

МНОГОСЛОЙНАЯ МОДЕЛЬ ГИПЕРТЕКСТА

И.С. Демин

Финансовая академия при Правительстве РФ, г.Москва

Представлена членом редколлегии профессором С.В. Мищенко

Ключевые слова и фразы: гипермедиа; гипертекст; гипертекстовые среды.

Аннотация: Рассматриваются особенности, и описываются недостатки гипертекстовой модели, положенной в основу организации информации в сети Интернет. Показывается, что в нынешнем его состоянии Интернет развивается по экстенсивному пути, постепенно исчерпывающем себя. Предложены основные принципы организации дополнительных структурно-семантических слоев гипертекста, предоставляющих принципиально новые возможности его организации. Подобная многослойная модель является развитием модели гипертекста, основанной на языке HTML.

С того времени, как Теодор Нильсен впервые использовал термин «гипертекст», прошло несколько десятков лет, и концепция гипертекста как способа нелинейной организации получила существенное развитие. Было предложено немало моделей гипертекста, некоторые из них, такие как ZOG, HyperCard, Intermedia или HyperTIES, сыграли важную роль в становлении нынешней концепции гипертекста, но сами стали, скорее, достоянием истории. Немногие из моделей продолжают развиваться и ныне, в их числе Xanadu – проект Т. Нильсена.

Наибольшее распространение, безусловно, получила модель гипертекста, основанная на языке HTML, и положенная в основу Интернета. Бурное развитие Интернета привело к тому, что само понятие «гипертекст» стало ему почти синонимично. Объединяя наиболее значительный в мировом масштабе объем информации в гипертекстовой форме, Интернет не мог не повлиять на формирование концепции гипертекста. Любой разговор о гипертекстах неизбежно ведется с оглядкой на Интернет, и любая гипертекстовая система неизбежно будет оцениваться в сравнении со всемирной сетью. Такое состояние дел неизбежно: слишком значительна роль Интернета в нашей жизни, и слишком велики ресурсы, объединяемые глобальной сетью, чтобы можно было не считаться с ее стандартами. Стандарты эти постоянно развиваются, и сейчас перед нами уже не тот Интернет, и не тот гипертекст, что десять лет назад. Современное состояние гипертекстовой среды Интернета, сложилось из результатов усилий многих теоретиков и практических разработчиков. Насколько далеко ушел современный Интернет от первоначальной модели гипертекста?

Часто упоминают о том, что Теодор Нильсен счел идею Интернета несовершенной. В рамках проекта Xanadu он стал развивать иную модель гипертекста, получившую воплощение, например, в гипертекстовой среде CosmicBook. Первоначальная идея Нильсена основывалась на возможности установления произвольных связей между двумя семантически близкими фрагментами текста. Предвосхитившее гипертекст гипотетическое устройство «Мемекс», задуманное Ваннева-

ром Бушем, должно было реализовать семантическую сеть объектов, представлявших собой темы научных изысканий. В. Буш считал, что сетевая организация более адекватна семантике предметной области, нежели существовавшие к тому времени стандартные структурные способы организации информации: систематические и предметные каталоги.

Идея «Мемекса», как и первоначальная идея гипертекста, базировалась на свободе установления связей. Предполагалось, что не только автор, но и любой пользователь с правами доступа может установить связи между любыми двумя фрагментами. У этой идеи, несомненно, была сильная сторона. В этом случае структура связей не устанавливалась одним автором (или группой авторов). Семантика сети представляла собой плод коллективного разума; предполагалось, что пользователь может устанавливать связи между любыми двумя фрагментами, если считает, что они семантически близки.

Очевидно, что ценность такой сети как способа организации научной информации будет проявляться лишь тогда, когда количество связей от каждого фрагмента не превышает разумные пределы, и когда они установлены профессионально. Слово «профессионально» означает здесь, что все авторы подчиняют свои действия некоторому единому семантическому принципу организации, а также то, что они достаточно компетентны, чтобы избегать несущественных ассоциаций. Слишком большое число связей приводит систему в состояние новой энтропии, а непрофессиональные связи нарушают единую семантику сети.

Похожую картину мы видим в среде CosmicBook. Все устанавливаемые связи там являются двунаправленными. Любая ссылка связывает якорь одного фрагмента с якорем другого, и, таким образом, обеспечивает взаимный переход. Активизация ссылки вызывает открытие окна связанного фрагмента, при этом связь графически изображается в виде «луча», соединяющего два якоря.

CosmicBook работает с ограниченными по размеру гипертекстами, называемыми «флайтами» (flight). Интернет же представляет собой, в сущности, единый гигантский гипертекст, на наиболее популярные ресурсы которого установлены сотни и даже тысячи гиперссылок. Если бы ссылки были двунаправленными, такие фрагменты были бы заполнены огромным числом загромождающих якорей, превышающих объем его основного содержания, и многократно увеличивающих трафик сети. При этом такие востребованные фрагменты оказались бы связанными со слишком большим числом других фрагментов, чтобы можно было сделать обоснованный выбор в пользу одного из них.

Кроме того, каждый из разработчиков гипертекстов, слагающих Интернет, основывается на своем понимании семантической близости фрагментов, на своих ассоциативных ценностях, в то время как связи принципиально одинаковы. Возникает непростая для пользователя задача выявления семантики связей, отбора тех из них, которые следует предпочесть для продолжения навигации в контексте решаемых задач.

Отметим, что, по крайней мере, «Мемекс» и задумывался как сеть, организуемая профессионалами, подчиняющими свои действия единой цели. Но попытка создать полностью свободный для установления новых связей гипертекст, в том случае, когда такие связи принципиально одинаковы, обречена на неудачу, поскольку вряд ли можно заставить всех пользователей работать в одном и том же семантическом слое, подчинять свои действия единой цели. Стало быть, сеть со свободно устанавливаемыми связями должна иметь ограничения по составу пользователей (или по их полномочиям). Заметим, что в первоначальных теоретических замыслах гипертекст таких ограничений не имел; очевидно, предполагалось, что число пользователей будет не слишком велико, чтобы сеть оказалась загромождена связями.

Появление и развитие Интернета направило концепцию гипертекста в иное русло. Доступность и демократичность этой сети отвергала какие-либо ограничения по числу или составу пользователей большинства гипертекстов, ее слагающих. Поэтому, сохранив произвольность связей, пришлось пожертвовать их свободой. Гипертекстовая среда Интернета позволяет пользователю установить входящую ссылку на любой из фрагментов, но исходящая ссылка должна размещаться только на таком фрагменте, который он создает сам, или имеет право изменять.

Это, казалось бы, небольшое ограничение на деле привело к существенным отличиям. Пользователь теперь не может соединить два существующих фрагмента гиперссылкой. Единственная возможность связать некоторые два фрагмента – создать свою собственную страницу с исходящими гиперссылками на эти фрагменты. Однако входящие ссылки никак не отражаются в Интернете, так что посетители указанных фрагментов вряд ли что-то узнают об их кем-то декларированной семантической близости. Таким образом, высказывать свое мнение пользователь может лишь «на своей территории», и это является одним из факторов быстрого роста объемов Интернета. Но Интернет растет лишьвширь, развиваясь в одной плоскости, не приобретая качественно новых черт. Экстенсивный путь развития неизбежно исчерпывает себя: расширение без приобретения новых возможностей делает систему плохо контролируемой. Мы видим это уже сейчас, когда огромные объемы информации, содержащейся в глобальной сети, затрудняют ее поиск и оценку.

Экстенсивный путь выразился и в том, что однонаправленные ссылки вызвали мощный рост дублирования информации. Одна из идей Т. Нильсена состояла в том, что цитирование должно исчезнуть – место цитат должны занять ссылки на первоисточники. В условиях однонаправленных ссылок эта идея выглядит неудачной: если пользователь перейдет по перекрестной ссылке на связанный ресурс, то обратный путь возможен лишь через бэктрекинг. В результате разработчики сайтов предпочитают копировать материалы с других ресурсов, а не ссылаться на них, чтобы не упустить тем самым пользователя.

Экстенсивный путь развития глобального гипертекста не является единственным. Существует и иной путь, состоящий в осуществлении предложенной в ранней концепции гипертекста возможности установления связей между двумя любыми фрагментами. В этом случае коллективное осмысление информации и определение связей между ее единицами приводит к усложнению структуры без образования новых фрагментов. Как мы уже сказали, увеличение количества ссылок чревато чрезмерным усложнением структуры гипертекста. Избежать этого можно, структурно отделяя вновь образуемые ссылки от уже существующих. Таким образом не будет нарушено правило «авторских ссылок», на котором основана гипертекстовая среда Интернета. Вновь образуемые ссылки будут расположены не внутри уже существующих фрагментов, но в некоторой надстройке, реализованной в ином структурно-семантическом слое.

Что нужно понимать под этим термином? Информация о вновь образуемых связях может размещаться в любом из узлов сети, подобно основным фрагментам Интернета. Однако для пользователя должно быть понятно, что дополнительные связи представляют собой нечто иное, нежели основные гиперссылки. Поясним это на примере. В гипертекстовых средах часто применяют специальные кнопки «вперед» и «назад», реализующие бэктрекинг – перемещение по так называемому историческому слою, то есть последовательности фрагментов, которые посетил пользователь в течение предыдущей работы. В гипертекстовых документах, имеющих форму большого документа, разбитого на главы или параграфы, также встречаются ссылки «вперед» и «назад». Однако семантика этих ссылок совсем иная. Если в первом случае «назад» означало «к предыдущему просмотренному

фрагменту», то во втором – «к предыдущей главе». Сведущий пользователь, как правило, верно определяет семантику в каждом конкретном случае, поскольку подобные ссылки выделяются структурно, например, на отдельной панели, в особом окне и т.д. Таким образом, можно говорить о структурно-семантических слоях, каждый из которых определяет особую организацию структуры гипертекста, например, последовательную, историческую, иерархическую и т.п.

Подобный слой можно применить и для отражения дополнительных ссылок. Область их применения чрезвычайно широка. Это, например, реализация образовательных трейлов, насущных для нужд обучения, формирование научных или иных обзоров. Большие возможности открывают новые структурно-семантические слои для организации доступа к большим объемам информации для разных категорий пользователей, в зависимости от их потребностей. Очевидно, что число таких слоев может быть значительным, и потому пользователь должен иметь возможность выбора одного из дополнительных слоев, и перехода от одного слоя к другому. Важно отметить, что выбор того или иного дополнительного слоя не сокращает возможностей работы с основным слоем, подобно тому, как реализация бэктрекинга никак не отражается на возможности перемещения по основным гиперссылкам.

Поскольку речь идет о создании новых слоев, сохраняющих свойства гипертекста, и в то же время обособленных от его основного слоя, мы назвали такую модель многослойной моделью гипертекста, а сам гипертекст – многослойным. Термин «многослойный» иногда применяют к гипертексту для указания неоднозначности его семантической организации, но, как можно видеть, в нашем случае он приобретает иное, конкретное значение. Отметим, что многослойная модель является развитием модели, основанной на языке HTML, которую уже вполне можно назвать традиционной.

Что же нужно для реализации модели многослойного гипертекста? Дополнительные слои принципиально не отличаются от основных гипертекстовых слоев, за исключением того, что они должны быть дополнены информацией о соответствии каждого фрагмента одному из фрагментов основного (или другого дополнительного) слоя. Они так же, как и основные фрагменты, могут размещаться в любом из узлов сети.

Но для того, чтобы пользователь имел возможность работать одновременно с двумя или несколькими слоями, необходимо особое программное средство – специализированный обозреватель. Как уже упоминалось, современные обозреватели поддерживают работу с некоторыми специализированными слоями, среди которых исторический слой, реализующий бэктрекинг, близкий ему слой журнализации, слой закладок и другие. Но эти слои являются функционально ограниченными и не гипертекстовыми в полном смысле этого слова. Специализированный обозреватель должен подобным же образом поддерживать работу с дополнительным слоем.

Важнейшая функция специализированного обозревателя при работе с дополнительными слоями – двусторонняя синхронизация слоев. Она означает, что навигация, производимая в дополнительном слое, должна вызывать соответствующий переход по фрагментам основного слоя. Наоборот, навигация в основном слое должна отражаться в смене фрагментов основного слоя.

Поскольку число основных фрагментов Интернета слишком значительно, построение дополнительных слоев возможно лишь для некоторых областей, определяемых прикладными задачами. Это значит, что навигация по основному слою может вывести пользователя за пределы слоя дополнительного. Тем самым вызывается специфическая задача диспетчеризации, решение которой должен взять на себя специализированный обозреватель.

По той же самой причине у пользователя может возникнуть потребность поиска возможных дополнительных слоев к области гипертекста, по которой он совершает навигацию. Следовательно, обозреватель должен поддерживать специфическую функцию поиска дополнительных слоев.

Помимо основных функций добавления комментариев, реализации новых гиперсвязей и выстраивания трейлов, дополнительные слои могут в существенной мере заменить существующие специализированные слои, такие как слой закладок или поисковый слой.

Таким образом, реализация рассмотренной многослойной модели гипертекста открывает новый путь развития Интернета, который при этом фактически приобретает новое измерение. Но, как мы видели, эта идея вырастает из простой попытки приблизить гипертекст к его первоначальной концепции.

Multi-layer Model of Hypertext

I.S. Demin

Finance Academy under Government of RF

Key words and phrases: hypermedia; hypertext; hypertext media.

Abstract: Peculiarities and drawbacks of hypertext model lying in the basis of Internet information organization are considered. It is shown that at present Internet is developing under extensive way, gradually exhausting itself. Basic principles of extra structural semantic layers of hypertext providing absolutely new possibilities for its organization are offered. Such multi-layer model continues the development of hypertext model based on HTML language.

Vielschichtiges Modell des Hypertextes

Zusammenfassung: Es werden die Besonderheiten und die Mängel des der Organisation der Information im Internet-Netz zugrunde gelegten Hypertextmodells betrachtet und beschrieben. Es wird aufgezeigt, daß sich in seinem jetzigen Zustand das Internet nach dem extensiven Weg, der allmählich sich erschöpft, entwickelt. Es sind die Hauptprinzipien der Organisation der zusätzlichen strukturno-semantischen Schichten des Hypertextes, die prinzipiell neue Möglichkeiten geben, angeboten. Solches vielschichtiges Modell ist eine Entwicklung des Modells des Hypertextes, das auf der Sprache HTML gegründet ist.

Modèle de l'hypertexte à plusieurs couches

Résumé: Sont examinées les particularités et les défauts des modèles de l'hypertexte qui sont mis à la base de l'organisation de l'information dans le réseau Internet. On montre que l'Internet se développe aujourd'hui par la voie extensive ce qui le pousse à dégrader. On a proposé les principes essentiels de l'organisation des couches supplémentaires sémantiques et structurelles de l'hypertexte présentant de nouvelles possibilités de son organisation. Un tel modèle est celui du développement du modèle de l'hypertexte fondé sur le langage HTML.