

Раздел 4. Компьютерные вычислительные сети и сетевые технологии обработки информации

Информационно-поисковые системы

Информационная вычислительная сеть – это совокупность технических, коммуникационных, программных средств, обеспечивающих эффективное использование ресурсов. Возможность построения таких сетей обеспечена наличием мощных телефонных сетей и развитием вычислительной техники, снижением ее стоимости. Ресурсы были объединены.

Вычислительные информационные сети позволяют:

- Построить мощные хранилища информации;
- Расширить перечень решаемых задач;
- Обеспечивают новые виды сервисного обслуживания (электронная почта);
- Снизить себестоимость обработки информации;

По охвату территории информационно-вычислительные сети подразделяются на:

- Локальные;
- Региональные;
- Федеральные;
- Глобальные.

Локальные сети объединяют компьютеры, расположенные недалеко – до 10 км. Локальная сеть обеспечивает совместное использование данных и оборудования, т.е. ресурсов. Различают следующие виды конфигурации локальных сетей: шина – все компьютеры подключены к одному кабелю; кольцо; звезда – при наличии главного компьютера – сервера; снежинка – объединение нескольких звезд.

Сейчас широкое распространение получили глобальные вычислительные сети. Глобальные вычислительные сети – это совокупность компьютеров и коммуникаций. Глобальная сеть – это телекоммуникационная структура, объединяющая локальные информационные сети, имеющие общий протокол связи, методы подключения и протоколы обмена. Основная задача глобальной сети – обеспечение доступа к сети и взаимодействие абонентов, рассредоточенных на большой территории. Крупнейшей является глобальная сеть Internet.

Internet появился в результате объединения сети Министерства обороны и спутниковой сети в США. Эта сеть была прародительницей Internet. Вначале услуги сети были дороги и пользоваться ими могли только крупные организации. Сейчас сеть доступна всем желающим. Спрос на услуги сети растет. В США проводится работа по подключению к сети начальной и средней школы. У нас подключение к сети стало возможным с понижением цен на компьютеры. Первые сети были реализованы в 60-х годах в США и СССР (сеть Академии наук). В 1961 г. по заданию министерства обороны США была создана сеть – прародительница Интернет. В 1975 г. эта сеть стала рабочей и получила дальнейшее развитие.

По способу передачи информации различают следующие виды сетей:

1. **Спутниковая связь.** Большая пропускная способность, большие расстояния, высокая надежность, но очень дорога. На нее приходится 3% всего трафика – объем информации.
2. **Оптоволоконная связь.** Достоинства и недостатки те же, но быстрое развитие световодов ведет к снижению стоимости.
3. **Радиосвязь** используется реже из-за помех.
4. **Модемная связь.**

Передача данных организована на основе протокола – правила обмена данными. По сети передается и получается любая информация (программы, тексты, картинки, погода, электронные версии газет). В сети подключены компьютеры различных типов. Для общения выработан единый стандарт (протокол управления пересылками/протокол Internet). В сети **Internet** каждая машина имеет адрес, по которому к ней осуществляется доступ. Адреса присваиваются в виде набора цифр, но для пользователей существует доменная система имен. Специальная программа переводит доменные имена в цифровой вид.

В основе сети лежат миллионы компьютеров, соединенных постоянно (их называют **серверами**). Их владельцами могут быть государственные и коммерческие организации, частные лица. Каждый сервер имеет **имя**, которое покупают за определенную сумму. Если к серверу подключены другие компьютеры, то он становится **провайдером**. Миллионы серверов объединяют десятки миллионов пользователей, они являются узлами, через которые информация передается или хранится. Сервером может быть только быстродействующая машина с жестким диском большого объема. Функции сервера:

1. **Хранение информации**: клиент может создать на жестком диске сервера свою папку и хранить там свою информацию, организуется виртуальный сервер.
2. **Передача информации**: сервер является частью сети и через него передается информация, если он находится на ее пути. Для выбора маршрута существуют специальные средства - **маршрутизаторы**.

В **Internet** существуют различные службы:

1. WWW – всемирная паутина;
2. FTP – файловый серверы;
3. Электронная почта.

Самой распространенной службой является **всемирная паутина**. Пользователи иногда считают, что это и есть **Internet**, но это одна из служб. Основные понятия – Web-страница и Web-документ. Это документ в специальном формате, он выглядит как обычный текст, но содержит специальные коды. Для просмотра Web-документов существуют специальные программы браузеры. В современных компьютерах устанавливаются версии ПО,

располагающие встроенными браузерами, обычно – Internet Explorer. В данной службе существует такое мощное средство, как **гипертекст**. Это ссылки на другие документы, текстовые и графические. В Web-документе это подчеркнутый текст или картинка в рамке. Если щелкнуть по такому объекту, то загружается документ, который указан в ссылке. Гипертекст облегчает работу в сети, благодаря этому можно путешествовать от одного документа к другому.

Служба пересылки файлов. На FTP-серверах хранятся файловые архивы. В них содержатся тексты программ, исполнимые файлы и т.д. Попадая в область FTP-сервера, можно получить доступ к библиотекам университетам, можно запросить справки, документы, учебники, техническую информацию. Работать с FTP-архивами можно из всемирной паутины, задавая адрес с **ftp://** .

Электронная почта. В **Internet** электронная почта является важным информационным ресурсом. Это технология оперативной связи между клиентами, является услугой вычислительных сетей. Используется для автоматизации конторских работ, для безбумажных почтовых отношений. Использование электронной почты предполагает передачу сообщений через специальные «почтовые ящики». Это область памяти сети, специальный файл для хранения почты. Он состоит из 2-х корзин: отправления и получения.

Преимущества Эл.почты : скорость обмена, относительно низкая стоимость, оперативность. Она похожа на обычную, письмо готовится на рабочем столе с помощью программы подготовки почты или с помощью текстового редактора. Затем пользователь вызывает программу отправки почты, которая отсылает сообщение на почтовый сервер адресата. Там специальная программа сортирует и рассылает по ящикам конечных пользователей. Почтовые серверы постоянно подключены к сети, а клиенты подключаются время от времени. У Эл.почты существует свой протокол. Основа почты – система адресов. Для доставки почты нужен точный адрес. Он состоит из 2-х частей, разделяемых знаком **@** . Например: SAutoKol@sci.smolensk.ru.

Программные поисковые сервисы. Информационно-поисковые системы.

Современные условия деятельности требуют постоянного анализа нормативно-технической и правовой литературы. Значительное время работники тратят на поиски и обработку специальной информации. Для поиска информации в постоянно увеличивающемся информационном пространстве могут использоваться две принципиально разные информационно-поисковые системы (ИПС): информационно-поисковые системы для глобальной сети и справочно-правовые системы (СПС). Обе системы развиваются и функционируют независимо друг от друга. Совместное использование этих систем позволяет быстро и качественно решить задачу поиска информации при решении широкого круга задач.

Информационно-поисковые системы появились на свет достаточно давно, в конце 70-х - начале 80-х годов. Создание систем хранения информации осуществляется двумя основными способами: с использованием иерархической и гипертекстовой моделей. В иерархической модели используется многоуровневая рубрикация при классификации информации. Для поиска документа используется его краткое описание, составляемое при вводе информации в систему. Современная гипертекстовая модель позволяет в электронных документах использовать ссылки на другие документы.

Информационно-поисковая система (ИПС) — это система, обеспечивающая поиск и отбор необходимых данных в специальной базе с описаниями источников информации (индексе) на основе информационно-поискового языка и соответствующих правил поиска. Главной задачей любой ИПС является поиск информации релевантной информационным потребностям пользователя. Очень важно в результате проведенного поиска ничего не потерять, то есть найти все документы, относящиеся к запросу, и не найти ничего лишнего. Поэтому вводится качественная характеристика процедуры поиска — релевантность.

Релевантность — это соответствие результатов поиска сформулированному запросу. Релевантность можно представить как меру близости между реально полученными документами и тем, что следовало бы получить из системы.

Типы ИПС для глобальной сети

1. **Каталог** — это сайт с собранными ссылками на сайты, которые располагаются по тематическим разделам. Каталоги в интернет являются древовидной структурой. При заходе на титульный лист, возникает перечень разделов и тем, подразделов.

Поиск в каталоге удобен и проводится посредством последовательного уточнения тем. Тем не менее, каталоги поддерживают возможность быстрого поиска определенной категории или страницы по ключевым словам с помощью локальной поисковой машины. База данных ссылок (индекс) каталога обычно имеет ограниченный объем, заполняется вручную персоналом каталога. Некоторые каталоги используют автоматическое обновление индекса. Результат поиска в каталоге представляется в виде списка, состоящего из краткого описания (аннотации) документов с гипертекстовой ссылкой на первоисточник. Недостатками каталогов является наличие устаревшей информации; пользователю нужна конкретная информация, а не сайт.

Примеры популярных каталогов:

1. **Yahoo**. Он содержит более миллиона ресурсов и поддерживается 150 редакторами.
2. **List.Ru** (www.list.ru): 100 тысяч ссылок, классифицированных по 18 каталогам.
3. **Mail.ru**. Каталог содержит более 300 тыс. ссылок.



2. **Поисковая машина** — поисковая система с базой данных, содержащей информацию об информационных ресурсах. База данных, содержащая информацию об Web-страницах, статьях, формируется программой-роботом. Поиск в такой системе проводится по запросу, состоящему из набора ключевых слов или фразы, заключенной в кавычки. Зарубежные поисковые машины: Google, Altavista, Российские поисковые машины: Yandex, Рамблер.

Процесс работы (который не заметен пользователю) и внутреннее устройство:

Состав:

1. **Паук** - программа, которая, скачивает найденные в интернете файлы в базу данных.
2. **Индексатор** — программа, анализирующая скаченную информацию пауком.
3. **Веб-сервер** — сайт, к которому обращается пользователь за поиском.
4. **Паук «путешественник»** — программа, путешествующая по ссылкам в интернете и определяющая дальнейший путь паука.
5. **База данных** — хранилище скаченных пауком и проиндексированных индексатором информационных страниц.
6. **Система выдачи результатов** — выдает из базы данных список веб-страниц при запросе пользователя.

Работа систем:

В интернете существует множество различных поисковых систем, но принцип их работы во многом схож. Сначала на сайт попадает паук и проводит ряд необходимых подготовительных мероприятий. Затем он скачивает страницу, разбирая ее на необходимые ему части. После этого скопированную страницу посещает индексатор, он выполняет свой алгоритм действий. После чего страница попадает в базу данных. Конечный процесс - это выдача результатов пользователю. База данных ИПС обновляется чаще, чем в каталогах.

Одну из первых строчек в рейтинге ИПС занимает **Google** по удобству, скорости, результативности. С 2001 г. эта ИПС обеспечивает поиск на 80 языках. В этой системе используется механизм индексации и сортировки данных, поэтому в отличие от других систем, в первых десяти результатах не бывает случайных ссылок.

3. **Метапоисковые системы** (поисковые службы) — системы способные послать запросы пользователя одновременно нескольким поисковым серверам, затем объединить полученные результаты и представить их пользователю в виде документа со ссылками. Пример: Metacrawler (метакраулер). В отличие от отдельных поисковых систем, метапоисковые системы не имеют собственных баз данных.

4. **«Электронные библиотеки»** — вид информационных систем, снабженных программным интерфейсом поиска и дистанционного доступа к информации, в которых документы хранятся в электронной, «оцифрованной» форме, и могут быть прочитаны с помощью современных средств коммуникации. Характерной чертой электронных библиотек является наличие единого механизма доступа к разнородной информации, хранящейся в базе данных: текстовой, графической, звуковой и видео. Пример: Русская виртуальная библиотека, Альдебаран.

Теория информационного поиска

Способы реализации поиска зависят от вида, назначения конкретной ИПС. Прочитать каждый документ, чтобы найти необходимый, практически невозможно. Поэтому на протяжении истории развития информационного поиска разрабатывались и совершенствовались методы поиска. Каждому документу, вводимому в поисковый массив, ставится в соответствии поисковый образ документа (ПОД), который представляет собой характеристику, отражающую основное смысловое содержание документа (ключевые слова). Процедура образования ПОД называется *индексированием*. На сегодняшний день нет другого способа быстрого поиска данных, кроме поиска по ключевым словам.

В виде такой же краткой характеристики – поискового предписания или поискового образа запроса (ПОЗ) должны быть сформулирован и информационный запрос. Благодаря этому процедура поиска может быть сведена к простому сопоставлению поисковых образов документов с поисковым предписанием. Точность поиска определяется степенью совпадения поисковый образ документа и поискового образа запроса.

Применение справочно-правовых систем

Справочно-правовые системы (информационно-правовые системы) — это компьютерные базы данных, содержащие нормативные документы органов власти, а также другие документы, их разъясняющие или иллюстрирующие их применение. При использовании справочно-правовой системы существенно экономится время, которое специалист тратит на поиск и проверку нужного нормативного акта.

Использование компьютерных технологий для работы с законодательной информацией началось еще во второй половине 1960-х гг. Ранее справочные системы создавались в основном в виде электронных картотек («электронных каталогов»). Современные СПС имеют

более сложную структуру, работают в диалоговом режиме. Сопровождением подобных систем занято большое количество компаний. Сервисные фирмы осуществляют поставку и текущее сопровождение систем. Крупные фирмы производят разработку программного обеспечения, наполнение и совершенствование СПС. К ним относятся:

- Консультант Плюс (АО «Консультант Плюс»);
- ГАРАНТ (НПП «Гарант-Сервис»);
- «Кодекс» (Центр компьютерных разработок).

Сложности обслуживания таких систем связаны с тем, что новые документы связываются сотнями и тысячами ссылок как с документами, введенными ранее, так и между собой, если для них готовятся новые редакции, примечания, комментарии и т. д., то процесс параллельной обработки такого количества документов становится серьезной проблемой.

В современных информационно-поисковых системах используются виды поиска:

- по реквизитам документов;
- по тексту документа;
- по специализированным классификаторам.

Поиск по реквизитам документа позволяет найти документ с определенным номером, выпущенный определенным органом и т. д. Для повышения эффективности поиска можно задавать интервалы номеров или дат.

Поиск по тексту документов заключается в простом переборе массива документов и нахождении среди них тех, которые содержат слово, заданное пользователем. Таким образом, описанный поиск позволяет найти документы по какой-либо теме. Недостатком поиска по тексту является попадание в список найденных документов большого количества документов, удовлетворяющих запросу, но пользователю в данный момент не нужных. Уменьшить их количество может опыт составления запросов к системе.

Справочно-правовая система «Консультант Плюс»

Справочная правовая система Консультант Плюс является самой популярной СПС (по исследованию ВЦИОМ 2012 г.). Систему Консультант Плюс используют в качестве надежного помощника многие специалисты: юристы, бухгалтеры, руководители организаций, а также специалисты государственных органов, ученые и студенты. В ней содержится огромный массив справочной правовой информации. Начало проекта Консультант Плюс - январь 1992 г.

В СИСТЕМЕ КОНСУЛЬТАНТ ПЛЮС - 20 478 808 ДОКУМЕНТОВ

Разделы системы Консультант Плюс:

ЗАКОНОДАТЕЛЬСТВО	<ul style="list-style-type: none"> • Федеральное законодательство • Законодательство 83 субъектов РФ 	4 806 365 документов
	<ul style="list-style-type: none"> • Документы высших судебных органов 	14 992 366 документов

СУДЕБНАЯ ПРАКТИКА	<ul style="list-style-type: none"> • Подборки судебных решений 	тов
ФИНАНСОВЫЕ И КАДРОВЫЕ КОНСУЛЬТАЦИИ	<ul style="list-style-type: none"> • Путеводитель по налогам • Путеводитель по сделкам • Путеводитель по кадровым вопросам 	327 742 документов
КОММЕНТАРИИ ЗАКОНОДАТЕЛЬСТВА	<ul style="list-style-type: none"> • Постатейные комментарии к законам и кодексам • Книги, монографии, учебные пособия 	67950 документов

И другие разделы по отдельным отраслям хозяйства.

Система позволяет производить следующие операции с текстом документа:

- Просмотр текста, оглавления;
- Выставление закладок в тексте;
- Выделение фрагментов текста;
- Конвертация в формат MS Word;
- Поиск по ключевым словам;
- Сохранение и печать документа.

Контрольные вопросы

1. Что такое Интернет?
2. Каких типов бывают коммуникационные сети?
3. Что такое провайдер?
4. Что такое гипертекстовая система WWW?
5. Внутреннее устройство поисковой машины.
6. Типы ИПС в Интернет.