



**Некоммерческое
акционерное
общество**

**АЛМАТИНСКИЙ
УНИВЕРСИТЕТ**

*Кафедра
информационных
систем*

WEB-ТЕХНОЛОГИИ

Конспект лекции для студентов специальностей

5B070300 - Информационные системы, 5B060200 - Информатика

Алматы 2015

СОСТАВИТЕЛИ: Имангалиев Ш.И., Дауренбаева Н.А. Web-технологии. Конспект лекции для студентов специальностей 5В070300 – Информационные системы, 5В060200 – Информатика. - Алматы: АУЭС, 2015. – 54 с.

Настоящий конспект лекции составлен в соответствии с программой курса «Web-технологии» для студентов специальностей 5В070300 - Информационные системы, 5В060200 – Информатика.

Библиография – 7 названий.

Рецензент: канд.техн.наук., доц. К.С. Чежимбаева

Печатается по плану издания некоммерческого акционерного общества «Алматинский университет энергетики и связи» на 2015 г.

© НАО «Алматинский университет энергетики и связи», 2015 г.

Содержание

Лекция №1. Основы WEB-технологии. Стандарты WEB и их эволюция.	7
Лекция №2. Взаимодействие компьютеров через Интернет. Технология доступа к сети Интернет.	10
Лекция №3. Модели коммуникации Интернет. Pull- и Push-модели коммуникации. Коммуникационные характеристики сервисов Интернет.	14
Лекция №4. WEB-серверы в Интернет. Генерация содержимого на стороне клиента.	19
Лекция №5. Статические и динамические WEB-сайты. Технология CGI.	22
Лекция №6. Технологии поиска в Интернет. Принципы поиска нужной информации.	27
Лекция №7. Поисковые машины. Механизм работы поисковой машины. Построение индекса.	29
Лекция №8. Кэширование WEB-страницы. WEB-ускорители.	33
Лекция №9. Электронная почта. Сервисы Интернет. Обмен файлами.	36
Лекция №10. Механизм работы сервиса FTP. FTP – клиенты.	39
Лекция №11. Защита информации в Интернет. Криптография и Интернет.	43
Лекция №12. Электронная цифровая подпись.	46
Лекция №13. Идентификация пользователей в Интернет. Механизм Cookies.	48
Лекция №14. Технологии создания клиентских приложений Интернет.	50
Лекция №15. Перспективы развития WEB-технологий и образовательный процесс.	52
Список литературы.	51

Лекция №1. Основы WEB-технологии. Стандарты WEB и их эволюция

Цель лекции: ознакомление с основными понятиями Интернета (понятие WEB).

Содержание лекции: понятие WEB, появление стандартов WEB, проект стандартов, развитие стандартов, формирование W3C.

Интернет в настоящее время является самым большим и популярным межсетевым объединением в мире. Он соединяет десятки тысяч компьютерных сетей и миллионы пользователей во всем мире. При этом объединены компьютеры тысяч различных типов, оснащенные самым разным программным обеспечением. Пользователи Интернет могут не обращать внимания на все эти различия.

Понятие WEB или Интернет происходит от термина «Interconnected networks» (связанные сети), т. е. с технической точки зрения – это глобальное сообщество малых и больших сетей. В более широком смысле – это информационное пространство, распределенное среди миллионов компьютеров во всем мире, которые постоянно обмениваются данными. Часто под словом Интернет также понимают само информационное содержимое Сети.

Интернет – это технология, которая смогла вобрать в себя целый арсенал уникальных возможностей. Интернет одновременно является мощнейшим и наиболее независимым информационным ресурсом, надежным и оперативным средством связи, основой для развития информационных технологий и творческого самовыражения миллионов людей во всем мире.

Основная задача Интернета это – круглосуточная, высоконадежная связь. Любые два компьютера (или других устройства), подключенных к Интернету, могут связаться друг с другом в любой момент времени. В дальнейшем, при использовании слова «Сеть» в качестве синонима слова Интернет, будем понимать под Сетью именно возможность связать любые два компьютера через Интернет *и обеспечить их* взаимодействие. Каждый компьютер, подключенный к Интернету – это часть Сети.

Появление стандартов WEB.

Во время войны браузеров компании Microsoft и Netscape сосредоточились на реализации новых свойств, а не на решении проблем тех свойств, которые они уже поддерживали, и добавляли собственные свойства и создавали свойства, которые прямо конкурировали с существующими свойствами другого браузера, но были реализованы несовместимым образом.

Разработчики в это время были вынуждены иметь дело со все возрастающими уровнями беспорядка при попытке создания WEB-сайтов,

иногда вынужденные создавать два различных, но по сути дублирующих друг друга сайта для двух основных браузеров, а иногда вынужденные поддерживать только один браузер, блокируя для других использование своего сайта. Это была ужасная работа, и неизбежная отрицательная реакция разработчиков была недалеко.

Формирование W3C.

В 1994 г. Тим Бернерс-Ли основал World Wide WEB Consortium (W3C) в Массачусетском технологическом институте (Massachusetts Institute of Technology) при поддержке CERN, DARPA (в которую была переименована ARPA) и Европейской Комиссии. Консорциум W3C видел свою задачу в стандартизации протоколов и технологий, которые используются для создания WEB, чтобы информационное содержание было доступно как можно большему числу жителей всего мира.

В течение нескольких следующих лет W3C опубликовал несколько спецификаций (называемых «рекомендациями»), включая HTML 4.0, формат изображений PNG, и Каскадные таблицы стилей (CSS), версия 1 и 2.

Однако W3C не принуждает использовать свои рекомендации. Производители должны подчиняться документам W3C, только в том случае, если они хотят пометить свой продукт, как соответствующий W3C. Практически это не слишком ценно с позиции продаж, так как подавляющее большинство пользователей WEB не знают, и, вероятно, не беспокоятся, о том, что такое W3C. Поэтому «война браузеров» продолжилась с прежней силой.

Проект стандартов WEB.

В 1998 г. на рынке браузеров доминировали Internet Explorer 4 и Netscape Navigator 4. Была выпущена бета-версия Internet Explorer 5, в которой был реализован новый и патентованный динамический HTML. Это означало, что профессиональным разработчикам WEB требовалось знать пять различных способов записи JavaScript.

В результате группа профессиональных разработчиков и дизайнеров WEB решили объединиться. Эта группа назвала себя «WEB Standards Project».

(WaSP). Идея состояла в том, что называя документы W3C стандартами, а не рекомендациями, они смогут убедить компании Microsoft и Netscape поддержать их. Начальным методом распространения призыва к действию было использование традиционного приема рекламы, называемого «roadblock», когда компания проводит рекламу через все информационные каналы одновременно, поэтому независимо от того, как зритель будет переключаться между каналами, он получит одно и то же сообщение. Группа WaSP опубликовала статью одновременно на различных сайтах, ориентированных на разработку приложений WEB, включая <http://builder.com>, Wired online, и некоторых популярных списков рассылки.

Кроме того они стали высмеивать компании, которые присоединялись к W3C (и другим организациям стандартизации), но затем сосредотачивались

больше на создании новых средств, чем на реализации для начала основ, с правильностью которых они согласились.

И последнее, группа WaSP показывает также помощь, а не только критикует. Семь членов сформировали группу «CSS Samurai», которая идентифицировала десять основных проблем поддержки CSS в Opera и Internet Explorer (компания Opera исправила свои ошибки, Microsoft — нет).

Развитие стандартов WEB.

В 2000 г. компания Microsoft выпустила Internet Explorer 5 Macintosh Edition. Это был очень важный этап, так как этот браузер, устанавливаемый в то время в Mac OS, и использовавшийся по умолчанию, имел также приемлемый уровень поддержки рекомендаций W3C. Вместе с приличным уровнем поддержки CSS и HTML в Opera, это способствовало общему положительному движению, когда разработчики и дизайнеры WEB впервые почувствовали удобство при проектировании сайтов в соответствии со стандартами WEB.

Группа WaSP убедила компанию Netscape отложить выпуск версии 5.0 браузера Netscape Navigator, пока она не станет более согласованной со стандартами (эта работа сформировала основу того, что стало очень популярным браузером Firefox). WaSP создала также «Dreamweaver Task Force» («Рабочую группу по Dreamweaver»), чтобы стимулировать компанию Macromedia изменить свой популярный инструмент разработки приложений WEB для обеспечения поддержки создания соответствующих стандартам сайтов.

Популярный сайт «A List Apart» по разработке WEB-приложений был переделан в начале 2001 г., и в статье, описывающей как и почему было сформулировано: через шесть месяцев, через год, или максимум через два года, все сайты будут спроектированы с учетом этих стандартов. Можно наблюдать, как наши навыки становятся устаревшими, или начать уже сейчас изучать методы на основе стандартов. Это было немного слишком оптимистично — не все сайты даже в 2008 г. созданы с учетом стандартов WEB. Но многие люди услышали. Старые браузеры потеряли свою долю рынка, и еще два очень высококлассных сайта были перепроектированы с помощью стандартов WEB: журнал Wired в 2000 г., и ESPN в 2003 г. стали лидерами отрасли в поддержке стандартов WEB и новых методов. Также в 2003 г. Дейв Ши открыл сайт с названием «CSS Zen Garden». Он должен был оказать на профессионалов WEB большее влияние, чем что-либо другое, иллюстрируя как можно изменить весь дизайн, изменяя только стиль страницы, контент может остаться тем же самым. С тех пор в сообществе профессиональных разработчиков WEB стандарты стали обязательными. И в данной серии лекций мы предоставим вам прекрасные основы этих методов, чтобы вы могли разрабатывать WEB-сайты также искусно, содержательно, доступно и в соответствии со стандартами, как и большие компании.

Контрольные вопросы

1. Основная задача интернета. Появление стандартов WEB.
2. Для чего нужны стандарты WEB?
3. Какими качествами должен обладать стандарт Интернета?

Лекция №2. Взаимодействие компьютеров через Интернет. Технология доступа к сети Интернет

Цель лекции: ознакомление с основными технологиями доступа к сети Интернет.

Содержание лекции: полоса пропускания, широкополосный доступ, Асимметричный DSL (ADSL), DSL, ISDN технологии, ISDN-технология, мультиплексор доступа.

DSL-технологии.

Обычный телефон использует лишь низкочастотный диапазон линии. Однако медный двухжильный телефонный провод может передавать гораздо больше данных, если вместо «телефонной» узкой полосы пропускания (3,1 кГц) использовать более широкую полосу. *Полоса пропускания* или *ширина полосы пропускания* - это разность между наиболее высокой и наиболее низкой частотами в диапазоне частот передающего канала. Аналоговая передача сигнала по телефонной линии занимает диапазон от 300 Гц до 3,4 кГц. Цифровой сигнал требует большего диапазона частот. Чем выше скорость передачи, тем больший требуется диапазон частот. Поэтому термином «полоса пропускания» все чаще обозначается верхняя граница скорости передачи данных по компьютерной сети. Термин *широкополосный доступ* означает, что канал предоставляет расширенную полосу частот для передачи информации. Высокая скорость передачи информации достигается благодаря тому, что при использовании широкой полосы частот несколько сигналов могут быть переданы параллельно по одному физическому каналу на нескольких различных частотах, позволяя передавать в единицу времени большее количество информации. Передача нескольких сигналов по одному физическому каналу путем разделения его на подканалы называется *мультиплексированием*. Говоря о частотном мультиплексировании, имеют в виду частотное разделение на подканалы.

Термин «широкополосный доступ» можно пояснить на основе аналогии с транспортными магистралями. Чем больше полос в автостраде, тем большее количество машин может проезжать по ней в единицу времени.

Под термином «узкополосный доступ» обычно подразумевается канал, достаточный для передачи голоса. Также, встречается термин «неширокополосный доступ». Он означает, что скорость передачи данных по каналу не более 64 Кбит/с. Различные источники дают разные определения широкополосному доступу, например, словарь Jupiter Communications определяет его как канал со скоростью передачи не менее 256 Кбит/с. Обычно широкополосный доступ позволяет передавать в одном канале различные сигналы и одновременно пользоваться телефоном и Интернетом.

Одна из технологий, которая позволяет использовать телефонную сеть (изначально приспособленную для передачи голосового сигнала) для высокоскоростной передачи цифровых данных, носит название *DSL*. Аббревиатура *DSL* происходит от «Digital Subscriber Line» (цифровая абонентская линия). Данная технология позволяет использовать более высокие частоты для передачи данных без ущерба для телефонного разговора. Существует целое семейство технологий под единым названием *xDSL* (приставка *x* указывает на конкретную спецификацию семейства *DSL*).

DSL-технология весьма перспективна. Она позволяет одновременно иметь доступ в Интернет и разговаривать по телефону. Скорость *DSL*-подключения намного выше по сравнению с обычным модемом. Для *DSL* не требуется прокладка новых проводов, поскольку используется оборудование телефонной линии.

Асимметричный DSL (ADSL).

При работе в Интернете, основной поток информации идет из Сети к пользователю, в Сеть передается гораздо меньший объем данных. Действительно, например, когда вы просматриваете те *WEB*-страницы, вы посылаете небольшой запрос, а из сети получаете не только текст, но и изображения. То есть информационный обмен является асимметричным. Очевидно, для передачи подобного трафика нужен асимметричный канал. Наглядно представить асимметричный обмен данными можно на базе аналогии с автомагистралью, в которой в одну сторону движется малый поток машин, в обратную сторону – большой. В этом случае логично организовать узкую полосу на въезд и широкую на выезд.

ADSL (Asymmetrical Digital Subscriber Line) или асимметричная *DSL*, позволяет передавать пользователю данные со скоростью, на порядок превышающей скорость передачи данных от пользователя. При этом сигнал от пользователя в Сеть передается на более низких частотах, чем сигнал из Сети пользователю.

Благодаря использованию новейших технологий скорость передачи данных по *ADSL* в десятки раз выше по сравнению с имеющимися видами коммутируемого доступа: около 8 Мбит/с в направлении к абоненту (нисходящий поток) и более 1 Мбит/с от абонента к сети (восходящий поток).

При этом одна и та же линия может параллельно использоваться для передачи голоса и цифровых данных. Высокая скорость позволяет комфортно

работать с WEB-сайтами с мультимедиа-информацией, быстро перекачивать большие файлы и полноценно использовать интерактивные приложения.

Достоинством ADSL является легкость установки: используются уже имеющиеся провода телефонного кабеля, соединяющие пользователя и телефонную компанию. ADSL предоставляет постоянный доступ в Интернет, при этом пользователи ADSL не разделяют полосу пропускания с другими абонентами. Однако ADSL имеет и свои недостатки, прежде всего - ограничение по дальности. При использовании ADSL-технологии скорость передачи потока данных в обратном направлении существенно зависит от расстояния. Если при расстоянии 3 км можно получить скорость около 8 Мбит/с в направлении к абоненту, то на расстоянии 5 км — только 1,5 Мбит/с. В стандарте ADSL 2 реализованы скорости 10 Мбит/с «нисходящего» и 1 Мбит/с «восходящего» потока при дальности до 3 км.

ПК пользователя подключается к ADSL-модему. Не вдаваясь в технические подробности, можно сказать, что принцип действия ADSL-модема заключается в том, что диапазон частот в интервале от 24 кГц до 1100 кГц разбивается на 4000-герцовые полосы, на каждый из которых назначается виртуальный модем. Таким образом, каждый из этих виртуальных модемов работает со своим диапазоном частот.

ADSL-модем подключается к частотному разделителю. Частотный разделитель (или сплиттер) представляет собой фильтр низких частот, разделяющий низкочастотный сигнал обычной телефонной связи (спектр голосовых сигналов) и высокочастотный ADSL-сигнал. Конструктивно частотный разделитель выполняется в виде блока, имеющего три гнезда, одно - для подключения ADSL-модема, другое - для подключения телефонного оборудования, а третье - для подключения к линии ADSL. Частотный разделитель позволяет подключить к одной линии и компьютер, и телефон. Таким образом, по одной линии могут передаваться и компьютерные сигналы, и аналоговые сигналы телефонной связи. Такой же частотный разделитель на телефонной станции позволяет делить низкочастотные и высокочастотные сигналы на другом конце абонентской линии. Голосовой (аналоговый) сигнал направляется в телефонную сеть общего доступа, а цифровой сигнал в мультиплексор доступа DSLAM (Digital Subscriber Line Access Multiplexer), откуда через сеть провайдера попадает в Интернет.

Мультиплексор доступа - это механизм, установленный в здании телефонной компании, который осуществляет подключение всех DSL-абонентов к одной высокоскоростной линии.

ADSL - весьма экономичная технология, обычно линия обходится потребителю намного дешевле, чем выделенный канал аналогичной пропускной способности.

ISDN-технология.

Выделенная телефонная линия - это арендованная телефонная линия связи, постоянно соединяющая двух абонентов. Наиболее распространенной

технологией выделенной линии является технология ISDN (Integrated Services Digital Network). ISDN - это стандарт цифровой передачи данных. Основным компонентом любой ISDN-линии является однонаправленный bearer-канал, или В-канал, с пропускной способностью 64 Кбит/с. По этому каналу могут передаваться цифровые данные и соответственно оцифрованные аудио- и видеоданные. Для расширения полосы пропускания В-каналы группируются по два, и в состав группы каналов, включается D-канал (16 Кбит/с), управляющий передачей данных. Передача информации может осуществляться по обычному медному проводу. Домашние пользователи, которые устанавливают ISDN-адаптер вместо модема, могут получить доступ в Интернет со скоростью до 128 Кбит/с. ISDN требует установки адаптеров на обоих концах линии передачи.

Таблица 2.1 – Максимальные скорости доступа в Интернет при использовании различных технологий

<i>Технология доступа</i>	<i>Максимальная скорость</i>
Коммутируемая телефонная линия	56 Кбит/с
ISDN	128 Кбит/с
Кабельное телевидение	36 Мбит/с
ADSL 2	1 Мбит/с (запрос), 10 Мбит/с (ответ)
Радиоканал в диапазоне 2,4 ГГц	До 54 Мбит/с
Спутниковый доступ	Несколько Мбит/с в зависимости от системы
T1	1,544 Мбит/с
T3	44,74 Мбит/с
GSM-data	9,6 Кбит/с
GPRS	160 Кбит/с (теоретический максимум)
UMTS	2,048 Мбит/с

ISDN-адаптеры иногда называют ISDN-модемами. Следует отметить, что данный термин не совсем точен, поскольку ISDN-адаптеры не выполняют функцию модуляции и демодуляции, линия ISDN изначально является цифровой. ISDN-канал обычно предоставляется телефонными компаниями. По линии ISDN можно вести телефонные разговоры и одновременно передавать данные в Интернет. В принципе, под выделенной линией обычно

понимается любой постоянный канал доступа в Интернет, поэтому в дальнейшем под выделенной линией будем понимать именно это. Ниже приводится таблица максимальных скоростей доступа к Интернету по различным технологиям.

Контрольные вопросы

1. Сколько всего компьютеров может быть одновременно подключено к сети интернет?
2. Какие сервисы предлагает сеть интернет?
3. Из каких частей состоит URL-адрес?
4. Какое сетевое оборудование используется для подключения компьютера к сети интернет?
5. Технологии, лежащие в основе сети интернет.

Лекция №3. Модели коммуникации Интернет. Pull- и Push-модели коммуникации. Коммуникационные характеристики сервисов Интернет

Цель лекции: ознакомление с моделями коммуникации Интернет.

Содержание лекции: Pull- и Push-модели коммуникации. Коммуникативные характеристики сервисов Internet. Мультимедиа. Сетевая навигация. Эффект присутствия.

Рассмотрим коммуникационные характеристики Интернет. В основе традиционной коммуникационной модели лежит процесс «один-ко-многим». В зависимости от средства коммуникации информация может быть представлена в статическом (текст, графика) и/или динамическом (аудио, видеоизображение, анимация) виде. Главной особенностью, лежащей в основе взаимодействия традиционных средств отсутствие интерактивного взаимодействия. В отличие от этой модели, в основе Интернет лежат два других принципа. Во-первых, при коммуникации в Интернет взаимодействие происходит через ее среду, которая вносит значительный вклад в происходящее взаимодействие. Эта модель подчеркивает, что первоначальное общение происходит не между отправителем и получателем информации, а скорее между пользователем и некой средой, коммуникационным пространством, причем оба участника диалога являются как отправителями, так и получателями информации. В данной модели сделан шаг от простого обмена информацией между «передатчиком» и «приемником» к созданию информационной среды, которая затем испытывается и модифицируется участниками диалога.

Интернет представляет собой многонаправленную коммуникационную модель «многие-ко-многим», в которой каждый абонент сети имеет возможность обращаться к другим отдельным абонентам или группам либо от своего имени, либо от имени группы. В данной модели средством коммуникации является распределенная компьютерная сеть, а информация, наряду с обычным представлением, может быть представлена в гипермедийном виде. Данная модель показывает, что интерактивное взаимодействие возможно как с другими пользователями Интернет, так и с самой средой непосредственно (взаимодействие со средой), причем последний вид взаимодействия превалирует. Благодаря наличию такого взаимодействия передатчик информации одновременно является и ее потребителем. В такой модели информация и содержание не просто передается от отправителя к ее потребителю, но вместо этого сама среда создается ее участниками и затем ими же испытывается. Коммуникационная модель для традиционных средств массовой коммуникации не содержит контура обратной связи, в то время как модель среды Интернет включает в себя ярко выраженные обратные связи. Примерами реализации обратных связей с потребителями могут быть электронная почта, данные о регистрации пользователей, файлы «cookie», подписка или регистрация на WEB-серверах. Наличие обратных связей значительно повышает эффективность использования коммуникационного средства взаимодействия и возможности адекватно реагировать на события, происходящие во внутренней и внешней среде.

Pull- и Push-модели коммуникации Интернет.

С этой точки зрения традиционные средства массовой информации (СМИ: телевидение, радио и т. д.) реализуют push-модель доставки информации потребителям, в которой потребители занимают пассивную роль и обладают только достаточно ограниченной возможностью выбора каналов информации.

В противоположность традиционной push-модель доставки информации, в основе Интернет лежит pull-модель, в которой информация предоставляется по запросу (demand pull). Эта особенность среды Интернет связана с активной ролью потребителей, обусловленной контролем над поиском информации и образу навигации за счет использования механизмов поиска или других источников URL (uniform resource locator). В Интернет также существует возможность реализации push-модели. Интернет может следовать этой модели при помощи push-технологии, в соответствии с которой пользователям не требуется производить поиск информации в Интернет, а достаточно подписаться на каналы по интересующей тематике, а информация в дальнейшем будет автоматически доставляться на компьютеры подписчиков каналов.

Технологии среды Интернет развиваются в направлении тесного взаимодействия push- и pull-моделей. Благодаря высокой функциональности Интернет всегда будет существовать возможность навигации пользователей в

целях поиска необходимой информации или каких-либо ресурсов. С другой стороны, пользователи всегда будут иметь возможность выбора поставщиков информации и тематики принимаемых push-каналов.

Интерактивность.

Важной характеристикой среды Internet является интерактивность. Интерактивность - это характеристика протекания процесса коммуникации, определяемая положением по отношению друг к другу коммуникационных сообщений. Для интерактивного взаимодействия характерным является необходимость ответной реакции на приход сообщения или информации, и, плюс к этому, ответ должен находиться в контексте предыдущих сообщений.

Исходя из вышесказанного, для среды Интернет интерактивность можно определить как способность отвечать пользователю, подобно некоторому лицу, участвующему в диалоге. Тем самым, интерактивность расширяет и дополняет представление функции компьютера как участника диалога, способного оценить действия пользователя и отвечать в соответствии с этими оценками. В процессе взаимодействия с использованием компьютерной гиперсреды обращение к клиенту может быть модифицировано самой компьютерной гиперсредой на самой ранней стадии установления отношений в зависимости от того, как клиент отреагировал на предшествующие действия. Возвращаясь к основополагающей модели, лежащей в основе Internet, можно сказать, что интерактивность в среде Internet воплощается уже не на уровне персонального общения через среду, а на уровне взаимодействия непосредственно с самой средой.

Интерактивный характер взаимодействия позволяет существенно повысить эффективность взаимодействия участников коммуникации. На практике, например, это может сократить время, требуемое для получения необходимой информации для заключения сделки или покупки. В то же время, интерактивный характер среды предлагает возможности приспособления связи точно к особенностям индивидуального клиента.

Гипертекст.

Не менее важной особенностью среды Internet является ее гипертекстовая природа. Первая в мире система, основанная на гипертексте, была предложена более пятидесяти лет назад одним из создателей первой ЭВМ (Ванневар Буш). Система называлась Метех и состояла из «устройства, в котором пользователь хранит отдельные книги, записи и коммуникации, и которое механизировано таким образом, что обращение к нему осуществляется гибко и быстро». Нельсон (Nelson) в 1967 году описывал гипертекст как сеть указателей и ассоциаций, организованных по возможности так же, как это делает человек, организуя и связывая в своем сознании отдельные фрагменты информации. Бомман (Bomman) в 1993 году дал следующее определение гипертекста: «Гипертекст предполагает концепцию непоследовательной записи информации, согласно которой пользователь связывает информационные фрагменты между собой с помощью указателей и

связей. В гипертекстовой системе информация представлена в виде узлов и связей».

Гипертекст характеризуется нелинейной сетевой формой организации материала, разделенного на фрагменты, для каждого из которых указан переход к другим фрагментам по определенным типам связей. Особенностью гипертекстовой технологии является то, что она ориентирована на обработку информации не вместо, а вместе с человеком, и, как следствие, становится авторской. Удобство ее использования состоит в том, что пользователь сам определяет подход к освоению или созданию материала с учетом своих способностей, знаний и квалификации. Гипертекст содержит не только информацию, но и аппарат ее эффективного поиска. Таким образом, гипертекстовое представление информации обладает преимуществами перед обычным способом, позволяя производить более эффективное представление и усвоение информации.

Мультимедиа.

Благодаря появлению средств мультимедиа стала возможной компьютерная интеграция статического (текст, графики, рисунки) и динамического (аудио, видео, анимация) содержимого. Комбинация узлов и связей в системе гипертекста со средствами мультимедиа позволило создать новую среду, принципиально отличную от простого сочетания этих двух компонентов.

Гипермедиа – это способ представления информации и доступа к ней. Ее концепция базируется на модели информационного пространства, представленного в виде графа, узлы которого содержат информацию, а семантические связи представлены дугами графа. Информация, хранящаяся в истинной системе гипермедиа, должна быть представлена всеми возможными формами, которые может воспроизвести современный компьютер. Таким образом, гипермедиа совмещает в себе радио (аудио), телевидение (динамическое изображение), прессу (текст, рисунки, фотографии) и компьютер (видеотерминал) посредством гипертекстовых связей, лежащих в основе компьютерной гиперсреды.

Эффект присутствия.

Важной чертой коммуникационной модели Интернет является понятие «эффект присутствия», характеризующее процесс общения клиента с окружающей средой, состоящей из: окружающей обстановки, помещения, рабочего места с компьютером и т. п., и среды, создаваемой компьютерными гиперсредствами. Соотношение степеней воздействия на клиента этих двух составляющих определяет степень «эффекта присутствия». Следует отметить, что «эффект присутствия» по своей природе в большей степени связан с интерактивным общением и с самой компьютерной гиперсредой, нежели с общением через нее. Важным свойством среды при этом является ее «прозрачность» для клиента, которая при персональном общении позволяет протекать диалогу наиболее естественно, а при общении с гиперсредой

позволяет последней реализовать в полной мере присущие ей свойства «виртуальной реальности».

Сетевая навигация.

Сетевая навигация может быть определена как процесс самоопределяемого движения в компьютерной гиперсреде. Метод нелинейного поиска и затребования информации, лежащий в ее основе, предоставляет клиенту как свободу выбора, так и возможность контроля, осуществляемого на более высоком уровне по сравнению с весьма ограниченными средствами навигации, доступными в обычных средствах массовой информации (телевидение, печать).

Таблица 3.1 – Характеристики коммуникационных сервисов Internet

№	Сервис	Междичностное взаимодействие	Машинное взаимодействие	Модель коммуникации	Вид представления информации	Симметричность обратной связи	Возможность интерактивного взаимодействия
1	Электронная почта	Да	Да	Один-к-одному Один-ко-многим	Текст, графика, звук	Да	Нет
2	Списки рассылки	Да	Да	Многие-ко-многим	Текст	Да	Нет
3	Конференции	Да	Да	Многие-ко-многим	Текст	Да	Нет
4	Разговорные комнаты	Да	Да	Многие-ко-многим	Текст	Да	Да
5	WWW	Да	Да	Многие-ко-многим	Текст, графика, звук, видео	Нет	Да
6	Программы голосовой связи	Да	Да	Один-к-одному	Звук	Да	Да
7	Аудио- и видео-конференции	Да	Нет	Многие-ко-многим	Звук, изображение	Да	Да

Коммуникативные характеристики сервисов Internet.

Поскольку Интернет представляет собой комплекс сервисов, то при выборе и применении каждого из них необходимо представлять их индивидуальные характеристики. В таблице 1 представлены результаты анализа коммуникационных сервисов Интернет по виду коммуникации, лежащей в основе коммуникационной модели, виду представляемой информации, симметричности обратной связи и возможности интерактивного взаимодействия.

Представленные данные показывают широкий диапазон видов коммуникаций в среде Интернет в зависимости от используемых средств. При решении конкретных задач необходим отдельный анализ каждого из сервисов Интернет на предмет необходимости его использования и оценки его эффективности.

Контрольные вопросы

1. Как компьютеры взаимодействуют через Интернет?
2. Форматы других приложений или подключаемых модулей.
3. Как определяется сетевая навигация?
4. Коммуникативные характеристики сервисов Интернет.
5. Что является важной чертой коммуникационной модели Интернета?

Лекция №4. WEB-серверы в Интернет. Генерация содержимого на стороне клиента

Цель лекции: ознакомление с различными технологиями доступа в Интернет, иметь представление о программе CGI, которая запускается WEB-сервером в реальном времени.

Содержание лекции: «Последняя миля». Сеть с коммутацией каналов. Модем. Генерация содержимого на стороне клиента.

Большинство из этих пользователей не имеет канала для высокоскоростного доступа к Сети, аналогичного тому, который имеет провайдер. Технологии, по которой осуществляется связь между абонентом и точкой присутствия провайдера, получила название *технологии последней мили*. Название это условно, а на практике расстояние, о котором идет речь, конечно, не обязательно равняется одной миле. В большинстве случаев «последняя миля» является узким местом в цепи каналов доставки информации из Сети на компьютер пользователя. Технология последней мили доступа в Интернет может быть самой различной (это может быть и

высокоскоростная выделенная линия, и медленное сеансовое подключение по телефонной сети общего пользования), но принцип «чем выше скорость доступа, тем выше плата за использование канала» является универсальным. Существует множество технологий, которые позволяют использовать уже имеющуюся инфраструктуру – телефонные линии, сети кабельного телевидения и др., для осуществления доступа в Интернет.

Различные технологии доступа в Интернет: модемный доступ по телефонной линии, доступ по радиоканалу, доступ по сети кабельного телевидения, подключение посредством силовой сети и, наконец, с помощью спутникового канала. Несмотря на столь различные способы соединения с Интернетом, для пользователя разница заключается лишь в устойчивости соединения, скорости передачи данных и времени отклика. Все ресурсы Интернета доступны ему независимо от средств, с помощью которых он соединен с провайдером. Рассмотрим различные варианты подключения к Сети более подробно.

Различные варианты доступа в Интернет.

Доступ по коммутируемой телефонной линии. Доступ по коммутируемой телефонной линии с помощью модема - это пока наиболее распространенный в СНГ способ доступа для домашних пользователей Интернета.

Сеть с коммутацией каналов. Поясним принцип работы коммутируемой телефонной сети или сети с коммутацией каналов. Когда вы дозваниваетесь кому-либо по телефону, между вами устанавливается некий физический канал связи, составленный из отдельных участков, которые соединяются коммутаторами. Участки каждого абонента до первого коммутатора являются неразделяемыми. То есть каждый из абонентов использует эти участки единолично. Остальные участки сети между коммутаторами являются разделяемыми, т. е. могут быть в разное время заняты разными пользователями. Но если какой-либо участок занят одним пользователем, то другой в это время его использовать уже не может. Подключение по коммутируемой телефонной линии - это так называемое временное (сеансовое) соединение. Телефонная линия может быть занята телефонным разговором либо передачей данных с помощью модема. Слово «модем» образовано от слов «МОДулятор» и «ДЕМОдулятор». Модем используется для передачи данных по телефонной линии. В самом общем плане *модем* - это устройство, которое позволяет передать цифровой код (определенное чередование последовательности единиц и нулей) в виде чередования звуков двух разных частот и провести обратное преобразование - перевести колебания звука в цифровую информацию. После того как два модема установили соединение, модулятор генерирует основной сигнал с постоянным периодом колебания (который называется *сигналом несущей частоты* или просто *несущей*) и добавляет второй сигнал, изменяющий несущую. Таким образом, один из параметров несущей изменяется в

соответствии с изменением передаваемого сигнала. Демодулятор анализирует, насколько поступающий сигнал отличается от несущей, и восстанавливает основной сигнал. Иными словами, передающий модем модулирует несущую частоту полезным сигналом и передает высокочастотный сигнал (совместимый с сигналом, передаваемым по телефонной линии), а модем-адресат демодулирует сигнал обратно в цифровую форму соответствующей ASCII-кодировке (American Standard Code for Information Interchange - стандартный код для обмена информацией - набор из 128 кодов символов для машинного представления букв латинского алфавита, чисел и других символов, каждому из которых соответствует конкретное 7-битовое двоичное число. Восьмой бит служит для контроля правильности переданного кода).

При использовании двух различных несущих сигнала данные могут передаваться одновременно в обоих направлениях. Первые модемы были разработаны в исследовательском центре Bell Laboratories и получили название Bell 103. Эти устройства использовали две пары частот: пара частот на один модем, и пара частот на второй. Передающий модем посылал данные, переключаясь между частотами 1,07 и 1,27 кГц, а отвечающий модем посылал данные, переключаясь между частотами 2,025 и 2,225 кГц.

Модемы Bell 103 работали со скоростью 300 бит/с, что позволяло передавать примерно 30 символов в секунду. Такая скорость была вполне приемлемой в период, когда использовался обмен только текстовыми сообщениями, поскольку она даже превышает скорость, с которой человек может читать текст. Модемы, передающие данные со скоростью 300 бит/с, просуществовали до начала 1980-х годов. Однако как только люди начали передавать изображения, такой скорости оказалось недостаточно.

В середине 1980-х годов появились модемы, передающие данные со скоростью 1200 бит/с. В начале 1990-х годов был взят рубеж в 9,6 Кбит/с, и скорость продолжала увеличиваться: 19,2; 28,8; 33,6 Кбит. В 1998 году появились модемы, работающие со скоростью 56 Кбит/с.

Современные модемы – это достаточно сложные устройства, которые используют более сложные схемы модуляции, и, кроме того, упаковывают данные при передаче, однако основной их принцип работы остался прежним. Клиентский компьютер выдает в коммуникационный порт последовательность нулей и единиц, которые могут передавать любые команды и данные. Модем принимает данные, разделяет их на команды и на информацию, которую необходимо передать по телефонной линии, и осуществляет передачу.

Пользователь через модем подключается к телефонной сети общего пользования, а ISP, используя другой модем, выделяет цифровой сигнал и осуществляет связь с Интернетом.

Следует добавить, что современные модемы для коммутируемых линий бывают *внутренние* и *внешние*. Внутренние модемы выполняются в виде

платы расширения, а внешние в виде отдельного устройства с автономным блоком питания.

Обычные современные модемы, использующие аналоговые телефонные каналы, имеют максимальную скорость передачи не более 33 600 бит/с, а приема - 56 Кбит/с. То есть примерно в 200 раз быстрее, чем модель Bell 103, однако эти скорости сегодня мало кого устраивают. Далее рассмотрим альтернативные технологии, позволяющие получить более высокую скорость доступа в Интернет.

Контрольные вопросы

1. Различные технологии доступа в Интернет.
2. Различные варианты доступа в Интернет.
3. Конфигурирование безопасной директории WEB-содержимого.
4. Технология последней мили доступа в Интернет.

Лекция №5. Статитические и динамические WEB-сайты. Технология CGI.

Цель лекции: ознакомление со статитическими и динамическими страницами, а также одной из технологией, которая позволяет добавлять динамическое содержание в WEB-страницы.

Содержание лекции: достоинства и недостатки WEB-сайтов. Технология CGI.

Статический WEB-сайт — сайт, состоящий из статичных html (htm, dhtml, xhtml) страниц составляющих единое целое. Содержит в себе (в виде HTML-размеченных) текст, изображения, мультимедиа содержимое (аудио, видео) и HTML-теги. Теги бывают как служебные, предназначенные для обозревателя, так и предназначенные для размещения, формирования внешнего вида и отображения информации. Все изменения на сайт вносятся в исходный код документов (страниц) сайта, для чего необходимо иметь доступ к файлам на WEB-сервере.

Обычно статический WEB-сайт представляет собой набор файлов HTML, размещенных на WEB-сервере. Достоинства и недостатки представлены в таблице 5.1.

Динамический WEB-сайт — сайт, состоящий из динамичных страниц — шаблонов, контента, скриптов и прочего, в виде отдельных файлов. Страница сайта, показываемая в итоге браузеру пользователя, формируется на

стороне сервера динамически, по запросу, из страницы-шаблона и отдельно хранимого содержимого (информации, скриптов и др.).

Таблица 5.1- Достоинства и недостатки

<i>Достоинства</i>	<i>Недостатки</i>
Для создания страницы не нужно знать языки WEB-программирования	Невозможность динамической генерации содержимого
Хорошая кэшируемость страницы	Невозможность полноценной поддержки посетителей (выбор внешнего вида, поддержка браузеров, cookie)
Быстрая скорость загрузки страниц	Для наполнения сайта информацией необходимо получать доступ к файлам-страницам посредством FTP, или сторонним WEB-скриптам, позволяющим редактировать страницы
Минимальные требования к веб-серверу и минимальная нагрузка на него	При большом количестве страниц (файлов), если возникает необходимость внести однотипные изменения (дизайн, оформление, добавление новых разделов) необходимо использовать стороннее ПО (утилиты)
Простота переноса на другой сервер или локальный компьютер	Сайт-визитка — содержит самые общие данные о владельце сайта (организация или индивидуальный предприниматель). Вид деятельности, история, прайс-лист, контактные данные, реквизиты, схема проезда. Специалисты размещают своё резюме. То есть подробная визитная карточка
Возможность прямого просмотра файла в браузере, без использования промежуточного ПО (WEB-сервера, CGI)	Каталог продукции — в каталоге присутствует подробное описание товаров/услуг, сертификаты, технические и потребительские данные, отзывы экспертов и т. д. На таких сайтах размещается информация о товарах/услугах, которую невозможно поместить в прайс-лист

Редактирование собственно содержимого (как и страницы-шаблона) может производиться как средствами самого сайта, так и с применением стороннего ПО. Возможность править все страницы предоставляется только определенной категории пользователей (например администраторам, или же зарегистрированным пользователям). Как правило для отображения любого количества однотипных страниц используется одна страница-шаблон, в которую подгружается соответствующее содержимое, это позволяет единомоментно корректировать внешний вид сайта (множество всех его страниц), редактируя всего лишь один шаблон.

В некоторых случаях к правке определенного веб-контента допускаются анонимные пользователи, что бывает реже (например, на форумах — добавление сообщений). Процесс получения содержимого сайта обычно выглядит следующим образом:

- 1) Генерация содержимого на стороне сервера.
- 2) Передача сгенерированной странички клиенту.
- 3) Генерация содержимого на стороне клиента.

Генерация содержимого на стороне сервера.

Сервер получает запрос от Клиента (например, *page.ru/index.php*) и запускает обработку файла-скрипта (в данном случае - *index.php*) интерпретатором. Языки программирования на Серверной стороне используются разные, наиболее часто встречаются, например: Java, PHP, Perl, ASP.NET, Ruby и другие.

Именно на этой стадии происходит выборка необходимой информации из баз данных и наполнение ею страницы, после чего готовая страница передаётся Клиенту.

Генерация содержимого на стороне клиента.

После того, как страница получена Клиентом с Сервера, браузер обрабатывает её и отображает Пользователю, при этом *исполняя скрипты* Клиентской стороны, если они были указаны в странице и получены.

На Клиентской стороне используется JavaScript, который может использоваться как для минимальных, невидимых глазу операций (например, проверки на правильность задания паролей при регистрации на сайтах), так и глобальных последовательностей и приложений.

Комбинированная генерация.

Чаще всего в жизни встречается именно комбинация этих двух методов генерации — весь «новый интернет» основан на нём, это, и «умная строка с подсказкой вариантов» у поисковых систем, и «редакторы» в форумах, и всплывающие меню, и многое другое.

После набора в адресной строке браузера необходимого URL, браузер получает информацию об используемом протоколе (HTTP) и имени сервера. Браузер обращается к DNS-серверу, для того чтобы перевести имя сервера в IP-адрес. На основе полученного IP-адреса браузер устанавливает связь с

искомым WEB-сервером и, используя протокол HTTP, запрашивает искомый ресурс. Сервер посылает браузеру HTML-страницу, которая хранится на сервере. Браузер считывает HTML-тэги, воссоздает страницу на экране вашего компьютера, и вы видите результат вашего запроса. Обычно даже простая WEB-страница содержит не только текст, но и графику, т. е. состоит из нескольких файлов разного типа. Для того чтобы браузер мог распознавать эти файлы, сервер информирует о том, какой файл будет отправляться (текст в формате HTML, или графика в формате JPG и т. д.), и после этого передается содержание файла. Несмотря на тот факт, что обычно WEB-страница состоит из многих файлов, за время одного запроса WEB-сервер посылает только один файл. То есть, получив HTML-текст и найдя в нем, например, ссылку на графический элемент, браузер посылает новый запрос на сервер (причем необязательно на тот же самый сервер, с которого пришел HTML-текст). Для того чтобы скачать каждый новый файл, браузер должен послать новый HTTP-запрос. При этом, поскольку современные серверы и браузеры работают в многопоточном режиме, одновременно может выполняться несколько запросов. Если запрашиваемого ресурса по данному адресу нет, то WEB-сервер выдаст сообщение 404/File not found (404/Файл не найден).

Технология CGI.

Страницы, которые были созданы по вашему запросу динамически - «сформированы на лету». Например, любая книга отзывов предоставляет вам определенную форму, в которую вы добавляете свой текст и затем, когда вы в следующий раз открываете эту страницу, она содержит новое сообщение. Одной из технологий, позволяющих добавлять динамическое содержание в WEB-страницы, является CGI (Common Gateway Interface). Она позволяет ассоциировать с тем или иным URL не статический документ, а программу, в результате работы которой могут быть сформированы данные в реальном времени. Например, если вы хотите предоставлять мгновенные погодные данные в некотором районе, вы должны каждый раз генерировать новую страницу. Это может быть реализовано на основе CGI-технологии. В момент обращения на сервер на нем запускается CGI-программа, которая обращается к цифровому измерительному устройству и выдает сведения о температуре, давлении и так далее. Каждый раз, обращаясь на этот адрес за справкой, вы будете получать текущие значения. Другой пример: когда вы заполняете некоторый запрос в поисковой системе, то в ответ получаете результат работы CGI-программы в виде набора искомых адресов.

CGI-программу можно представить как часть WEB-сервера, которая запускается WEB-сервером в реальном времени. Сервер передает запросы пользователя CGI-программе, которая их обрабатывает и возвращает результат своей работы на экран пользователя. С точки зрения клиента нет никакой разницы, что находится по адресу конкретного URL - статический документ или CGI-программа. Браузер воспринимает информацию одинаково вне зависимости от того, сформирована ли она «на лету» или это статическая

страница. Результат работы CGI-программы имеет тот же формат, что и статический документ. Под термином CGI понимают не только программу, но и протокол. В этом смысле CGI - это стандартный способ для WEB-сервера передать пользовательский запрос прикладной программе и получить данные для передачи их пользователю. Протокол CGI пересылки данных между сервером и приложением является частью протокола HTTP.

Большая часть CGI-программ является CGI-скриптами. *Скриптом* называют набор инструкций, которые интерпретируются или выполняются другой программой. Некоторые языки, такие как Perl, JavaScript, задумывались именно как скриптовые языки или, как их еще называют, языки описания сценариев. В принципе CGI-программы могут быть написаны как на скриптовых языках, так и на компилируемых: C, C++, Delphi.

Альтернативой технологии CGI является технология компании Microsoft, называемая Active Server Page (ASP), которая построена по тому же принципу: скрипт, включенный в WEB-страницу, выполняется на сервере до того, как страница отсылается пользователю. Существует также ряд других технологий, работающих по тому же принципу.

В отличие от схемы, где динамическое содержание формируется на стороне сервера, можно также генерировать динамическое содержание страниц уже на стороне клиента.

В последнем случае активные документы хранятся на WEB-серверах и скачиваются на локальный компьютер так же как и статические страницы, а после того, как программа активного документа оказывается на локальном компьютере, она выполняет на нем определенные вычисления, и результаты этих вычислений отображаются на экране. Соответственно, скорость отображения данных на экране не зависит от скорости связи с удаленным сервером, поскольку вывод активного документа происходит только после его загрузки на локальный компьютер. Следует при этом отметить, что загрузка программы, служащей для формирования документа, может занимать весьма ощутимое время. Для создания активных документов могут использоваться различные технологии: это могут быть приложения, написанные на JavaScript, Java-апплеты и элементы управления ActiveX.

Контрольные вопросы

1. Генерация содержимого на стороне сервера.
2. Генерация содержимого на стороне клиента.
3. Комбинированная генерация.
4. Достоинства и недостатки статического WEB-сайта.

Лекция №6. Технологии поиска в Интернет. Принципы поиска нужной информации

Цель лекции: ознакомление с принципами поиска информации и механизмом работы поисковой машины.

Содержание лекции: Поиск в WEB. Индексированные каталоги. Тематические коллекции ссылок. Подбор доменного имени. Поисковые машины. Поиск по индексу.

В Интернете размещены миллионы сайтов, причем наряду с актуальной информацией соседствует много устаревших ресурсов. Интернет - это наиболее демократичный источник информации, где нет единоличного управления. Каждый может разместить в Сети собственный ресурс. В результате мало кто озабочен тем, чтобы избежать дублирования информации или следовать стандартам. Не зря бытует мнение, что в Сети есть все, но найти там что-либо непросто. Видимо, для того чтобы находить, нужно уметь искать. В настоящем разделе сделан обзор инструментов поиска в сети Интернет, объясняется механизм работы поисковых систем и даются практические рекомендации по оптимизации поиска.

Для поиска в Интернете предназначены различные инструменты: поисковые машины (поисковики), индексированные каталоги (рубрикаторы), рейтинги, метапоисковые системы и тематические списки ссылок, онлайн-энциклопедии и справочники. При этом для поиска информации разного типа наиболее эффективными оказываются различные категории инструментов поиска. Рассмотрим каждую категорию в отдельности.

Индексированные каталоги

Каталог представляет собой данные, сгруппированные по темам в виде иерархических структур. Тематические разделы первого уровня определяют наиболее популярные, максимально широкие темы, такие как «спорт», «отдых», «наука», «магазины» и т. д. В каждом таком разделе есть подразделы. Таким образом, вы можете уточнять интересующую вас область, путешествуя по дереву каталога, постепенно сужая область поиска. Например, при поиске учебных заведений цепочка может иметь вид: *Образование -> Учебные заведения -> Высшие учебные заведения -> Институты*. Дойдя до нужного подкаталога, вы находите в нем набор ссылок. Обычно в каталоге все ссылки являются профильными, поскольку составлением каталогов занимаются не программы, а люди. Очевидно, что если вы ищете общую информацию по некоторой широкой теме, то целесообразно обратиться к каталогу. Если же вам необходимо найти конкретный документ, то каталог окажется малоэффективным поисковым средством.

Помимо каталогов общего профиля, в Сети достаточно много специализированных. В случае, если внутри отдельной темы каталога

находится огромное количество ресурсов, возможно дополнительное ранжирование по популярности. Например, в каталоге Яндекса ранжирование идет по индексу цитирования. (*Индекс цитирования* - это число ссылок на ваш сайт с других сайтов в Сети).

Помимо каталогов, в Сети существуют рейтинги. От каталога рейтинг отличается тем, что в нем описание ресурсов делают непосредственно их владельцы, а в каталоге - авторы, т. е. его редакторы.

Тематические коллекции ссылок

Тематические коллекции ссылок - это списки, составляемые группой профессионалов или даже коллекционерами-одиночками. Очень часто узкоспециализированная тема может быть раскрыта одним специалистом лучше, чем группой сотрудников крупного каталога.

Подбор доменного имени

Каталог - удобная система поиска, однако если вам нужно попасть на сервер компании Intel или IBM, то вы вряд ли станете обращаться к каталогу. Угадать название соответствующих сайтов нетрудно: www.intel.com, www.ibm.com.

Аналогично, если вам необходим сайт, посвященный погоде в мире, его логично искать на сервере www.weather.com. В большинстве случаев найти сайт с ключевым словом в названии предпочтительнее, чем документ, в тексте которого это слово многократно используется. При поиске малоизвестной компании интуитивный подбор адреса может успешно конкурировать с другими приемами поиска, поскольку при подобной системе поиска вы можете установить соединение даже с сервером, который не зарегистрирован ни в одной поисковой системе. Однако очевидно, что подобное угадывание не совсем эффективный способ и, если вам не удастся подобрать искомое имя, придется обратиться к поисковой машине.

Контрольные вопросы

1. Какой вид поиска является самым быстрым и надежным?
2. Где пользователь может найти адреса WEB-страниц?
3. Каково основное назначение поисковой системы?
4. Из каких частей состоит поисковая система?
5. Какие поисковые системы вы знаете?
6. Какова технология поиска по рубриктору поисковой системы?
7. Какова технология поиска по ключевым словам?

Лекция №7. Поисковые машины. Механизм работы поисковой машины. Построение индекса

Цель лекции: ознакомиться с механизмом работы поисковой машины.

Содержание лекции: Поиск по индексу. Механизм работы поисковой машины. Поиск по индексу.

Поисковые машины.

В ответ на запрос вы обычно получаете длинный список документов, многие из которых не имеют никакого отношения к тому, о чем вы спрашивали. Такие документы называются *нерелевантными* (от англ, *relevant* - подходящий, относящийся к делу). Таким образом, *релевантный* документ - это документ, содержащий искомую информацию. Очевидно, что от умения грамотно составлять запрос зависит процент релевантных документов в списке найденных ссылок. Доля релевантных документов в списке всех найденных поисковой машиной документов называется *точностью поиска*. Нерелевантные документы называют *шумовыми*. Если все найденные документы релевантны (шумовых нет), то точность поиска составляет 100%. Если найдены все релевантные документы, то *полнота поиска* - 100%.

Таким образом, качество поиска определяется двумя взаимозависимыми параметрами: точностью и полнотой поиска. Увеличение полноты снижает точность и наоборот.

Механизм работы поисковой машины.

Поисковые системы можно сравнить со справочной службой, агенты которой обходят предприятия, собирая информацию в базу данных. При обращении клиента в службу информация выдается из этой базы. Данные в базе устаревают, поэтому агенты их периодически обновляют. Некоторые предприятия сами присылают данные о себе, и к ним приезжать агентам не приходится. Иными словами, справочная служба имеет две функции: создание и постоянное обновление данных в базе и поиск информации в базе по запросу клиента.

Аналогично, поисковая машина состоит из двух частей: так называемого *робота*, который обходит все предписанные серверы, формируя базу данных, и *поискового механизма*. У термина робот множество синонимов, помимо робота его также называют *сетевым агентом*, или, подчеркивая тот факт, что он ползает по сети, - *пауком* или *червяком*.

База робота в основном формируется им самим (робот сам находит ссылки на новые ресурсы) и в существенно меньшей степени - владельцами ресурсов, которые регистрируют свои сайты в поисковой машине. Помимо робота, формирующего базу данных, существует программа, определяющая рейтинг найденных ссылок.

Принцип работы поисковой машины сводится к тому, что она опрашивает свой внутренний каталог (базу данных) по ключевым словам, которые пользователь указывает в поле запроса, и выдает список ссылок, ранжированный по релевантности.

Следует отметить, что поисковая система оперирует именно внутренними. Несмотря на то, что база данных поисковой машины постоянно обновляется за счет опроса узловых адресов в Сети, внутренние ресурсы поисковой машины и ресурсы Сети несопоставимы, и поэтому всегда есть вероятность, что машина даст устаревший адрес или не найдет нужный ресурс. Проблема состоит не только в ограниченности внутренних ресурсов, но и в том, что скорость робота ограничена. Увеличение внутренних ресурсов поисковой машины не решает проблемы в силу того, что скорость обхода конечна. При этом нельзя сказать, что поисковая машина внутри имеет копию определенной части исходных ресурсов Интернета, разложенных по каталогу. Полная информация (исходные документы) хранится отнюдь не всегда, чаще хранится лишь ее часть - так называемый *индексированный список* или *индекс*, который гораздо компактнее текста документов.

Для построения индекса исходные данные преобразуются таким образом, чтобы объем базы был минимальным, а поиск осуществлялся очень быстро и давал максимум полезной информации. Объясняя, что такое индексированный список, можно провести параллель с его бумажным аналогом - так называемым *конкордансом*, т. е. словарем, в котором в алфавитном порядке перечислены слова, употребляемые определенным писателем, а также указаны ссылки на них и частота их употребления в произведениях писателя.

Очевидно, что конкорданс (словарь) гораздо компактнее исходных текстов произведений, и найти в нем нужное слово намного проще, нежели перелистывать книгу в надежде наткнуться на ключевое слово.

Построение индекса.

Сетевые агенты, или роботы-пауки, «ползают» по Сети, анализируют содержимое WEB-страниц и собирают информацию о том, что и на какой странице было обнаружено. При нахождении очередной HTML-страницы большинство поисковых систем фиксирует слова, картинки, ссылки, скрипты и другие элементы (в разных поисковых системах по-разному), содержащиеся на ней. Причем при отслеживании слов на странице фиксируется не только их наличие, но и местоположение, т. е. где эти слова находятся: в заголовке (title), подзаголовках (subtitles), в метатэгах (meta tags) или в других местах. Обычно фиксируются значимые слова, а союзы и междометия типа «а», «но» и «или» игнорируются. Метатэги позволяют владельцам страниц самим определять ключевые слова и тематику, по которым индексируется страница. Это особенно актуально в случае, когда ключевые слова имеют несколько значений. Метатэги могут сориентировать поисковую систему при выборе из нескольких значений слова на единственно правильное. Однако метатэги

работают надежно только в том случае, когда заполняются честно. Недобросовестные владельцы WEB-сайтов помещают в свои метатэги наиболее популярные в Сети слова, не имеющие ничего общего с темой их сайта только для того, чтобы заманить к себе новых посетителей и повысить рейтинг посещаемости своего ресурса. Исключение из поиска подобных сайтов - это еще одна задача, которую должна решать хорошая поисковая система. Каждый робот поддерживает свой собственный список ресурсов, наказанных за недобросовестную рекламу.

После сбора информации на предписанных WEB-страницах идет индексация полученных данных. Роботы-пауки просматривают информационное наполнение WEB-страниц и создают индексированную базу поиска по ключевым словам, а затем по запросу пользователя система выдает ранжированный по релевантности список сайтов.

Очевидно, что если вы ищете сайты по ключевому слову «цветы», то поисковый механизм должен найти не просто все страницы, где находится слово «цветы», а те, где это слово имеет отношение к теме сайта. Для того чтобы определить, насколько то или иное слово имеет отношение к профилю некоторой WEB-страницы, необходимо оценить, насколько часто оно встречается на странице, есть ли по данному слову ссылки на другие страницы или нет. Короче говоря, необходимо ранжировать найденные на странице слова по степени важности. Словам присваиваются весовые коэффициенты в зависимости от того, сколько раз и где они встречаются (в заголовке страницы, в начале или в конце страницы, в ссылке, в метатэге и т. п.). Каждый поисковый механизм имеет свой алгоритм присваивания весовых коэффициентов - это одна из причин, по которой разные поисковые машины по одному и тому же ключевому слову выдадут различные списки ресурсов. Поскольку страницы постоянно обновляются, то и процесс индексирования должен выполняться постоянно. Роботы-пауки путешествуют по ссылкам и формируют файл, содержащий индекс, который может быть довольно большим. Для уменьшения его размеров прибегают к минимизации объема информации и сжатию файла. После обработки данные сохраняются в базе, которая постоянно обновляется. Имея несколько роботов, поисковая система может обрабатывать сотни страниц в секунду. Сегодня мощные поисковые машины хранят сотни миллионов страниц и получают десятки миллионов запросов ежедневно.

При построении индекса решается также задача снижения количества дубликатов - задача нетривиальная, учитывая, что для корректного сравнения нужно сначала определить кодировку документа. Еще более сложной задачей является отделение очень похожих документов (их называют «почти дубликаты»), например таких, в которых отличается лишь заголовок, а текст дублируется. Подобных документов в Сети очень много - например, кто-то списал реферат и опубликовал его на своем сайте под другим заголовком. Современные поисковые системы позволяют решать все эти проблемы.

Поиск по индексу.

Поиск по индексу заключается в том, что пользователь формирует запрос и передает его поисковой машине. При использовании нескольких ключевых слов весьма полезно применение языка запросов, основу которого составляют булевы операторы.

Наиболее часто используемые булевы операторы:

- AND - все термины, соединенные «AND», должны присутствовать в предлагаемом документе. Некоторые поисковые системы используют значок «+» вместо «AND»;

- OR - как минимум одно из ключевых слов, соединенных «OR», должно присутствовать в искомом документе;

- NOT- ключевое слово (слова), следующее за «NOT», не должно появляться в искомом документе. Некоторые поисковые системы используют значок «-» вместо «NOT»;

- FOLLOWED BY - одно из ключевых слов должно следовать непосредственно за другим;

- NEAR - одно из слов должно отстоять на определенное количество слов от другого;

- *кавычки* - слова внутри кавычек - это фрагмент текста, который должен быть найден в пределах документа или файла. Отметим, что семантика языка запросов может несколько отличаться в конкретных поисковых машинах, но обычно описана в справке поисковой машины.

Текст, в пределах которого проверяется логическая комбинация, называется *единицей поиска*. Это может быть предложение, абзац или весь документ. В разных поисковых системах могут использоваться различные единицы поиска. Поиск в пределах предложения возможен для тех систем, которые имеют в индексе подробный адрес.

После того как пользователь передал запрос поисковой системе, она обрабатывает синтаксис запроса, сравнивает ключевые слова со словами в индексе. Затем составляется список сайтов, отвечающих запросу, они ранжируются по релевантности и формируется результат поиска, который выдается пользователю.

Контрольные вопросы

1. Качество индекса.
2. Принципы работы поисковой машины.
3. Исходные данные для построения индекса.
4. Что такое «единица поиска»?

Лекция №8. Кэширование WEB-страницы. WEB-ускорители

Цель лекции: ознакомление с технологиями кэширования.

Содержание лекции: кэширование WEB-страниц. Прокси-сервер. Срок годности документов. WEB-ускорители.

Возникает вопрос: как ускорить загрузку. Одно из решений состоит в том, чтобы не скачивать данные с удаленных серверов по нескольку раз. При навигации в WEB часто возникает необходимость как в книге «перелистнуть страницу назад», т. е. загрузить страницу повторно. Очевидно, что когда вы в браузере нажимаете кнопку *Назад* для того, чтобы вернуться к просмотренной ранее странице, нет смысла повторно скачивать ее с сервера. Логично сохранить просмотренный документ на диске вашего компьютера в специальной буферной зоне (кэше) и быстро отобразить его, если пользователь захочет вернуться по ссылкам назад. Именно так и поступают современные браузеры. Предположим, что вы просматриваете страницы в последовательности 1-2-3 и страница 3 имеет ссылку на страницу 4. В этом случае обычно переход со страницы 3 на страницу 2 будет быстрее (загрузка из кэша), чем на страницу 4 (загрузка из Сети).

В современных браузерах (таких как Internet Explorer и Netscape Navigator) существует возможность настройки объема папки, в которой сохраняются кэшированные документы.

Прокси-сервер.

Технология кэширования имеет свои ограничения. Например, если вы и ваш коллега подключены к Интернету через одного и того же провайдера. В какой-то момент времени вы запрашиваете документ, который только что скачивал ваш коллега по каналу того же провайдера. Нерационально несколько раз обращаться в Сеть за одним и тем же документом. Логично кэшировать данные на уровне сервера провайдера.

Решение этой задачи провайдеры часто возлагают на так называемый прокси-сервер (от англ. *Proxy*-заместитель, посредник). *Прокси-сервер* – это сервер, который выполняет роль посредника между рабочей станцией пользователя и Интернетом, и позволяет повысить безопасность, административный контроль и использовать функции кэширования.

Прокси-сервер получает запрос от пользователя на выполнение какого-либо интернет-сервиса, например, запрос на просмотр WEB-странички. Если прокси выполняет функции кэширования, он смотрит, нет ли запрашиваемой странички в локальном кэше сервера. Если такая страница есть, она возвращается пользователю без передачи запроса дальше в Интернет. Если странички в кэше нет, прокси-сервер выступает как клиент от лица пользователя, использует свой собственный IP-адрес и запрашивает необходимую страничку в Интернете. Когда страничка возвращена, прокси-

сервер передает ее пользователю. Функция кэширования прокси-сервера может существенно ускорить работу с Интернетом и снизить объем сетевого трафика. Кэширование WEB-страниц является также одной из мер повышения надежности работы Сети, поскольку позволяет компенсировать временную неисправность нужного WEB-узла.

Обычно на компьютере, на котором работает программа прокси-сервера, выделяется значительное дисковое пространство. Крупные провайдеры могут выделять под кэширование документов десятки и даже сотни гигабайт памяти.

Таким образом, быстрое отображение искомого документа возможно за счет того, что постоянно запрашиваемые документы могут храниться не только на сервере разработчика документа, но также «ближе к пользователю» на прокси-сервере провайдера или непосредственно на компьютере пользователя.

Следует отметить, что когда мы говорим о провайдерах, у которых имеются тысячи клиентов, оказывается, что значительное количество обращений запрашивают ограниченный круг популярных в Сети документов. Прокси-сервер крупного ISP, имеющего дело с тысячами клиентов, удовлетворяет около 50% обращений непосредственно со своего диска. Обычно сервер ISP связан с серверами других интернет-провайдеров региона. Поэтому если интересующий клиента документ не находится в кэше прокси-сервера провайдера, он может быть найден на других прокси-серверах. Таким образом, количество пользователей кэш-пространства увеличивается во много раз. Кэширующие прокси-серверы могут быть установлены не только ISP-провайдерами, но и крупными компаниями, которые хотят снизить нагрузку на сеть и оптимально обрабатывать запросы.

Срок годности документов.

Для того чтобы избежать ситуации получения устаревшего кэшированного файла, существуют правила, позволяющие определить, в каких случаях можно представить на запрос кэшированный документ, а когда нет. Часть этих правил прописана в протоколе HTTP, другие задаются в настройках браузера и администратором прокси-сервера. Есть ряд критериев, по которым происходит принятие решения о кэшировании на прокси-сервере. Например, если объект помечен как конфиденциальный, то он не будет кэшироваться. Аналогично существуют критерии, по которым происходит решение о возможности отправки кэшированного документа на запрос клиента. Например, если срок годности у документа истек, то сервер не отправит его клиенту, а запросит, есть ли на сервере разработчика сайта более свежий документ. Вполне возможно, что документ с истекшим сроком годности на самом деле не обновлялся, т. е. не устарел, и его можно предоставлять на запрос клиента.

WEB-ускорители.

Мы уже говорили о том, что стандартные браузеры используют технологию кэширования для ускорения загрузки WEB-страниц. Однако существуют специализированные программы, которые могут ускорить процесс загрузки лучше, чем это делает Internet Explorer. Ускорение может достигаться за счет оптимизации параметров сетевого подключения, за счет кэширования данных и за счет отсеивания «лишней» информации. В некоторых программах реализовано сразу несколько из вышеперечисленных механизмов.

Ускорение за счет оптимизации процесса кэширования.

Говоря о кэшировании данных для ускорения загрузки WEB-страниц, следует выделить три направления

- кэширование IP-адресов;
- оптимизация дискового кэша на локальном компьютере;
- упреждающее чтение.

Поясним, в чем состоит суть каждого из методов.

Как известно, после того как вы сообщите браузеру URL того или иного ресурса, он обращается к DNS-серверу для перевода доменного имени в IP-адрес. Типична ситуация, когда приходится опросить целую цепочку DNS-серверов, прежде чем удастся найти необходимую информацию. Обычно IP-адреса не кэшируются на компьютере пользователя, и адрес, которым вы пользовались несколько минут назад, запрашивается заново, на что уходит лишнее время. Таким образом, первая возможность увеличения скорости отображения WEB-страниц заключается в использовании программы, которая выполняет более оптимальный опрос DNS-серверов и организывает *кэширование IP-адресов*.

Получив IP-адрес, обозреватель проверяет, не хранится ли копия указанной страницы в кэше компьютера, и если принимается решение, что документ из кэша пригоден, он предоставляется пользователю. Этот процесс тоже требует времени, и его ускорение - отдельная задача, которую можно поручить сторонней программе. Большинство программ этого назначения создают свой собственный кэш, который занимает больше места, но работает быстрее. В этом и состоит второе направление ускорения.

Для того чтобы пояснить третий механизм (упреждающее чтение), опять вернемся к работе обычного браузера. После того как документ отображен браузером, обычно наступает пауза. Пока пользователь читает документ, процесс загрузки прекращается, т. е. система простаивает. Упреждающее чтение позволяет использовать соединение во время паузы, в течение которой подгружаются документы, ссылка на которые имеется на текущей странице. Загрузка этих документов в фоновом режиме позволяет отобразить их намного быстрее в случае запроса пользователем. В этом состоит суть третьего механизма - *упреждающего чтения* или *префетчинга* (от англ. *prefetch* — принести заранее). В некоторых источниках данный механизм называют *упреждающей загрузкой* (от англ. *pre-loading*).

Контрольные вопросы

1. WEB-ускорители. Ускорение за счет оптимизации процесса кэширования.
2. Стандартные браузеры для чего используют технологию кэширования?
3. Как можно ускорить загрузку?

Лекция №9. Электронная почта. Сервисы Интернет. Обмен файлами.

Цель лекции: ознакомление с сервисами Интернет и услугой глобальной сети-электронная почта, а также особенностями посылки письма по электронной почте.

Содержание лекции: адрес электронной почты. Два сервиса - два протокола. Основные функции почтовых клиентов. Отправление и доставка письма. Передача специальных символов и вложенных сообщений.

Самой распространенной услугой в глобальной сети является электронная почта. В настоящее время свой адрес электронной почты имеют приблизительно 350 миллионов человек. Посылка письма по электронной почте обходится значительно дешевле посылки обычного письма. Кроме того, сообщение, посланное по электронной почте, дойдет до адресата за несколько часов, в то время как обычное письмо может добираться до адресата несколько дней, а то и недель.

В момент регистрации доступа в Интернет ISP предоставляет пользователю дисковое пространство под почтовый ящик, адрес почтового ящика (E-mail Account Address), имя пользователя (E-mail Account Login Name) и пароль (E-mail Account Password) для предотвращения несанкционированного доступа к почте. Адрес электронной почты имеет формат: имя_пользователя@имя_домена, например, khan@aipet.almaty.kz. Часть слева от значка «@» - это имя почтового ящика (Account Name) на сервере, из которого владелец этого адреса забирает письма (в данном примере - khan). Как правило имя пользователя совпадает с именем почтового ящика. Часть справа от значка «@» называется *доменом* и указывает на местонахождение этого почтового ящика. Следует отметить, что носителем адреса электронной почты вовсе не является конечный пункт доставки, т. е. компьютер, на котором вы просматриваете почту. Каждый пользователь может зарегистрировать несколько адресов на разных почтовых серверах. Сегодня существуют сотни серверов, которые бесплатно предоставляют услуги электронной почты, так что любой пользователь Интернета может

свободно зарегистрировать почтовые ящики в разных частях света. Поэтому адрес электронной почты определяет не адрес вашего домашнего компьютера, а адрес сервера, на котором Вы будете получать почту.

Предположим, что у некоего пользователя зарегистрировано два адреса: *dasha@ny-server.com* и *dasha@almaty-server.ru*. Из каждого ящика он получает письма на свой домашний компьютер, но одно он копирует с сервера в Нью-Йорке, а другое - с сервера в Алматы. В связи с этим говорят о направленной адресации электронной почты, имея в виду, что адрес определяет маршрут доставки.

Два сервиса - два протокола.

Электронная почта построена по принципу клиент-серверной архитектуры. Пользователь общается с клиентской программой, которая, в свою очередь, общается с почтовым сервером. Очевидно, что процедура отправки и получения почты в разной степени требует идентификации личности, причем здесь просматривается полная аналогия с обычной почтой: вам необходим ключ для того, чтобы забрать письмо из своего почтового ящика. Ключ в данном случае есть средство авторизации, удостоверяющее, что вы имеете право получить почту. При отправлении письма авторизация необязательна. В связи с тем, что отправка и получение требуют разной степени идентификации личности, существуют и два разных протокола - на отправку и на прием писем. Несмотря на то, что при отправлении письма авторизация необязательна, сегодня большинство провайдеров предпочитает ее вводить из соображений безопасности и борьбы со спамом.

Для передачи писем используются протокол SMTP (Simple Mail Transfer Protocol - простой протокол пересылки почты) и соответственно SMTP-серверы. Чаще всего отправка почты происходит с почтового сервера вашего интернет-провайдера, хотя, в принципе, это не обязательно. Обычно SMTP-серверы не требуют идентификации, поэтому вы можете отправить письмо с любого такого сервера. Для приема почтовых сообщений в настоящее время наиболее часто используется протокол POP3 (Post Office Protocol - протокол почтового офиса), который забирает почту только конкретного пользователя с сервера и потому требует авторизации прав доступа к ящику.

Основные функции почтовых клиентов.

Базовые функции всех почтовых клиентов - это прием сообщений, обеспечение их просмотра в автономном режиме, сортировка сообщений, автоматизация создания ответных сообщений и поддержка адресной книги.

Отправление и доставка письма.

После того как вы подключились к Интернету, клиентская программа соединяется с почтовым сервером и передает ему почтовый адрес отправителя и текст самого сообщения. При отправке почты клиентская программа взаимодействует с сервером исходящей почты, т. е. SMTP-сервером, по протоколу SMTP, подключаясь к порту с номером 25. Имя SMTP-сервера, которому передается почта, должно быть заранее введено в клиентскую

почтовую программу при ее настройке. Рассмотрим этот процесс более подробно. После того как письмо доставлено на сервер отправителя, SMTP-сервер отправителя должен связаться с сервером получателя. Для этого ему необходимо знать IP-адрес SMTP-сервера, получателя. Чтобы узнать этот адрес, он обращается к DNS-серверу и запрашивает IP-адрес SMTP-сервера получателя. Если по какой-либо причине SMTP-сервер отправителя не может связаться с SMTP-сервером получателя, послание встает в очередь для отправки. Обычно через каждые 15 минут будет производиться попытка переслать письмо из очереди. Примерно через четыре часа отправителю будет послано сообщение о проблеме, а спустя пять дней большинство серверов прекращают попытки и возвращают отправителю недоставленную почту, так что бесследно письмо исчезнуть не может. Если посмотреть на протокол общения почтового клиента с SMTP-сервером, можно убедиться, что он достаточно прост. Здесь используются интуитивно понятные команды: типа HELO (представляюсь), MAIL FROM (определяю отправителя), RCPT TO (определяю получателя) и т. п.

Как только SMTP-серверу отправителя удастся связаться с SMTP-сервером получателя, он передает послание. Сервер получателя определяет, существует ли пользователь с именем указанным в адресе и передает послание в его ящик. После того как почта оказалась на сервере, получатель имеет возможность скачать ее оттуда. При получении почты, накопившейся в вашем почтовом ящике, клиентская программа взаимодействует с POP3-сервером по протоколу POP3. Поскольку при входе в Интернет вы сообщаете свой пароль и логин, то обычно при получении почты дополнительного введения этих данных не требуется. Ваш почтовый клиент связывается с POP3-сервером и передает команды, которые определяют передачу копий посланий электронной почты на локальную машину.

Как правило, после получения почты сообщения на сервере удаляются, если в почтовом клиенте не определено обратное. Как и SMTP-сервер, POP3-сервер при «разговоре» с почтовым клиентом получателя использует определенный набор интуитивно понятных команд: USER (введите ID пользователя), PASS (введите пароль), QUIT (покиньте POP3-сервер), DELE (удалите сообщение). Следует отметить, что если SMTP-протокол относительно прост и всего лишь передает на сервер содержимое письма с адресами получателей, то работа с POP3-сервером гораздо более сложная, он позволяет осуществлять частичное скачивание писем, предоставляет возможность отказаться от больших посланий. Для приема почты может использоваться и более современный протокол - IMAP, который дает возможность выборочно копировать пришедшие вам письма с почтового сервера на ваш компьютер, работать с папками и т. д.

Когда электронное письмо не доходит до адресата, это может произойти по нескольким причинам. Возможно, сервер получателя не принимает письма от данного отправителя. Кроме того, послание будет отправлено назад, если

сервер перегружен и не может обработать запрос. Даже если письмо не отвергается сервером получателя, оно может вернуться, если, например, вы ошиблись при написании имени получателя - система сообщит, что на сервере не существует получателя с таким именем. И, наконец, может возникнуть ситуация, когда получатель с указанным именем хотя и существует, но его почтовый ящик переполнен.

Передача специальных символов и вложенных сообщений.

Протокол SMTP позволяет передавать не все символы ASCII, а только те из них, которые характерны для английского алфавита и определяются базовой таблицей кодирования ASCII. Как известно, ASCII определяет две таблицы кодирования: базовую семибитную (символы с номерами 0—127) и расширенную восьмибитную (128—255).

Символы кириллицы относятся к расширенной таблице кодирования, и для протокола SMTP они являются специальными символами. При передаче специальных символов обычно используется стандарт MIME (Multipurpose Internet Mail Extensions - многоцелевые расширения почтового стандарта Интернета), который позволяет передавать специальные символы, а также кодировать нетекстовую информацию для передачи по электронной почте, например графические файлы, которые вы помещаете во вложении к письму.

Изначально электронная почта не предназначалась для пересылки файлов, поэтому для того, чтобы переслать произвольный файл, необходима его перекодировка в аналог текстового сообщения. Стандарт MIME *позволяет определить по заголовку файла, какой тип данных он содержит* (текст, графика, музыка и т. д.), и предоставляет единый метод кодирования вложенных файлов. Когда используются стандарты MIME, то проблем с кодировкой кириллицы обычно не возникает.

Контрольные вопросы

1. Базовые функции всех почтовых клиентов.
2. Как осуществляется доставка и отправление письма?
3. Передача специальных символов и вложенных сообщений.
4. Назовите несколько причин, когда электронное письмо не доходит до адресата.

Лекция №10. Механизм работы сервиса FTP. FTP – клиенты

Цель лекции: ознакомление с основными назначениями FTP и механизмом работы сервиса FTP.

Содержание лекции: протокол передачи файлов FTP (FILE TRANSFER PROTOCOL), формат адреса FTP-ресурса, доступ с помощью ftp.exe.

FTP (File Transfer Protocol - протокол передачи файлов) был реализован в 1971 году и активно используется по сей день. FTP базируется на протоколах базового уровня TCP/IP и представляет собой способ для обмена файлами между компьютерами в Интернете.

Основное назначение FTP - пересылать (копировать, передавать) файлы в Интернете с удаленного компьютера на локальный (Download) и с локального компьютера на удаленный (Upload). С помощью FTP-протокола можно также удалять и переименовывать свои файлы на удаленном компьютере. Нужно отметить, что в отличие от WWW-серверов, которые предоставляют информацию только для чтения, FTP-серверы позволяют не только скачивать информацию, но и добавлять информацию на сервер. В частности, чтобы передать файлы, из которых конструируются WEB-страницы, чаще всего пользуются именно FTP-протоколом. Естественно, что права изменять информацию на сервере должны предоставляться только владельцам данной информации. Права пользователя при обращении к FTP-серверу регламентируются логином и паролем, но это не означает, что в Сети нет серверов, доступных широкой аудитории. На многие серверы можно получить доступ под условным именем (например, anonymous или guest), и в этом случае вам будет доступен ограниченный набор файлов и, как правило, только для считывания.

Протокол FTP применяется давно и на первый взгляд предельно прост. Однако эта простота кажущаяся и многие начинают испытывать проблемы с установлением FTP-соединения, особенно когда сервер или клиент находятся за брандмауэром или NAT'ом. Поэтому сегодня мы поговорим об особенностях работы протокола FTP в различных режимах.

Протокол FTP является старейшим сетевым протоколом (создан в 1971 году), но, тем не менее, широко используется по сей день. Важной особенностью протокола является то, что он использует несколько соединений: одно для управляющих команд, остальные для данных. Причем соединений для передачи данных может открываться несколько, в каждом из которых файлы могут передаваться в обоих направлениях. Именно с этой особенностью и связан ряд проблем.

В зависимости от способа установления соединения для передачи данных различают активный и пассивный режимы работы FTP. В активном режиме сервер сам устанавливает соединение передачи данных к клиенту, в пассивном наоборот. Рассмотрим эти режимы более подробно.

Формат адреса FTP-ресурса.

В общем случае формат URL FTP-ресурса имеет вид:

ftp://<user>:<password>@<host>:<port>/<url-path>

где:

- user - имя пользователя (логин);
- password - пароль;
- host - доменное имя или IP-адрес сервера;
- url-path - путь к файлу.

Обычно порт не указывают, в этом случае используется стандартный - 21. Для обращения к серверам общего доступа в качестве имени пользователя достаточно указать anonymous, а в качестве пароля - свой почтовый адрес. Для анонимного FTP в URL принят упрощенный синтаксис: ftp: //<host>/<url-path>, т. е. при отсутствии имени автоматически будет вставлено anonymous.

FTP-клиенты.

Служба FTP основана на клиент-серверной архитектуре. На клиентском компьютере запускается программа-клиент, которая соединяется с сервером и передает или получает файлы. Доступ на FTP-сервер может осуществляться с помощью FTP-клиентов разных классов, а именно:

- консольный клиент - программа ftp.exe, поставляемая вместе с операционной системой Windows,
- FTP-клиент, встроенный в браузер (например, Internet Explorer или Netscape Navigator);
- FTP-клиент, встроенный в файловый менеджер (в частности, FAR или Windows Commander);
- FTP-клиент, встроенный в HTML-редактор (HomeSite, Dreamweaver или FrontPage);
- специализированные программы (например, CuteFTP или WS FTP).

Доступ с помощью ftp.exe.

Ftp.exe- это консольный клиент, который не имеет графического интерфейса и наименее удобен. Мы приводим его только потому, что он полезен для уяснения вопроса - «как это работает».

Запуск ftp.exe вызывает интерактивный сеанс: программа предоставит вам командную строку и будет ждать запросов. Для ведения диалога существует несколько десятков команд, некоторые из них приведены ниже:

- cd [удаленная_директория] - сменить каталог;
- mkdir [имя_директории] - создать каталог на удаленной машине;
- get [удаленный_файл] [локальный_файл] - передать копию удаленного файла на ваш компьютер. Если имя локального файла не было задано, то оно совпадает с именем удаленного файла;
- reget [удаленный_файл] [локальный_файл] - докачать удаленный файл в том случае, когда часть его уже есть на локальной машине. Команда особенно полезна для получения больших файлов при возможных разрывах соединения;
- put [локальный_файл] [удаленный_файл] - переслать файл на удаленную систему. Если имя удаленного файла не указано, то оно совпадает с именем на локальной системе;

- delete [удаленный_файл] - стереть файл с именем удаленный_файл;
- close - оборвать FTP-сеанс с удаленным сервером и вернуть к командному интерпретатору;
- bye - закончить работу с FTP-сервером (приводит к выходу из интерпретатора).

Полный список команд и дополнительные сведения можно получить в соответствующем RFC (Аббревиатура RFC означает Request for Comments (запрос на пояснения) - серия документов, которая публикуется организацией Internet Engineering Task Force (IETF) и затрагивает самые различные вопросы развития Интернета и его технологий).

Следует отметить, что FTP-протокол предоставляет долговременное интерактивное соединение, позволяет вести диалог и поддерживать соединение, пока не будет дана команда о прекращении сеанса связи. FTP-протокол понимает два типа файлов - текстовые и двоичные, причем все нетекстовые файлы воспринимаются как двоичные. Работа через FTP-соединение осуществляется так же, как обмен текстовыми сообщениями (служебными командами), и непосредственно теми файлами, которые перемещаются между двумя компьютерами. Консольная форма работы менее удобна для пользователей, но позволяет лучше понимать, какие команды выполняет компьютер.

Служба архивов FTP.

FTP-архивы являются одними из основных информационных ресурсов Internet. Фактически, это распределенный депозитарий текстов, программ, фотографий и прочей информации, хранящейся в виде файлов на различных компьютерах во всем мире.

Информация в FTP-архивах разделена, в основном, на три категории: Защищенная информация, режим доступа к которой определяется ее владельцами и разрешается по специальному соглашению с потребителем. К этому виду ресурсов относятся коммерческие архивы (к примеру, коммерческие версии программ в архивах *ftp.microsoft.com*), закрытые национальные и международные некоммерческие ресурсы (например, работы по международным проектам CES или IAEA), частная некоммерческая информация со специальными режимами доступа (например, частные благотворительные фонды). Информационные ресурсы ограниченного использования, к которым относятся, например, программы класса shareware. В данный класс могут входить ресурсы ограниченного времени использования или ограниченного времени действия. Свободно распространяемые информационные ресурсы или freeware, если речь идет о программном обеспечении. К этим ресурсам относится все, что можно свободно получить по сети без специальной регистрации. Это может быть документация, программы или что-либо еще. Следует отметить, что свободно распространяемое программное обеспечение не имеет сертификата качества, но его разработчики открыты для обмена опытом.

Из выше перечисленных ресурсов наиболее интересными являются две последних категории, которые, как правило, оформлены в виде FTP-архивов.

Технология FTP была разработана в рамках проекта ARPA и предназначена для обмена большими объемами информации между машинами с различной архитектурой. Главным в проекте было обеспечение надежной передачи, поэтому с современной точки зрения FTP кажется перегруженным излишними редко используемыми возможностями. Стержень технологии составляет FTP-протокол.

Контрольные вопросы

1. Основное назначение FTP.
2. Назовите типы файлов, который понимает протокол FTP.
3. Как осуществляется доступ с помощью ftp.exe.
4. Формат адреса FTP-ресурса. FTP-клиенты.
5. Служба архивов FTP. Технология FTP

Лекция №11. Защита информации в Интернет. Криптография и Интернет.

Цель лекции: познакомиться с технологией шифрования (криптография) важной информации для защиты ее от изменений и неавторизованного доступа и изучить способы закодирования исходного текста (алгоритмы шифрования)

Содержание лекции: криптография. шифрование с симметричным и асимметричным ключом.

Для обеспечения секретности применяется шифрование, или криптография, позволяющая трансформировать данные в зашифрованную форму, из которой извлечь исходную информацию можно только при наличии ключа. В основе шифрования лежат два основных понятия: алгоритм и ключ. Алгоритм – это способ закодировать исходный текст, в результате чего получается зашифрованное послание. Зашифрованное послание может быть интерпретировано только с помощью ключа. Очевидно, чтобы зашифровать послание, достаточно алгоритма. Однако использование ключа при шифровании предоставляет два существенных преимущества. Во-первых, можно использовать один алгоритм с разными ключами для отправки посланий разным адресатам. Во-вторых, если секретность ключа будет нарушена, его можно легко заменить, не меняя при этом алгоритм шифрования. Таким образом, безопасность систем шифрования зависит от секретности используемого ключа, а не от секретности алгоритма шифрования. Многие алгоритмы шифрования являются общедоступными.

Количество возможных ключей для данного алгоритма зависит от числа бит в ключе. Например, 8-битный ключ допускает 256 (2⁸) комбинаций ключей. Чем больше возможных комбинаций ключей, тем труднее подобрать ключ, тем надежнее зашифровано послание. Так, например, если использовать 128-битный ключ, то необходимо будет перебрать $2^{128} = 10^{40}$ ключей, что в настоящее время не под силу даже самым мощным компьютерам. Важно отметить, что возрастающая производительность техники приводит к уменьшению времени, требующегося для вскрытия ключей, и системам обеспечения безопасности приходится использовать все более длинные ключи, что, в свою очередь, ведет к увеличению затрат на шифрование. Поскольку столь важное место в системах шифрования уделяется секретности ключа, то основной проблемой подобных систем является генерация и передача ключа. Существуют две основные схемы шифрования: симметричное шифрование (его также иногда называют традиционным или шифрованием с секретным ключом) и шифрование с открытым ключом (иногда этот тип шифрования называют асимметричным). При симметричном шифровании отправитель и получатель владеют одним и тем же ключом (секретным), с помощью которого они могут зашифровывать и расшифровывать данные. При симметричном шифровании используются ключи небольшой длины, поэтому можно быстро шифровать большие объемы данных. Симметричное шифрование используется, например, некоторыми банками в сетях банкоматов. Однако симметричное шифрование обладает несколькими недостатками. Во-первых, очень сложно найти безопасный механизм, при помощи которого отправитель и получатель смогут тайно от других выбрать ключ. Возникает проблема безопасного распространения секретных ключей. Во-вторых, для каждого адресата необходимо хранить отдельный секретный ключ. В третьих, в схеме симметричного шифрования невозможно гарантировать личность отправителя, поскольку два пользователя владеют одним ключом. В схеме шифрования с открытым ключом для шифрования послания используются два различных ключа. При помощи одного из них послание зашифровывается, а при помощи второго - расшифровывается. Таким образом, требуемой безопасности можно добиться, сделав первый ключ общедоступным (открытым), а второй ключ хранить только у получателя (закрытый, личный ключ). В таком случае любой пользователь может зашифровать послание при помощи открытого ключа, но расшифровать послание способен только обладатель личного ключа. При этом нет необходимости заботиться о безопасности передачи открытого ключа, а для того чтобы пользователи могли обмениваться секретными сообщениями, достаточно наличия у них открытых ключей друг друга. Недостатком асимметричного шифрования является необходимость использования более длинных, чем при симметричном шифровании, ключей для обеспечения эквивалентного уровня безопасности, что сказывается на вычислительных ресурсах, требуемых для организации процесса шифрования.

Криптография (от греч. *Cryptos*- тайный) - это наука и технология шифрования важной информации для защиты ее от изменений и неавторизованного доступа. Криптография служит не только для перевода текстов в нечитаемую зашифрованную форму, но позволяет также решить задачи аутентификации и идентификации пользователей при работе их в Сети.

Криптография является основой безопасных коммуникаций. Следует отметить, что в Сети мы общаемся не только с людьми, но и с различными службами. Например, когда мы собираемся скачать с какого-либо сервера программу, нам тоже важно знать, что данный сервер принадлежит именно фирме-разработчику, а не фирме-пирату, которая может заразить ваш компьютер вирусом. Сообщение можно зашифровать и частным, и публичным ключом, а расшифровать только вторым ключом из пары. То есть сообщение, зашифрованное частным ключом, можно расшифровать только публичным, и наоборот. Частный ключ известен только владельцу, и его нельзя никому передавать, в то время как публичный ключ распространяется открыто всем корреспондентам.

Пару ключей - частный и публичный - можно использовать как для решения задач аутентификации, так и секретности (конфиденциальности).

Для того, чтобы обеспечить конфиденциальный обмен сообщениями в переписке двух лиц, необходимо иметь две пары ключей.

При шифровании с помощью ключевой пары у вас нет необходимости рассылать свой публичный ключ всем корреспондентам. Гораздо удобнее выложить этот ключ в Сети на сервере с открытым доступом. Тогда каждый может скачать данный ключ и послать вам секретное сообщение, которое не прочтет никто кроме вас.

Шифрование с симметричным и асимметричным ключом.

Следует отметить, что алгоритм асимметричного шифрования сильно проигрывает алгоритму симметричного шифрования в плане эффективности, поэтому многие существующие системы шифрования используют комбинацию асимметричной и традиционной симметричной систем шифрования. Шифрование с открытым ключом используется для передачи симметричного ключа, который служит непосредственно для шифрования передаваемой информации.

Контрольные вопросы

1. Что произойдет если секретность ключа будет нарушена?
2. Что является основой безопасных коммуникаций?
3. Шифрование с симметричным и асимметричным ключом.

Лекция №12. Электронная цифровая подпись

Цель лекции: иметь представление о методе аутентификации отправителя или автора подписи, подтверждающий, что содержание документа не было изменено.

Содержание лекции: односторонняя хеш-функция, цифровая подпись.

Для того, чтобы объяснить механизм цифровой подписи, необходимо ввести понятие односторонней хеш-функции. *Односторонняя хеш-функция* - это функция, которая преобразует исходное послание любой длины в строку символов ограниченной длины - так называемый дайджест послания. При использовании 16-байтной хеш-функции вы на выходе получите запись длиной 16 байт.

Даже если послание, безопасность которого мы хотим обеспечить, должным образом зашифровано, все равно остается возможность модификации исходного сообщения или подмены этого сообщения другим. Одним из путей решения этой проблемы является передача пользователем получателю краткого представления передаваемого сообщения. Подобное краткое представление называют контрольной суммой, или дайджестом сообщения. Контрольные суммы используются при создании резюме фиксированной длины для представления длинных сообщений. Алгоритмы расчета контрольных сумм разработаны так, чтобы они были по возможности уникальны для каждого сообщения. Таким образом устраняется возможность подмены одного сообщения другим с сохранением того же самого значения контрольной суммы. Однако, при использовании контрольных сумм возникает проблема передачи их получателю. Одним из возможных путей ее решения является включение контрольной суммы в так называемую электронную подпись. При помощи электронной подписи получатель может убедиться в том, что полученное им сообщение послано не сторонним лицом, а имеющим определенные права отправителем. Электронные цифровые подписи создаются шифрованием контрольной суммы и дополнительной информации при помощи личного ключа отправителя. Таким образом, кто угодно может расшифровать подпись, используя открытый ключ, но корректно создать подпись может только владелец личного ключа. Для защиты от перехвата и повторного использования подпись включает в себя уникальное число - порядковый номер. более подробно о электронной цифровой подписи (ЭЦП) читайте в разделе курса ОСВМ Аутентификация информации. С 2012 года в Казахстане функционирует свой Национальный удостоверяющий центр, в котором граждане и организации Казахстана могут получить свои закрытые ключи и программные пакеты для формирования своей цифровой подписи для ведения электронных юридических операций с ее использованием. Необходимо,

однако, заметить, что данная система еще очень сырая и позволяет злоумышленникам легко воспользоваться чужой электронной подписью для совершения подложных сделок. Это происходит по следующим причинам: выдача файлов электронных подписей совершается «вручную», то есть оператор, выдающий подпись всегда может иметь ее копию на своем USB-носителе, пароль выдается на всех один (!!!) - 123456. При попытке смены пароля файл электронной подписи записанный на стираемом носителе, фатально повреждается, и, если владелец подписи не сделал предварительно копию, то он лишается возможности подписывать документы до получения новой подписи из НУЦ, ЦОН или у программистов районного налогового комитета.

На сегодня уже известно множество фиктивных сделок с недвижимостью и иным имуществом, а также незаконным получением документов с помощью чужой электронной подписи. Поэтому, единственно разумным действием, предохраняющим человека от лишения его собственности.

Хеширование – это одностороннее, т. е. необратимое, в отличие от шифрования, преобразование. По дайджесту послания нельзя восстановить исходное сообщение, но можно его однозначно идентифицировать.

Рассмотрим конкретный пример. Пусть А, применив хеш-функцию, получил дайджест в виде некоего кода (*код1*). Затем, используя свой личный ключ, А зашифровывает дайджест, который и становится аналогом подписи данного документа. Далее А отправляет открытый текст послания и цифровую подпись В. Абонент В расшифровывает подпись публичным ключом, который он получил от А и убеждается, что письмо действительно от А. Таким образом, аутентификация состоялась. Затем необходимо удостовериться, что послание не было изменено в процессе доставки. В результате расшифровки В получает дайджест послания - *код1*. Затем В, используя ту же хеш-функцию, что и А, хеширует текст послания и получает дайджест в виде некоего *кода2*. Если *код1* и *код2* совпадают, то В делает вывод о том, что текст послания не был изменен на этапе доставки.

Таким образом, *цифровая подпись* или *электронная подпись* – это метод аутентификации отправителя или автора подписи, подтверждающий, что содержание документа не было изменено. Цифровая подпись может быть поставлена как в зашифрованном, так и в открытом послании.

Контрольные вопросы

1. Как объясняется механизм цифровой подписи?
2. Как создаются электронные цифровые подписи?
3. Дайте определение понятию «хеширование».

Лекция №13. Идентификация пользователей в Интернет. Механизм Cookies

Цель лекции: ознакомиться с механизмом работы cookies.

Содержание лекции: применение cookies. Преимущества и недостатки технологий на основе cookies

Как работают cookies.

Cookies - небольшие текстовые записи, которые сохраняются браузером на компьютере пользователя по запросу сервера. Эта текстовая информация хранится на компьютере пользователя до тех пор, пока она не будет удалена. *Cookies*, в которых хранится история обращений пользователя к WEB-серверу, посылают многие серверы. Это позволяет четче отслеживать характер обращений на сайты и предоставлять пользователям сайтов и их владельцам дополнительные сервисы. Например, владельцам сайта важно знать, сколько человек просмотрело их сайт, кто пришел на сайт впервые, а кто - повторно. Данная задача легко может быть решена с помощью технологии cookies. Программное обеспечение WEB-сайта может сгенерировать уникальный ID-номер каждого посетителя сайта, отослать его пользователю, и он будет храниться на пользовательском компьютере. При повторном посещении эти данные считываются, и WEB-сайт узнает своего посетителя. В дальнейшем по мере посещения данной странички информация в ваших cookies может изменяться. При этом следует отметить, что сервер может ассоциировать разные cookies с различными частями WEB-сайта. Cookies позволяют серверам настроить работу узла таким образом, чтобы при следующем посещении пользователя могли обслужить индивидуально. Например, не показывать ему несколько раз одну и ту же рекламу или позволить ему просматривать страницы с учетом его предыдущих настроек. Сервер не помнит о предыдущих обращениях пользователя и введенных им настройках. Использование технологии cookies позволяет запоминать пользователя, считывая из файлов cookies его имя, персональные настройки и т. д. Различают постоянные и временные файлы cookies. Временные файлы cookies хранятся на компьютере пользователя только в течение текущего сеанса браузера, а постоянные сохраняются на компьютере пользователя и считываются WEB-узлом при повторном посещении. С помощью технологии cookies владелец сайта может проследить за перемещениями пользователя по сайту, получить сведения о его предпочтениях и использовать их в рекламных целях. В cookies могут храниться такие сведения о пользователе, как определенные индивидуальные настройки, выбранные им при посещении узла, а также личные идентификационные данные, которые были внесены пользователем в соответствующие регистрационные формы. Следует отметить, что cookies, полученные с сервера, могут содержать только те

персональные данные, которые пользователь сам послал данному серверу. Для хранения cookies, браузер должен вести определенную базу данных на компьютере пользователя. Если вы пользуетесь Microsoft Internet Explorer, то на вашем компьютере эта база данных представлена папкой с файлами, имена которых состоят из имен пользователя и сервера, и может иметь, например, следующий вид: khan@spylog[2].txt.

Например, файл khan@spylog[2].txt имеет вид:

```
VISIDSVWDOBAOFOTIYBYDIPspylog.com/10241362503808295671022  
16316196829530892*
```

В данной текстовой строке содержится имя сайта spylog.com, который посещен, и служебная информация. Обычно браузер имеет ограничения на максимальный размер cookies (не более 4 Кбайт), общее количество cookies (не более 500) и максимальное количество cookies, хранимых для одного сервера (около 20).

Cookies имеют небольшой размер и не могут содержать большой объем данных. Обычно информация о том, какие сайты посещал тот или иной пользователь, какая реклама ему уже показывалась, список заказанных им товаров и т. п. хранится на WEB-сервере, а cookies используются для поиска этой информации. Технология cookies позволяет делать персонализированные страницы. Например, если вы задали свои координаты при регистрации на некотором сайте, который показывает погоду, то он сможет сообщать вам погоду специально для вашего региона.

Сайты электронной коммерции могут использовать cookies для создания покупательской корзины. Информация обо всех продуктах, которые вы добавляете в вашу корзину, хранится в базе данных WEB-сервера вместе с вашим ID. Когда вы хотите получить счет, серверу известны все заказанные вами товары, поэтому ему несложно получить из базы данных информацию об их стоимости и выставить вам счет. При посещении WEB-сайтов с чужого компьютера следует помнить, что они могут внести на него нежелательные cookies (т. е. ассоциированные с настройками гостя, а не хозяина компьютера). В этом случае можно запретить размещение cookies или удалить их по окончании работы.

Удаление cookies безопасно с точки зрения работоспособности вашего компьютера. Однако следует иметь в виду, что если вы удалите cookies, принадлежащие сайту, который вы часто посещаете, то вместе с этими cookies могут быть потеряны ваши индивидуальные настройки.

Какие проблемы возникают cookies?

Если несколько человек обращаются на некоторый сайт, пользуясь одним ПК, то посещаемый WEB-сервер будет принимать разных посетителей за одно лицо и путать их персональные настройки.

Если вы используете для доступа в Интернет разные компьютеры (один в офисе, другой дома, третий в командировке), то в каждом случае на всех трех компьютерах будут разные cookie-файлы. Кроме того, cookies могут

стираться (например, при переустановке системы). Поэтому если вы удалите cookies и обратитесь на ранее посещенный сайт, он воспримет вас как нового пользователя.

Всех этих проблем можно избежать, регистрируясь на сайте. В этом случае сервер хранит ваши cookies вместе с логином и паролем и вне зависимости от того, с какого компьютера вы осуществляете доступ, вам будет предложена правильная информация.

Поскольку cookies хранятся в текстовом виде и пользователь может их удалить или видоизменить, разработчики WEB-приложения обычно не хранят в cookies важной информации и организуют работу WEB-сервера так, чтобы оно могло функционировать и при запрете на помещение cookies на диск пользователя. Наличие cookies на диске позволяет проанализировать, какие сайты посещал пользователь, а иногда также и персональные данные, сообщенные пользователем на конкретный сайт.

Контрольные вопросы

1. Как работают cookies?
2. Какие проблемы возникают у cookies?
3. Что потребуются сделать для хранения cookies?
4. С какой точки зрения безопасно удаление cookies?

Лекция №14. Технологии создания клиентских приложений Интернет.

Цель лекции: ознакомиться с основными технологиями клиентских приложений Интернет.

Содержание лекции: dynamic HTML. Macromedia Flash.

До появления версий 4.0 Internet Explorer и Netscape Navigator сценарии могли изменять содержание и внешний вид страниц, только используя метод write.document при загрузке страницы.

Любой сценарий, выполняемый позже, может посылать команды браузеру, такие как запрос на загрузку новой страницы, вывод окна предупреждения или изменение значения поля формы, но как только страница загружена, сценарий не может изменить HTML-код. Как же сделать WEB-страницы интерактивными, если их невозможно изменить в ответ на действия пользователя? Динамический HTML, реализованный в Netscape от 4 версии и в большей степени в Internet Explorer от 4 версии, устраняет эти ограничения. С использованием динамического HTML сценарии могут вставлять блоки HTML, удалять и заменять их или изменять свойства объектов после

отображения страницы на экране. Браузер автоматически обновляет новые свойства и (или) новый HTML-код. Динамический HTML строится на двух принципах.

1) Объектная модель документа (DOM) описывает способ организации и названия объектов в браузере, а также определяет, какие объекты и свойства могут быть изменены и какие значения они могут принимать.

2) Событийная модель описывает способ передачи управления сценариям - какие действия посылают сценарии на выполнение.

Macromedia Flash.

World Wide WEB (Всемирная паутина) развивается очень стремительно. Традиционные WEB-сайты (WEB-sites) со статическими информационными страницами уже не привлекают как прежде внимание конечных пользователей. Поэтому WEB-дизайнеры и разработчики пытаются придать своим страницам неотразимый вид, вводя графику и файлы анимационного формата GIF. Несмотря на то, что применение графики в традиционных форматах сжатия изображений GIF, JPEG и PNG придает сайту некоторый эффект, тем не менее, загрузка таких страниц может отнять много времени из-за больших размеров графических файлов. Альтернативным инструментом создания анимации является программный продукт - Macromedia Flash, который поможет добиться поставленных при WEB-разработке целей. По сравнению с другими графическими приложениями, предназначенными для публикации результатов в WEB или вывода в файл, Flash обладает многими преимуществами, такими, как: применение векторной графики, которая в действительности представляет собой ряд математических формул, описывающих размер, цвет и местоположение формы. Векторная графика состоит из линий и кривых, тогда как растровая - из небольших точек растра, или пикселей. Применение математических формул является главным преимуществом отображения графики и сохранения небольшого размера файлов, что особенно важно для WEB. Возможность сжатия анимации, графики и звука. Такое сжатие файла обуславливает свойство потоковой передачи, которое позволяет отображать сайт в WEB-браузере до полной загрузки его содержимого. Это означает, что одна векторная графика сайта отображается, в то время как другая графика, звук и анимация все еще загружаются. Интерактивность во Flash может состоять из самых разных свойств. Помимо неограниченных возможностей применения простой графики, анимации или крупных заголовков, в любой объект или область сайта можно вводить полноценные навигационные свойства: кнопки, меню или фрагменты анимации, с помощью которых конечный пользователь перемещается по сайту. А задавая действия (actions), можно создавать интерактивные фильмы. Для создания интерактивных элементов управления во Flash предоставляется язык создания сценариев - ActionScript. Изображения и текст, разработанные во Flash, всегда отображаются ясно и четко, что опять же обусловлено векторной графикой. Flash помогает пользователю рисовать.

В частности, Flash может распознавать основные геометрические формы в процессе их создания. При этом она заменяет небрежно нарисованную форму правильной геометрической формой, которую, как предполагается, пользователь пытается нарисовать. Например, если быстро нарисовать овал или окружность, создав круглую форму, Flash сделает эту форму более плавной и круглой, доведя ее до идеального овала или окружности. Кроме того Flash помогает создавать прямые и плавные линии.

В состав Flash входит инструмент, который создает основу для большинства качественных Flash-фильмов. Этот инструмент – ActionScript – событийно-управляемый язык, встроенный во Flash. ActionScript делает ваши страницы интерактивными. Вы можете реагировать на события с мышки или с клавиатуры, можете выполнить какие-либо действия при проигрывании определенного кадра.

Для того, чтобы овладеть ActionScript в полной мере, желательно уже иметь опыт программирования (предпочтительно на C++, JavaScript). Однако одним из достоинств языка Flash является то, что вам не нужно быть профессионалом во Flash, или полностью знать ActionScript, чтобы писать на нем качественный код. Вы можете использовать лишь те возможности языка, которые сочтете необходимыми для своей работы.

Контрольные вопросы

1. Какие технологии существуют для создания клиентских приложений Интернет?
2. Альтернативный пакет для создания анимации.
3. Какой инструмент входит в состав Flash?
4. Достоинство и недостатки языка Flash.

Лекция №15. Перспективы развития WEB-технологий и образовательный процесс

Цель лекции: познакомиться с перспективами развития WEB-технологий для решения многообразных образовательных задач.

Содержание лекции: перспективы развития WEB-технологий. Типологии WEB-приложений в образовании. Мобильные устройства. Характеристика средств Интернета для онлайн-обучения.

WEB-технологии являются неотъемлемым атрибутом информационного общества и его базовым основанием. Эти технологии, о которых не слышали в конце прошлого века, работают практически во всех областях экономики, науки, культуры, социальных преобразований. WEB-технология в настоящее

время соединяет десятки тысяч компьютерных локальных, региональных, федеральных сетей и миллионы пользователей во всем мире. При этом сетью объединены компьютеры тысяч различных типов, оснащенных самым разным программным обеспечением. Интернет образует как бы ядро, обеспечивающее связь различных информационных сетей, принадлежащих различным учреждениям во всем мире.

Если ранее сеть использовалась исключительно в качестве среды передачи файлов и сообщений электронной почты, то сегодня решаются более сложные задачи распределенного доступа к ресурсам. Существуют различные оболочки, поддерживающие функции сетевого поиска и доступа к распределенным информационным ресурсам, электронным банкам данных.

Сеть Интернет когда-то служила исключительно оборонным, исследовательским и образовательным целям. Но постепенно она становилась все более популярной в деловом мире.

Компании привлекают быстрота, дешевая глобальная связь, удобство для проведения совместных работ, доступные программы, уникальная база данных сети. Они рассматривают глобальную сеть как дополнение к своим собственным локальным сетям.

В настоящее время в глобальной сети используются практически все известные линии связи: от низкоскоростных телефонных линий до высокоскоростных цифровых спутниковых каналов.

Фактически Интернет состоит из множества локальных и глобальных сетей, принадлежащих различным компаниям и предприятиям, связанных между собой различными линиями связи. Интернет можно представить в виде мозаики, сложенной из небольших сетей разной величины, которые Интернет предоставляет уникальные возможности дешевой, надежной и конфиденциальной глобальной связи по всему миру. Однако, как говорится, грядущее технологии начинается сегодня. И несомненно то, что многие пионерские разработки, которые сегодня существуют в типе пилотных образцов, вскоре станут доступны массово и придадут Интернету новое содержание. Одна из тенденций хранится в том, что доступ в Интернет постепенно делается широкополосным. Повсеместное распространение высокоскоростного доступа позволит больше расширить доступ, например, к интерактивным формам обучения и развлечений. Поток видео в ближайшем будущем будет повсеместно доступно по сети так же, как сегодня доступен текст. Распространенной практикой станет заказ обожаемого фильма и передачи в комфортное для вас время. Необходимость путешествовать по миру в целях получения воспитания существенно уменьшится. Следующее важное направление, которое кардинально изменит возможности глобальной Сети, - это более широкое распространение мобильных устройств с доступом в Интернет. Мобильные устройства позволят иметь доступ к Интернету в любом месте 24 часа в сутки. Распространение сетей класса 3G позволит иметь почти мгновенный

широкополосный доступ в Сеть с всевозможных смартфонов и коммуникаторов. Сформируется новое поколение людей, которые будут неразрывно связывать свой быт и отдых с Сетью посредством доступа к ней на базе мобильных устройств. Уже сегодня появился новый термин «always on generation» - «поколение, доступное по Сети в любое время».

Мобильные устройства.

Мобильные устройства создадут качественно новую потребность обращения к Сети. Доступность и распространенность беспроводных мобильных устройств приведет к появлению огромного числа людей, которые будут иметь возможность мгновенно (менее минуты) получать доступ в Сеть в нерабочее время. Эта группа людей будет иметь возможность применения мобильных устройств для принятия решения по выбору товаров, услуг в разных областях включая отдых и лечение. Это, в свою очередь, повлечет существенное изменение во взаимоотношениях целого ряда групп: продавец - покупатель, врач - больной и т. д. В перспективе мобильные устройства будут становиться все меньше и начнут имплантироваться под кожу. Внедрение мобильных устройств сможет радикально изменить взаимоотношения в сфере потребления интернет-услуг, что окажет на эту сферу даже большее влияние, чем само появление Интернета. Пользователи смогут иметь доступ к Сети в нерабочее время за счет использования проводного и беспроводного доступа к Сети посредством домашних ПК, интерактивного телевидения, мобильных телефонов, PDA, электронных игрушек, цифровых камер и прочих мобильных устройств.

Развитие мобильных устройств и мобильного бизнеса будет стимулироваться снижением цен на абонентскую плату и на стоимость аппаратной платформы. Новые технологии сделают возможным применение услуг, которые ранее были принципиально невозможны.

Типологии WEB-приложений в образовании.

Под использованием Интернета в образовании понимается применение различных WEB-технологий для решения многообразных образовательных задач: обучения, изучения и управления образовательным процессом.

Весьма распространенным является технологически-ориентированный подход. Так, проведенном ЮНЕСКО в 1995 г. статистическом исследовании опыта применения электронных средств коммуникации в образовании был использован следующий список применяемых коммуникаций: телефон, факс, аудиоконференция, видеоконференция, электронная почта, доступ к базам данных.

Различные сервисы Интернета могут быть использованы для обучения, в том числе в режиме онлайн. Пользование ими требует как от преподавателей, так и от обучаемых различной степени навыков и умений работы в информационной среде.

Таблица 15.1 -Характеристика средств Интернета для онлайнного обучения

Степень виртуальности (уровень навыков)	Средства Интернета
Слабая степень	Использование электронной почты или дискуссионных листов
Средняя степень	Дискуссионные листы и онлайнные материалы лекций, доставляемые через Интернет (WEB)
Сильная степень	Указанное выше плюс интерактивные WEB-пособия, специально созданные для онлайнного курса и взаимодействия с обучаемым (например, производство обучаемыми собственными WEB-страниц)
Экспертный уровень	Указанное выше плюс полностью виртуальная среда, предоставляющая участникам возможности совместной деятельности (типа Multi User Dimension, MUD)

Контрольные вопросы

1. Что понимается под «глобальной сетью»?
2. Из чего состоит Интернет?
3. Преимущества мобильных устройств.
4. Какую роль играет WEB-технология в образовании?

Шнар Ихсанович Имангалиев
Нуркамиля Алдангаровна Дауренбаева

WEB-ТЕХНОЛОГИИ

Конспект лекции для студентов специальностей

5B070300 – Информационные системы, 5B060200 – Информатика

Редактор Н.М.Голева

Специалист по стандартизации Н.К. Молдабекова

Подписано в печать _____.____._____.

Тираж 40 экз.

Объем 3,38 уч.-изд. л.

Формат 60x84 1/16

Бумага типографская №1

Заказ _____. Цена 1690 тенге

Копировально-множительное бюро
некоммерческого акционерного общества
«Алматинский университет энергетики и связи»
050013 Алматы, Байтурсынова, 126