный указатель ресурса* (URL - Uniform Resource Locator), который определяет местонахождение каждого файла, хранящегося на компьютере, подключенном к Интернету.

Адрес URL является сетевым расширением понятия полного имени ресурса в операционной системе. В URL, кроме имени файла и директории, где он находится, указывается сетевое имя компьютера, на котором этот ресурс расположен, и протокол доступа к ресурсу, который можно использовать для обращения к нему. Система адресации URL и адресация почтовой службы имеют сходную структуру.

**4. Поисковые системы**

*Поисковая* *система* — это программно-аппаратный комплекс, предназначенный для осуществления поиска в сети Интернет и реагирующий на запрос пользователя, задаваемый в виде текстовой фразы (поискового запроса), выдачей списка ссылок на источники информации, в порядке релевантности (в соответствии запросу).

**4.1 Краткая история развития поисковых систем**

В начальный период развития Интернет, число его пользователей было невелико, а объем доступной информации сравнительно небольшим. В большинстве своем, доступ к сети Интернет имели лишь сотрудники научно-исследовательской сферы. В это время задача поиска информации в Интернете не была столь актуальной, как в настоящее время.

Одним из первых способов организации доступа к информационным ресурсам сети стало создание открытых каталогов сайтов, ссылки на ресурсы, в которых группировались согласно тематике. Первым таким проектом стал сайт Yahoo.com, открывшийся весной 1994 года. После того, как количество сайтов в каталоге Yahoo значительно увеличилось, была добавлена возможность поиска нужной информации по каталогу. В полном смысле это еще не было поисковой системой, так как поисковая область была ограничена только ресурсами, присутствующими в каталоге, а не всеми Интернет-ресурсами.

Каталоги ссылок широко использовались ранее, однако практически полностью утратили свою популярность в настоящее время.

Первой полноценной поисковой системой стал проект WebCrawler, вышедший в свет в 1994 году.

В 1995 году появились поисковые системы Lycos и AltaVista. Последняя долгие годы была лидером в области поиска информации в сети Интернет.

В 1997 году Сергей Брин и Ларри Пейдж создали поисковую машину Google в рамках исследовательского проекта в Стэндфордском университете. В настоящий момент Google - самая популярная поисковая система в мире!

В сентябре 1997 года была официально анонсирована поисковая система Yandex, являющаяся самой популярной в русскоязычном Интернете.

В настоящее время существуют три основные поисковые системы (международные) – Google, Yahoo и MSN, имеющие собственные базы и алгоритмы поиска. Большинство остальных поисковых систем (коих насчитывается большое количество) использует в том или ином виде результаты трех перечисленных. Например, поиск AOL (search.aol.com) использует базу Google, а AltaVista, Lycos и AllTheWeb – базу Yahoo.

**4.2. Основные характеристики поисковой системы**

К основным характеристика поисковой системы относятся:

*1. Полнота*

Полнота - представляет собой отношение количества найденных по запросу документов к общему числу документов в сети Интернет, удовлетворяющих данному запросу. К примеру, если в Интернете имеется 100 страниц, содержащих словосочетание «как выбрать автомобиль», а по соответствующему запросу было найдено всего 60 из них, то полнота поиска будет 0,6. Очевидно, что чем полнее поиск, тем меньше вероятность того, что пользователь не найдет нужный ему документ, при условии, что он вообще существует в Интернете.

*2. Точность*

Точность - определяется степенью соответствия найденных документов запросу пользователя. Например, если по запросу «как выбрать автомобиль» находится 100 документов, в 50 из них содержится словосочетание «как выбрать автомобиль», а в остальных просто наличествуют эти слова («как правильно выбрать магнитолу и установить в автомобиль»), то точность поиска считается равной 50/100 (=0,5). Чем точнее поиск, тем быстрее пользователь найдет нужные ему документы, тем меньше различного рода «мусора» среди них будет встречаться, тем реже найденные документы не будут соответствовать запросу.

*3. Актуальность*

Актуальность - характеризуется временем, проходящим с момента публикации документов в сети Интернет, до занесения их в индексную базу поисковой системы. Например, на следующий день после появления интересной новости, большое количество пользователей обратились к поисковым системам с соответствующими запросами. Объективно с момента публикации новостной информации на эту тему прошло меньше суток, однако основные документы уже были проиндексированы и доступны для поиска, благодаря существованию у крупных поисковых систем так называемой «быстрой базы», которая обновляется несколько раз в день.

*4. Скорость поиска*

Скорость поиска тесно связана с его устойчивостью к нагрузкам. Поисковая машина должна отрабатывать запрос максимально оперативно, чтобы не тормозить вычисление следующих запросов.

*5. Наглядность*

Наглядность представления результатов является важным компонентом удобного поиска. По большинству запросов поисковая машина находит сотни, а то и тысячи документов. Вследствие нечеткости составления запросов или неточности поиска, даже первые страницы выдачи не всегда содержат только нужную информацию. Это означает, что пользователю зачастую приходится производить свой собственный поиск внутри найденного списка. Различные элементы страницы выдачи поисковой системы помогают ориентироваться в результатах поиска. подробные пояснения по странице результатов поиска.

**4.3. Состав и принципы работы поисковой системы**

В России основной поисковой системой является «Яндекс», далее - Rambler.ru, Google.ru, Aport.ru, Mail.ru. Причем, на данный момент, Mail.ru использует механизм и базу поиска «Яндекса».

Практически все крупные поисковые системы имеют свою собственную структуру, отличную от других. Однако можно выделить общие для всех поисковых машин основные компоненты. Различия в структуре могут быть лишь в виде реализации механизмов взаимодействия этих компонентов.

Поисковые системы обычно состоят из трех компонентов:

* модуль индексирования
* база данных
* поисковый сервер

*Модуль индексирования*

Модуль индексирования состоит из трех вспомогательных программ (роботов):

* Агент
* Кроулер
* Индексатор

Агент – программа, предназначенная для скачивания веб-страниц. «Паук» обеспечивает скачивание страницы и извлекает все внутренние ссылки с этой страницы. Скачивается html-код каждой страницы. Для скачивания страниц роботы используют протоколы HTTP. Работает «паук» следующим образом. Робот на сервер передает запрос “get/path/document” и некоторые другие команды HTTP-запроса. В ответ робот получает текстовый поток, содержащий служебную информацию и непосредственно сам документ.

Ссылки извлекаются из тэгов a, area, base, frame, frameset, и др. Наряду со ссылками, многими роботами обрабатываются редиректы (перенаправления). Каждая скачанная страница сохраняется в следующем формате:

* URL страницы
* дата, когда страница была скачана
* http-заголовок ответа сервера
* тело страницы (html-код)

Crawler (кроулер)– программа, которая автоматически проходит по всем ссылкам, найденным на странице. Выделяет все ссылки, присутствующие на странице. Его задача - определить, куда дальше должен идти паук, основываясь на ссылках или исходя из заранее заданного списка адресов. Crawler, следуя по найденным ссылкам, осуществляет поиск новых документов, еще неизвестных поисковой системе.

Indexer (робот- индексатор) - программа, которая анализирует веб-страницы, скаченные пауками. Индексатор разбирает страницу на составные части и анализирует их, применяя собственные лексические и морфологические алгоритмы. Анализу подвергаются различные элементы страницы, такие как текст, заголовки, ссылки структурные и стилевые особенности, специальные служебные html-теги и т.д.

Таким образом, модуль индексирования позволяет обходить по ссылкам заданное множество ресурсов, скачивать встречающиеся страницы, извлекать ссылки на новые страницы из получаемых документов и производить полный анализ этих документов.

*База данных*

База данных, или индекс поисковой системы — это система хранения данных, информационный массив, в котором хранятся специальным образом преобразованные параметры всех скачанных и обработанных модулем индексирования документов.

База данных отыскивает предмет запроса, основанный на информации, указанной в заполненной форме, и выводит соответствующие документы, подготовленные базой данных. Чтобы определить порядок, в котором список документов будет показан, база данных применяет алгоритм ранжирования. В идеальном случае, документы, наиболее релевантные пользовательскому запросу будут помещены первыми в списке. Различные поисковые системы используют различные алгоритмы ранжирования, однако основные принципы определения релевантности, следующие:

1. Количество слов запроса в текстовом содержимом документа (т.е. в html-коде).

2. Тэги, в которых эти слова располагаются.

3. Местоположение искомых слов в документе.

4. Удельный вес слов, относительно которых определяется релевантность, в общем количестве слов документа.

*Поисковый сервер*

Поисковый сервер является важнейшим элементом всей системы, так как от алгоритмов, которые лежат в основе ее функционирования, напрямую зависит качество и скорость поиска.

Поисковый сервер работает следующим образом:

* Полученный от пользователя запрос подвергается морфологическому анализу. Генерируется информационное окружение каждого документа, содержащегося в базе (которое и будет впоследствии отображено в виде сниппета, то есть соответствующей запросу текстовой информации на странице выдачи результатов поиска).
* Полученные данные передаются в качестве входных параметров специальному модулю ранжирования. Происходит обработка данных по всем документам, в результате чего, для каждого документа рассчитывается собственный рейтинг, характеризующий релевантность запроса, введенного пользователем, и различных составляющих этого документа, хранящихся в индексе поисковой системы.
* В зависимости от выбора пользователя этот рейтинг может быть скорректирован дополнительными условиями (например, так называемый «расширенный поиск»).
* Далее генерируется сниппет, то есть, для каждого найденного документа из таблицы документов извлекаются заголовок, краткая аннотация, наиболее соответствующая запросу и ссылка на сам документ, причем найденные слова подсвечиваются.
* Полученные результаты поиска передаются пользователю в виде SERP (Search Engine Result Page) – страницы выдачи поисковых результатов.

**ПРАКТИЧЕСКОЕ ЗАДАНИЕ**

***Создание аннотированного каталога медицинских ресурсов***

***Цель работы:*** создание подробного упорядоченного каталога медицинских сайтов.

*Каталог* — это коллекция аннотированных ссылок на сайты, объединенных общей тематикой.

*Цель каталога* — предложить пользователю представительную подборку

сайтов, чтобы облегчить поиск нужной информации.

*Порядок выполнения работы:*

1. Запустите любой доступный веб-обозреватель (Internet Explorer, Google

Chrome, Mozilla Firefox, Opera).

2. Воспользовавшись помощью поисковых систем (Google, Яндекс, Rambler и т.д.), найдите всевозможные медицинские интернет-ресурсы.

3. Подберите 10 наиболее интересных медицинских ресурса.

4. Создайте аннотированный каталог по схеме:

* Название сайта
* Электронный адрес ресурса
* Скриншот главной страницы
* Описание (на чем профилируется сайт, краткая характеристика и возможные сведения (город, информация о создателе сайта, и т.д.))
* Аннотация

Аннотация должна содержать следующие сведения:

* какую информацию содержит сайт (названия основных разделов и тематик)
* на какую аудиторию рассчитан
* полезен ли данный ресурс? Чем?
* удобство (особенности навигации)

*Пример:*

Название сайта: Русский Медицинский Журнал

Адрес ресурса: http://www.rmj.ru



Описание: первое независимое электронное издание для практикующих врачей. Самое читаемое издание в России. Современная полная информация о диагностике и лечении заболеваний для врачей всех специальностей.

Аннотация: на сайте размещены полные тексты журнала, архивы, каталоги статей, конференции. Имеется так же форум для врачей. Аудитория сайта — врачи общей практики, узкие специалисты, клинические ординаторы и аспиранты, студенты старших курсов медицинских институтов. Материал сайта полный, наглядный, хорошо структурирован и будет полезен для врачей всех специальностей. Так же существенным плюсом является хорошо налаженная навигация, наличие карты сайта и расширенного поиска.

Тестовое задание №1

Критерии оценивания:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Оценка | Отлично | Хорошо | Удовлетворительно | Неудовлетворительно |
| Кол-во верных ответов | 15-18 | 14-11 | 10-5 | 4-0 |

Тест №1

* 1. **Что такое информация?**

 1.Это смысл, содержащийся в объектах окружающей действительности

 2.Книжный фонд библиотеки

 3.Последовательность знаков некоторого алфавита

* 1. **Информацию, важную и существенную в настоящий момент, называют:**

 1.Полезной

 2.Актуальной

 3.Достоверной

 4.Объективной

* 1. **Сколько в 1 Байте содержится Бит?**

 1.1024 бит

 2.8 бит

 3.16 бит

* 1. **Что входит в состав МФУ (многофункциональное устройство)?**

 1.Принтер, сканер, факс

 2.Монитор, мышь, клавиатура

 3.Оперативная память, жесткий диск, центральный процессор

* 1. **Что такое система счисления?**

 1.Знаковая система, в которой приняты определенные правила записи чисел.

 2.Калькулятор

 3.Система, которая помогает счислять информацию

* 1. **Какие системы счисления выделяются?**

 1.Пятикратная, десятикратная, пятнадцатикратная

 2.Мировая, городская, сельская

 3.Двоичная, восьмеричная, шестнадцатеричная

* 1. **По форме представления информацию можно условно разделить на следующие виды:**

 1.Математическую, биологическую, психологическую

 2.Текстовую, числовую, графическую, звуковую

 3.Обыденную, научную, производственную

* 1. **В какой строке единицы измерения информации расположены по возрастанию?**

 1.Гигабайт, мегабайт, килобайт, байт, бит

 2.Бит, байт, мегабайт, килобайт, гигабайт

 3.Бит, байт, килобайт, мегабайт, гигабайт

* 1. **Информационные процессы – это:**

 1.Процессы строительства зданий

 2.Процессы химической и механической очистки воды

 3.Процессы сбора, хранения, обработки, поиска и передачи информации

* 1. **В какой строке единицы измерения информации расположены по убыванию?**

 1.Гигабайт, мегабайт, килобайт, байт, бит

 2.Бит, байт, мегабайт, килобайт, гигабайт

 3.Байт, бит, килобайт, мегабайт, гигабайт

* 1. **Под носителем информации принято подразумевать:**

 1.Линию связи

 2.Материальный объект, на котором можно тем или иным способом зафиксировать информацию

 3.Интернет

* 1. **Информацию, взятую из надежного источника, называют:**

 1.Полезной

 2.Актуальной

 3.Достоверной

* 1. **Из таблицы истинности что означает операция →?**

 1.Отрицание

 2.Дизъюнкция

 3.Следование

* 1. **Минимальная единица измерения информации:**

 1.Байт

 2.Бит

 3.Мбайт

* 1. **Перевод текста с английского языка на русский является процессом:**

 1.Передачи информации

 2.Обработки информации

 3.Поиска информации

* 1. **Какое из устройств компьютера не входит в состав системного блока?**

 1.Процессор

 2.Принтер

 3.Дисковод

* 1. **Для ввода звуковой информации служит:**

 1.Монитор

 2.Колонки

 3.Мышь

* 1. **С клавиатуры команда Вставить заменяется комбинацией клавиш:**

 1.Ctrl + V

 2.Ctrl + A

 3.Ctrl + C

Правильные ответы:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* | *12* | *13* | *14* | *15* | *16* | *17* | *18* |
| *1* | *2* | *2* | *1* | *1* | *3* | *2* | *3* | *3* | *1* | *2* | *3* | *3* | *2* | *2* | *2* | *2* | *1* |

**Контрольные вопросы**

1. Что относится к компьютерным сетям?

2. Что такое рабочая станция, сервер, клиент?

3. Как классифицируются компьютерные сети?

4. Что относится к сетевому оборудованию?

5. Что представляют собой локальные сети?

6. Какие существуют преимущества локальных сетей?

7. Чем отличается одноранговая ЛВС от ЛВС с выделенным сервером?

8. Какие существуют базовые топологии ЛВС?

9. Что такое глобальная сеть?

10. Как возникло слово "Интернет"?

11. Что такое "Всемирная паутина"?

12. Какие существуют протоколы передачи информации?

13. Что такое IP-адрес?

14. Как классифицируются домены первого уровня?

15. Что представляют собой поисковые системы?

16. Какие существуют основные характеристики поисковых систем?

17. Из каких компонентов состоит поисковая система?

18.Что такое агент, кроулер, индексатор? Какие функции они выполняют?

19. Какие существуют принципы определения релевантности?

20. Как работает поисковый сервер?

**Список использованной литературы**

1. Н. Угринович. “Информатика и информационные технологии. 10-11 классы”, М., ЛБЗ, 2013 г.
2. С. Симонович и др. “Общая информатика”, М., АСТпресс, 2012 г.
3. Шафрин Ю., Основы компьютерной технологии- М.: АБФ, 2013 г.
4. В. Т. Безручко “Практикум по курсу “Информатика”, М., Финансы и статистика, 2012 г.
5. С.В.Назаров. “Компьютерные технологии обработки информации”, М., АСТпресс, 2013 г.
6. http://www.pedsovet.info/ - методическое объединение преподавателей
7. http://innov.ncic.ru/bases/m.html - Образовательные инновации. Учебно-методические разработки.