

УДК 004.415.53(075)

<https://doi.org/10.32362/2500-316X-2021-9-2-7-21>

НАУЧНАЯ СТАТЬЯ

Некоторые формализованные подходы к оценке удобства интерфейса пользователя веб-приложений

**Б.М. Басок^{1,®},
С.Л. Френкель²**¹ МИРЭА – Российский технологический университет, Москва, 119454 Россия² Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, 119333 Россия® Автор для переписки, e-mail: basok@mirea.ru

Резюме. Предлагается подход, основанный на применении математических методов моделирования поведения проектируемого веб-приложения и поведения пользователя. Данный подход позволяет оценить удобство применения приложения. Рассматриваются модели интерфейсов пользователей (ИП), применяемые для оценки удобства навигации по сайту. Показывается возможность использования двух классов моделей ИП: структурно-логических (онтологических) и вероятностных. Модели первого класса используются для оценки степени понятности предлагаемой ИП навигации по поисковому веб-ресурсу в принятых терминах. Для этого используются математически-определенные меры сходства между элементами ИП и их связями и элементами гипотетической (но правдоподобной) модели представлений пользователя об области поиска в заданной предметной области. В частности, в работе обосновывается возможность использования в качестве таких мер сходства лексикографических метрик расстояния Жаккара и Левенштейна. Вероятностные модели основаны на цепях Маркова. С помощью данных моделей при наличии надежных статистических данных, собранных при опытной эксплуатации веб-приложения или его прототипа, предлагается оценить среднее число шагов пользователя, необходимое для решения конкретной задачи поиска информации об объекте предметной области. В работе приводятся некоторые рекомендации по использованию мер подобия для улучшения удобства использования ИП веб-приложений. Кроме того, приводятся данные о технике оценки вероятностей переходов цепей Маркова и семантических связей. Для пояснения особенностей предлагаемых в статье моделей анализа ИП приводится сайт Приемной комиссии РТУ МИРЭА (<https://priem.mirea.ru/>) с довольно простой организацией, обеспечивающей поиск информации о поступлении и обучении в Университете. Применение описанного в статье подхода наряду с традиционными методами тестирования удобства использования ИП позволит повысить общий уровень юзабилити приложений и тем самым сократить затраты на выявление и исправление связанных с этим ошибок.

Ключевые слова: веб-приложение, юзабилити-тестирование, онтология, цепи Маркова

• Поступила: 24.08.2020 • Доработана: 09.09.2020 • Принята к опубликованию: 10.10.2020

Для цитирования: Басок Б.М., Френкель С.Л. Некоторые формализованные подходы к оценке удобства интерфейса пользователя веб-приложений. *Российский технологический журнал*. 2021;9(2):7–21. <https://doi.org/10.32362/2500-316X-2021-9-2-7-21>

Прозрачность финансовой деятельности: Никто из авторов не имеет финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

RESEARCH ARTICLE

Formalized approaches to assessing the usability of the user interface of web applications

**Boris M. Basok^{1,®},
Sergey L. Frenkel²**

¹ MIREA – Russian Technological University, Moscow, 119454 Russia

² Federal Research Center Computer Science and Control, Russian Academy of Sciences, Moscow, 119333 Russia

® Corresponding author, e-mail: basok@mirea.ru

Abstract. The paper proposes an approach based on the use of mathematical methods for modeling the behavior of the designed web application and user behavior. This approach allows an evaluation of the usability of this application. The paper discusses the models of user interfaces (UI), which are used to assess the convenience of site navigation. The possibility of using two classes of UI models is shown – structural-logical (ontological) and probabilistic. Models of the first class are used to assess the degree of comprehensibility of the proposed UI navigation on the search web resource in the accepted terms. For this, mathematically defined measures of similarity between the elements of the UI and their relationships and the elements of a hypothetical (but plausible) model of the user's perception of the search area in a given subject area are used. In particular, the paper substantiates the possibility of using Jaccard and Levenshtein lexicographic distance metrics as such measures of similarity. Probabilistic models are based on Markov chains. With the help of these models, in the presence of reliable statistical data collected during the trial operation of a web application or its prototype, it is proposed to estimate the average number of user steps required to solve a specific problem of finding information about a subject area object. The paper provides some recommendations on the use of similarity measures to improve the usability of the UI of web applications. In addition, data on the technique for estimating the probabilities of transitions of Markov chains and semantic connections are presented. To clarify the features of the UI analysis models proposed in the article, the website of the RTU MIREA Admission Committee (<https://priem.mirea.ru>) is provided with a fairly simple organization that provides a search for information about admission and training at the University. The application of the approach described in the paper, along with traditional methods of testing the usability of UI, will increase the overall level of usability of applications, and thereby reduce the cost of identifying and correcting related errors.

Keywords: web application, usability testing, ontology, Markov chains

• Submitted: 24.08.2020 • Revised: 09.09.2020 • Accepted: 10.10.2020

For citation: Basok B.M., Frenkel S.L. Formalized approaches to assessing the usability of the user interface of web applications. *Rossiiskii tekhnologicheskii zhurnal = Russian Technological Journal*. 2021;9(2):7–21 (in Russ.). <https://doi.org/10.32362/2500-316X-2021-9-2-7-21>

Financial Disclosure: No author has a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

ВВЕДЕНИЕ

Важнейшей характеристикой программных систем, в том числе веб-приложений, является удобство их использования (юзабилити). В первую очередь, юзабилити программных систем определяется удобством использования их графического интерфейса пользователя (ИП). ИП считается удобным в том случае, когда для выполнения необходимых действий пользователю требуются достаточно малые

затраты времени для чтения, изучения и анализа контента данного интерфейса. В [1] рассматривалась задача оценки юзабилити ИП веб-приложений, находящихся в опытной или промышленной эксплуатации. Однако полученные экспериментальные данные показали, что, несмотря на тестирование юзабилити ИП на этапах проектирования и разработки, некоторые ошибки проектирования, осложняющие удобство использования веб-приложения, тем не менее, остаются. Цена обнаружения и ликвидации данных

ошибок при эксплуатации указанных веб-приложений при этом существенно возрастает, как минимум на один-два порядка.

Таким образом, для обеспечения высокого уровня юзабилити веб-приложений необходимо существенно повысить качество оценки удобства ИП на ранних этапах жизненного цикла программного обеспечения. В данной работе рассматривается подход, основывающийся на применении математических методов моделирования поведения проектируемой системы и поведения пользователя, позволяющих прогнозировать уровень юзабилити программного обеспечения. Применение данного подхода, наряду с традиционными методами тестирования удобства использования ИП, позволит повысить общий уровень юзабилити приложений, и тем самым сократить затраты на выявление и исправление связанных с этим ошибок.

В статье показывается возможность использования двух классов моделей ИП: структурно-логических (онтологических) и вероятностных. Модели первого класса используются для оценки степени понятности предлагаемой ИП навигации по поисковому веб-ресурсу в принятых терминах. Для этого используются математически-определенные меры сходства между элементами и их связями ИП и элементами гипотетической (но правдоподобной) модели представлений пользователя об области поиска в заданной предметной области. Другой класс рассматриваемых моделей – цепи Маркова, моделирующие переходы пользователя между элементами ИП при осуществлении навигации с помощью данного интерфейса. С их помощью при наличии надежных статистических данных, собранных при опытной эксплуатации веб-приложения, можно оценить среднее число шагов пользователя, необходимых для решения конкретной задачи поиска интересующей информации об объекте предметной области, обслуживаемой данной информационно-поисковой системой. Мы рассматриваем эти аспекты на примере анализа ИП сайта Приемной комиссии РТУ МИРЭА¹ с довольно простой организацией, обеспечивающего поиск информации о поступлении и обучении в Университете. В работе показано, что использование предлагаемого анализа позволяет несколько упростить пользователю поиск необходимой информации на примере этого сайта.

1. АНАЛИЗ ЗАДАЧ ЮЗАБИЛИТИ ИП В СОВРЕМЕННЫХ ИТ-СИСТЕМАХ

Независимо от вида и назначения программного обеспечения (ПО), причины возникновения проблем с удобством использования ИП кроются в разрыве

между местом и доступом к нужной информации и ожиданиям пользователей (в частности, посетителей сайтов). Эти недостатки в дизайне ИП приводят к тому, что пользователи либо не могут найти нужную информацию, либо тратят чрезмерное время на ее нахождение [2]. Поэтому разработчики должны создавать ИП, интуитивно понятные для конечных пользователей, чтобы они могли точно понимать ссылочное значение элементов интерфейса и выполнять намеченные задачи правильно, легко понимая его референтное значение [2], т.е. указание, куда ведет ссылка.

Степень «неудобства» ИП будем характеризовать трудоемкостью поиска необходимой информации с применением соответствующих виртуальных элементов ИП (гиперссылок, кнопок, движков и т.д.) в терминах характеристик тех или иных действий. К указанным характеристикам можно отнести, например, число ссылок, которые должен рассмотреть пользователь до получения результата, среднее число кликов от входа на домашнюю страницу до решения задачи поиска информации или распределение вероятности числа кликов (или определяемого ими времени) до начала решения задачи и т.д.

В нашем определении юзабилити ИП мы не будем фиксировать ту или иную меру трудоемкости, которую можно выбрать, исходя из требований к решаемой задаче (примеры рассматриваются ниже). Ограничимся пока ее неформальным определением как трудоемкости решения задачи поиска информации, выполняемой с помощью данного интерфейса и выражаемой в удобных для пользователя терминах, относящихся к конкретной предметной области.

Коль скоро речь идет о предсказании удобства использования, необходимы те или модели взаимодействия пользователя с ИП. Наиболее изученными и доступными являются марковские модели [2]. Однако их использование зависит от концептуальных моделей ИП, на которые опирается пользователь в своей работе. Это, в свою очередь, зависит от используемых моделей предметной области, для которой разрабатывается данный интерфейс. Это же относится к так называемым «онтологическим» моделям ИП [3].

Поэтому в настоящей статье основное внимание уделяется анализу свойств моделей ИП и предметных областей с точки зрения их приспособленности к оценке и прогнозу влияния качественных свойств данного интерфейса на его юзабилити.

2. ОПИСАНИЕ МОДЕЛЬНОГО ПРИМЕРА

Предлагаемый подход к оценке юзабилити ИП веб-приложения продемонстрируем на примере сайта приемной комиссии РТУ МИРЭА, домашняя страница которого представлена на рис. 1.

¹ URL: <https://priem.mirea.ru/>, дата обращения 20.08.2020.
[URL: <https://priem.mirea.ru/>. Accessed August 20, 2020.]

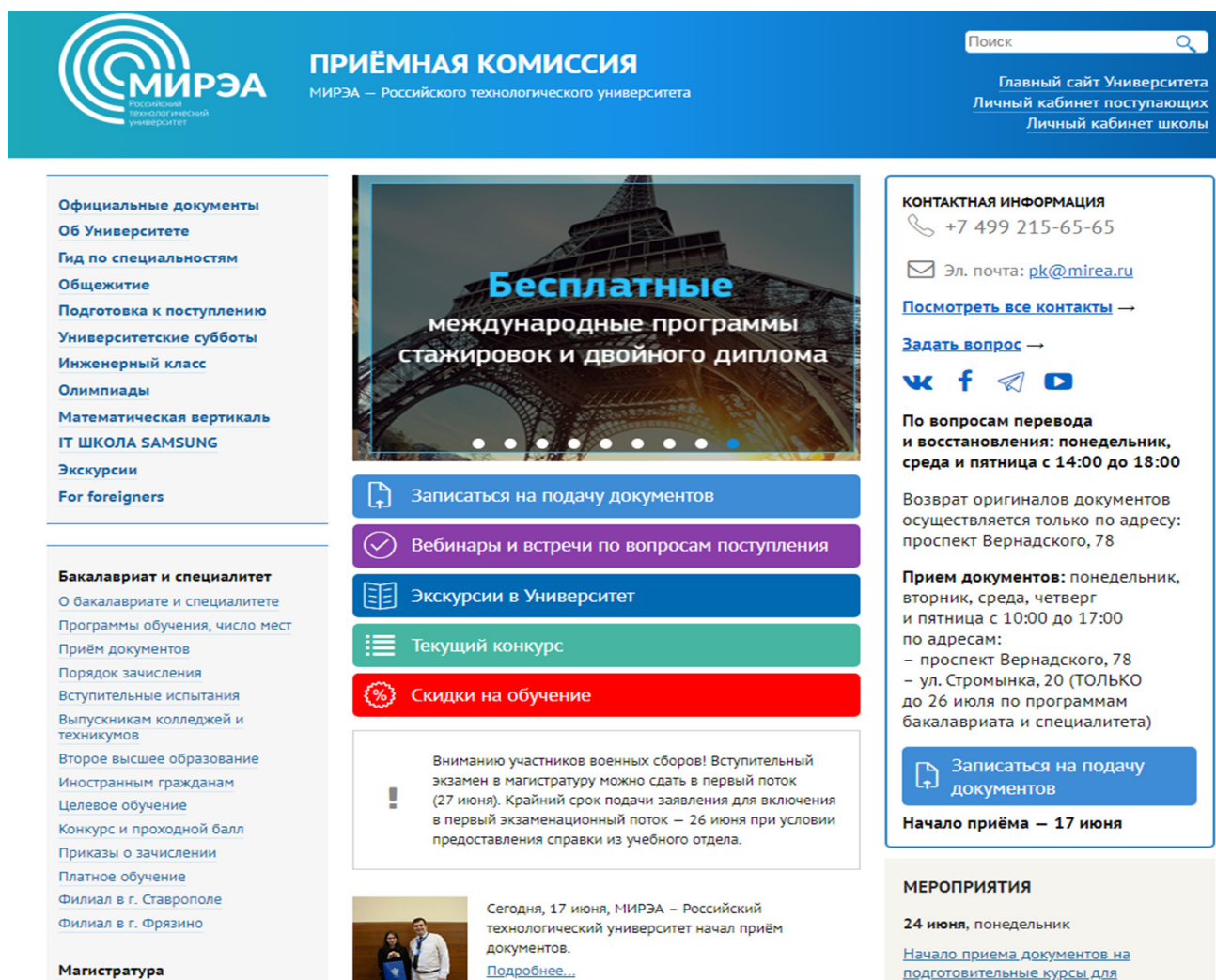


Рис. 1. Домашняя страница сайта «Приемная комиссия МИРЭА – Российского технологического университета»

Навигация по сайту зависит от тех задач, которые ставит перед собой пользователь, и эти задачи понятны из рис. 1, например, поиск условий поступления в Университет в бакалавриат или магистратуру.

Вопрос состоит в том, насколько длителен (трудоемок) путь (например, выраженный в количестве кликов) для того, чтобы добраться до требуемой информации, например, о бакалавриате. С этой точки зрения приемлемый уровень юзабилити ИП может быть получен, если обеспечена последовательная навигация по сайту, т.е. в результате выполнения каждого шага поиска пользователь получает информацию о том, что он может найти, сделав следующие шаги согласно явно представленной навигационной схеме. Соответственно, нам нужна модель поисковой деятельности, связанной с просмотром и пониманием страниц, и модель возможных отклонений из-за недостатков дизайна, определяющего схему навигации по этому сайту.

Модель навигации по сайту предполагает, что посетитель заходит на первую, как правило, домашнюю

страницу или страницу, на которую ссылается внешняя гиперссылка. Далее, пользователь повторяет следующую последовательность действий:

1. Оценивает содержание домашней страницы.
2. Реагирует в соответствии с необходимостью получения дополнительной информации.
3. Ждет загрузки следующей страницы.
4. Читает следующую страницу, пытается найти информацию, которая соответствует цели.

Соответственно, задачу обеспечения юзабилити ИП мы рассматриваем как задачу организации поиска текстовой информации по сайту, посвященному конкретной предметной области.

3. КОНЦЕПТУАЛЬНАЯ МОДЕЛЬ ПОИСКА ПО САЙТУ

В качестве концептуальной модели поиска по сайтам, аналогичным указанному выше примеру, используем модель, идентичную модели представления информации в библиотечных поисковых системах, предполагая, что этот поиск осуществляется

по запросу, имеющему иерархическую структуру с уровнями:

«категория» \rightarrow {«темы»} \rightarrow {«статьи»},

где «категория» наиболее общее понятие рассматриваемой предметной области.

В нашем примере разумно ввести категорию «Уровень образования», к которой относятся темы «Бакалавриат и специалитет», «Магистратура», «Аспирантура», «Среднее профессиональное образование», «Перевод и восстановление». Таким образом, перечисленные заголовки-темы соответствуют всем видам образования, предлагаемым в РТУ МИРЭА.

Можно также выделить категорию «Условия поступления», к которой относятся темы «Наличие мест», «Порядок поступления», «Вступительные испытания», «Целевое обучение», «Иностранцам гражданам», «Статистика прошлых лет», «Приказы о зачислении», «Платное обучение».

Множество {«статьи»} – это гиперссылки, указывающие на часть информации, необходимой пользователю. Например, гиперссылка «Системное программирование и компьютерные технологии» на рассматриваемом сайте указывает на описание направления подготовки бакалавриата 01.03.02 «Прикладная математика и информатика».

В общем случае темы могут представлять собой иерархии с различным числом уровней, и если число уровней больше 1, то вышестоящая тема может рассматриваться для нижестоящих тем как подкатегория рассматриваемой категории. В частности, тема «Бакалавриат и специалитет» может рассматриваться как подкатегория, а связанные с ней темы – как подкатегории нижнего уровня. Например, тема следующего уровня, представленная гиперссылкой «Программы обучения, число мест», и тема «Список направлений обучения (специальностей)», представленная ссылкой <https://priem.mirea.ru/first-degree/list/da>, являются подкатегориями нижнего уровня. Данная иерархия в конечном итоге приводит к статьям, подробно описывающим различные направления подготовки по выбранному профилю. Например, гиперссылка «Математическое моделирование и вычислительная математика» указывает на описание направления подготовки 01.03.02 «Прикладная математика и информатика» по профилю «Математическое моделирование и вычислительная математика».

Стратегия поиска. Поскольку поиск (соответственно, Навигация) по сайту зависит от интересов пользователя, рассматриваем следующую формализованную модель:

Имеется список целей $G = \{g_1, g_2, \dots, g_n\}$, набор целевых категорий $C = \{c_1, c_2, \dots, c_m\}$ и классы

пользователей $\{CU_s\}$, $s = 1, \dots, k$. Среди данных классов можно выделить пользователей, интересующихся средним профессиональным образованием; выпускников средних школ, интересующихся информацией о бакалавриате; студентов, интересующихся информацией о магистратуре и т.д.

Достигнуть цели g_i – значит найти некоторую информацию, называемую «статьей», в которой пользователь может найти сведения, соответствующие этой цели. Отметим, что говоря о цели, мы предполагаем, что посетитель знаком с предметной областью, отражаемой на сайте. Например, будучи выпускником средней школы, он понимает (сам, или с помощью родителей), какую специальность хочет получить, и как с ней связан тот формальный список изучаемых специальностей, который он может найти на сайте.

Если пользователь (абитуриент) строго ориентирован на данный уровень образования, то указанные выше подкатегории, можно рассматривать как категории, например, категорию «Бакалавриат и специалитет», которой соответствуют темы «Специальности», «Проходной балл и конкурс» и т.д.

Поскольку список целей и связанных с ними целевых категорий у каждой группы пользователей разный, разумно рассматривать характеристики юзабилити ИП относительно конкретной цели, которая подразумевает и специфический класс пользователей. Иными словами, будем рассматривать характеристики юзабилити ИП как численные значения $U(g_i, CU_s \in CU)$, g_i – элемент списка целей G , CU – классы пользователей, $s = 1, \dots, |CU|$ (например, «Выпускники средних школ», «Выпускники колледжей», «Бакалавры» и т.д.).

Пусть абитуриенты – выпускники средних школ решают задачу, на какую специальность подать документы, а именно, найти направление бакалавриата с подходящими предметами, проходными баллами и стоимостью обучения. В этом случае возможная навигация по сайту выглядит как граф, представленный на рис. 2. На данном рисунке обрывающиеся стрелки соответствуют просмотру других специальностей, отличных от «Прикладной математики и информатики». Номера в вершинах нужны для экономного обозначения дуг между вершинами E_{ij} , $i, j = 1, 2, \dots, 11$.

Модель процесса достижения пользователем целевой статьи состоит в следующем:

1. Выбор целевой категории: категория, которая соответствует цели, сначала выбирается из категорий, отображаемых на домашней странице.
2. Выбор темы: тема, которая соответствует цели, выбирается из списка тем в выбранной категории. В качестве альтернативы пользователь может вернуться к исходной выбранной категории (подкатегории), если нет очевидного

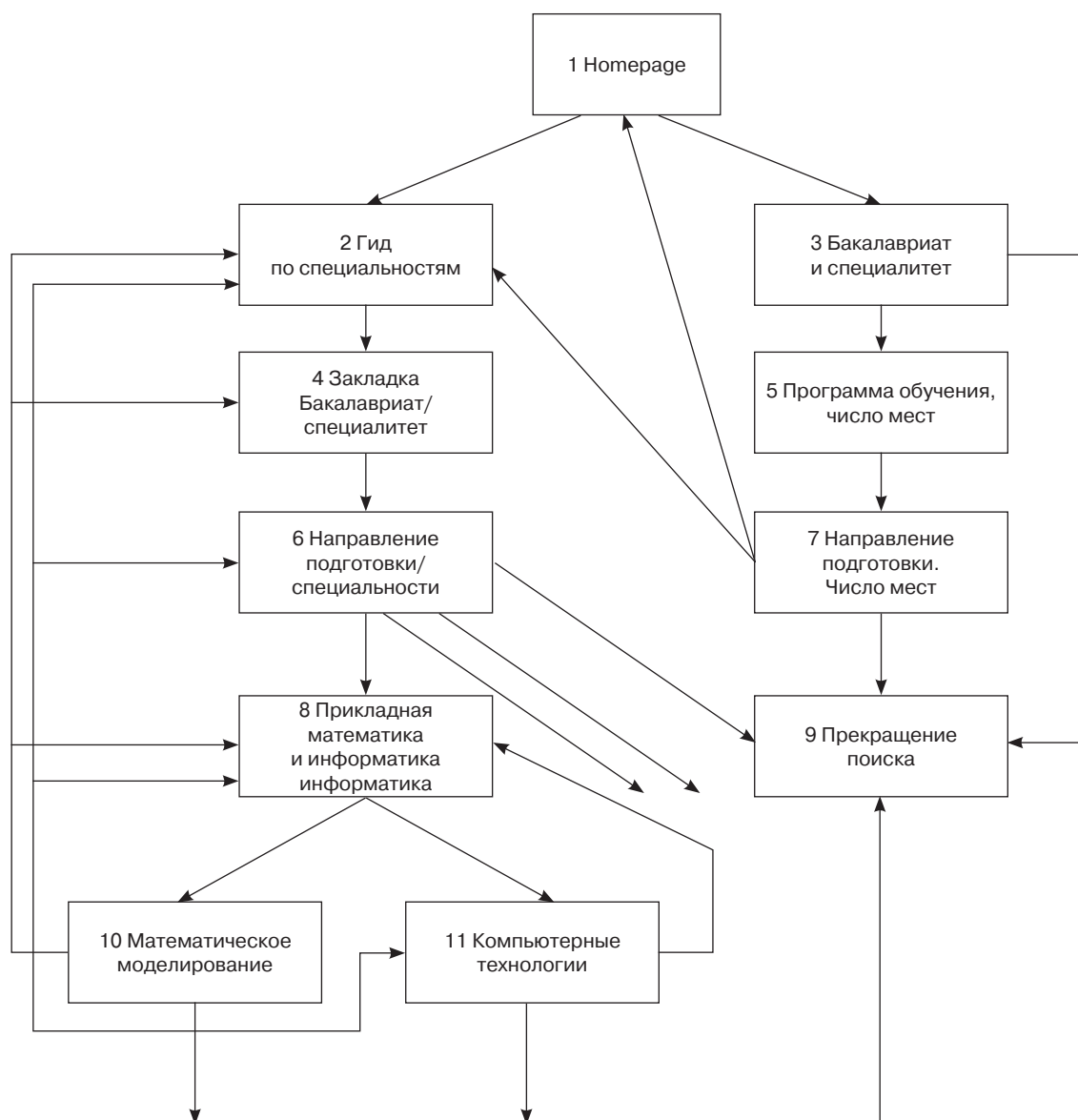


Рис. 2. Навигация по сайту для решения задачи поиска специальностей желающим получить степень бакалавра или специалиста

подходящего маршрута к целевой статье при выборе темы (например, дуги $E_{10,2}$, $E_{11,2}$ на рис. 2).

3. Выбор статьи: выбор темы делает доступным ссылки на статьи, названия которых связаны с этой темой для отображения. Пользователь определяет, есть ли заголовок, который соответствует целевой статье. Пользователь возвращается к выбору темы, если целевой статьи не найдено или он не понял названия статьи.

В указанных терминах проблема обеспечения юзабилити ИП состоит в ответах на вопросы:

- легко ли понять текст описания? (Понятная ли проблематика категорий/тем?);
- насколько сложно выбрать правильную категорию, и будет ли выбрана правильная тема в правильной категории, среди конкурирующих категорий/тем для конкретных целей.

С этой точки зрения основные требования к ИП состоят в следующем: обеспечить семантическую схожесть категорий, тем и статей, так, чтобы у пользователя, знающего предметную область, не возникало трудностей, связанных с несоответствием вышестоящих объектов (например, категорий) и нижестоящих (тем или заголовков статей) в описанной выше иерархической структуре элементов интерфейса. Мерой этого сходства могут быть как известные семантические меры, так и вероятности ошибок, вызванных неверным пониманием пользователем интерфейса, что, в свою очередь, может приводить к необходимости выполнения дополнительных шагов поиска.

Эта задача может быть решена либо структурно-логическими (семантико-онтологическим) методами (SO) [3], либо вероятностными методами с использованием моделей цепей Маркова.

4. МОДЕЛИ АНАЛИЗА СХОДСТВА СИМВОЛЬНЫХ (ТЕКСТОВЫХ) ДАННЫХ В SO

Рассмотрим, как семантическое сходство (similarity, подобие) искомой информации, тем или иным способом понимаемой пользователем, и заголовка гиперссылки, по которой она ищется, сказывается на выборе очередного шага при поиске по сайту сведений в интересующей предметной области. Поскольку этот поиск определяется ИП, то правильное понимание значения элементов данного интерфейса позволяет пользователям переходить непосредственно к интересующему контенту. Например, если на данном шаге поиска в понимании пользователя есть представление о такой сущности, как «Плата за обучение» по данному направлению, то для успешного нахождения этой информации в его сознании должно возникнуть представление о том, что выбранная и нажатая гиперссылка «подобна» данной сущности. В результате нажатия на гиперссылку пользователь получит интересующую его информацию, и переход по этой гиперссылке окажется успешным (в той или иной степени).

Таким образом, существенным аспектом уровня юзабилити ИП является оценка того, насколько гиперссылки на искомую информацию семантически связаны (с точки зрения пользователя) с темами, на которые они указывают.

Для задания мер сходства необходимо иметь модель представления ИП (структуру страниц и ссылок), рассматривая интерфейс как набор некоторых отношений информационных объектов (гиперссылок, заголовков, текстов). Для построения этой модели мы уточним и доопределим схему ИП, поддерживающего навигацию по сайту (рис. 2) как множество терминов, представляющих сущности в интересующей пользователя области.

В нашем рассмотрении область (домен) – это множество сущностей {«Направление подготовки», «Изучаемые предметы», «Профиль специальности», {«Условия поступления», «Условия обучения», «Плата за обучение»}, определяемых набором стандартных, понятных пользователю терминов.

В современной теории и практике информационно-поисковых систем такие модели часто строят с использованием представления знаний *онтологиями* [3]. В современной информатике онтология включает представление, формальное наименование и определение категорий, свойств и отношений между объектами, часто именуемыми *концептами* [3]. Другими словами, онтология является устойчивой языковой конструкцией, дающей традиционное понятное знакомому с предметной областью пользователю выражение соответствующей идеи. Это позволяет

значительно упростить анализ связей между рассматриваемыми сущностями.

Для рассматриваемого сайта это могут быть сущности, представленные такими гиперссылками и текстами, как «Программы обучения, число мест», «Вступительные испытания», которые так или иначе присутствуют в сознании пользователя, пытающегося найти ту или иную информацию по данным вопросам. Иными словами, онтология представляется концептуальной схемой (определенной на множестве концептов), которая состоит из структуры данных, содержащей все объекты данной области интересов пользователя, их связи и правила, принятые в данной области.

Одно из формализованных определений онтологии – это структура $O = \langle X, R, F \rangle$, где X – конечное множество понятий предметной области (на которой определяются категории); R – конечное множество отношений между понятиями; F – конечное множество функций интерпретации источника знаний.

Например, множество понятий – это уровни получаемого образования (бакалавриат, магистратура и т.д.), специальности, направления подготовки (фактически содержательное описание курсов), условия приема (число мест, стоимость обучения).

Примером отношений между понятиями является, например, изучаемый предмет по данной специальности, соответствующий уровню обучения, скажем, базовая дисциплина «Искусственный интеллект» по направлению подготовки «Прикладная математика и информатика» (профиль «Математическое моделирование и вычислительная математика»). И, наконец, интерпретация связана с пониманием пользователем предметной области и то, насколько применяемый интерфейс широко отражает это понимание.

Отметим, что в данной работе мы игнорируем формальный аспект использования категорий [4], поскольку он имеет значение только при разработке автоматических средств построения и анализа онтологий (частично об этом упоминается в подразделе 4.5.).

Таким образом, поскольку, как отмечалось выше, с точки зрения задач пользователя, ИП представляет собой отношение между элементами – терминами, являющимися гиперссылками на тексты, содержащими нужную информацию, то и систему этих отношений и терминов можно рассматривать как онтологию.

Построим, например, онтологию «Уровень образования». Будем рассматривать онтологию как иерархическую структуру, выражаемую направленным ациклическим графом (DAG) (рис. 3), в котором вершины Б/С (Бакалавриат и специалитет),

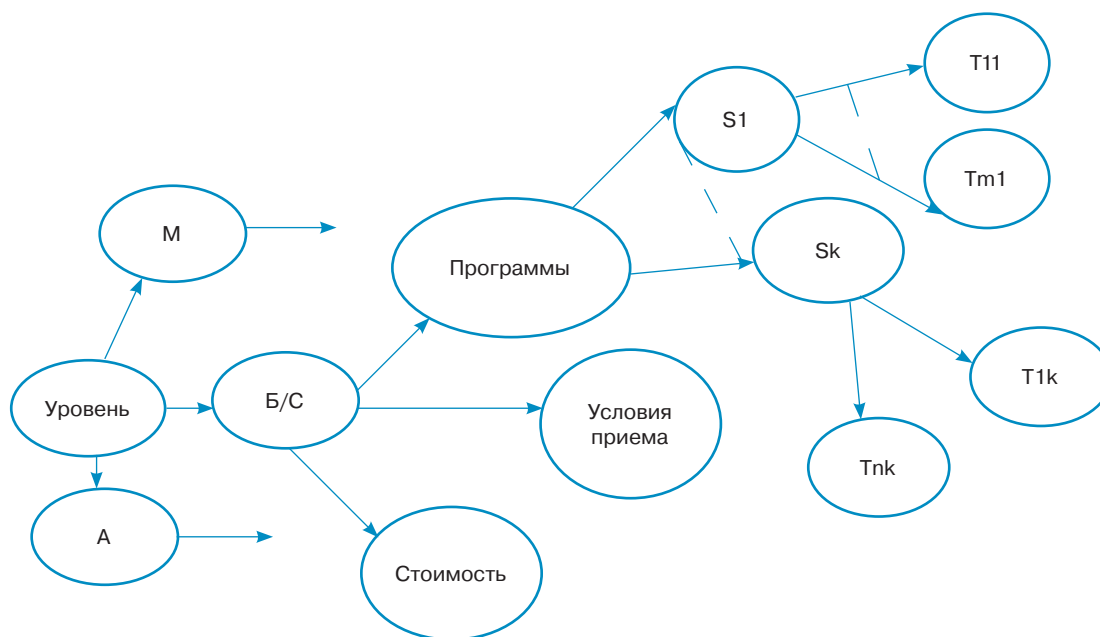


Рис. 3. Онтология области «Специальность» для уровня образования

М и А (соответственно, Магистратура и Аспирантура), вершины S_1, \dots, S_k соответствуют гиперссылкам и заголовкам текста, названным как соответствующие запросы пользователей. Например, S_1 = «Бакалавриат и специалитет», S_2 = «Программа обучения, число мест». Здесь k – число доступных направлений подготовки и специальностей на данном уровне образования, n – число изучаемых дисциплин для данного направления подготовки (специальности), $S = \{S_1, \dots, S_k\}$ – гиперссылки и заголовки текста на имеющиеся направления подготовки (специальности).

Каждая вершина задает темы категории «Уровень образования» вплоть до статей, находящихся по ссылкам T_{ij} , $i = 1, \dots, n$, $j = 1, \dots, k$. Для рассматриваемого примера – это описания программ и дисциплин по профилю специализации данного направления подготовки.

Если считать, что пользователь мыслит в указанных терминах компонентов онтологии, то проблема состоит в том, чтобы наименования элементов ИП (гиперссылок, прежде всего), весьма далекие в смысле семантики получаемого ответа, были бы для пользователя не похожими друг на друга. С другой стороны, необходимо чтобы ссылки, семантически схожие с темой запроса, который пользователь собирается сделать, были бы текстуально схожи с заголовком (названием) гиперссылки (например, «Программа обучения, число мест»).

С точки зрения юзабилити ИП [5] вопрос состоит в близости (в лучшем случае в совпадении) онтологии разработчика ИП и онтологии пользователя (как он представляет себе предметную область, в которой он ищет необходимую ему информацию). При

этом следует подчеркнуть, с точки зрения пользователя, онтология – это совокупность знаний, которыми необходимо овладеть, чтобы понять и правильно интерпретировать элементы ИП и их связи.

Связи элементов ИП во многом предопределяют характер пользовательского диалога, а именно, как пользователь реагирует на действия, которые предлагает ему выполнить данный интерфейс. То есть, необходимо оценить связь онтологий ИП со знаниями и понятиями пользователя о данной предметной области. Если считать, что онтологии как ИП, так и пользовательских представлений состоят из конечного числа концептов, то юзабилити ИП может характеризоваться числом совпадений между онтологиями.

При этом пользователь, чтобы понять назначение элемента ИП, должен не только правильно его интерпретировать и быть знакомым с концепциями, которые присутствуют в тексте (например, быть знакомым с концепцией трехуровневого образования «бакалавр-магистр-аспирант»), но должен также понимать возможности данного элемента при использовании его для навигации по сайту. Например, предполагается, что закладка «←» на веб-странице, соответствующей теме «Гид по специальности», понимается пользователем как указание для возврата к ранее посещенному контенту.

Также следует помнить, что некоторый элемент ИП потенциально может относиться к более чем одной онтологии («конфликт онтологий»), и пользователю неясно, к какой именно он относится. Например, в главном меню ссылка «Программа обучения, число мест» потенциально может относиться как к направлению подготовки/специальности, так

и профильным дисциплинам, о чем пользователь может не знать, и такая двусмысленность может вызвать у него неправильное понимание содержания, которое предлагает ссылка.

4.1. Краткое описание возможных подходов к оценке семантического подобия

Как следует из сказанного выше, для эффективного решения нашей задачи необходима оценка сходства текстов ссылок (мы не рассматриваем их графические представления) с учетом представления пользователя о данной предметной области, то есть семантики рассматриваемых данных.

Вопрос состоит в том, как оценить представление объектов (элементов) интерфейса, обеспечивающего доступ к поиску нужной пользователю информацию и знаний. Для этого, очевидно, необходимо уметь структурировать эти объекты и наделять их семантикой (смыслом в контексте рассматриваемых задач).

В литературе можно выделить следующие пять методов измерения сходства [6]:

- методы, основанные на понятии «расстояния» между объектами, наделенными семантикой;
- методы, основанные на информационном содержимом;
- методы, основанные на свойствах терминов;
- методы, основанные на иерархии онтологий;
- гибридные методы.

Далее рассмотрим лишь самые общие подходы к вычислению сходства, как меры расстояния, основанные на наиболее популярных моделях представления концептуальной информации.

4.2. Расстояние и подобие: формальные аспекты

Формально расстояние («расстояние в пространстве») – это функция D с неотрицательными действительными значениями, определенными на декартовом произведении $X \times X$, такими, что $D: X \times X \in R^+$. Она называется метрикой расстояния на X , если для каждого $x, y, z \in X$:

$$D(x, y) = 0, \text{ если } x = y \text{ (аксиома тождества);}$$

$$D(x, y) + D(y, z) \geq D(x, z) \text{ (неравенство треугольника); (1)}$$

$$D(x, y) = D(y, x) \text{ (аксиома симметрии).}$$

Множество X , снабженное метрикой, называется метрическим пространством. Метрики «подобия» $S(x, y)$ рассматриваются как инверсия к понятию расстояния, которое должно следовать этим правилам, будучи тем больше, чем меньше различия между

объектами x, y . В частности, $S(x, x) > S(x, y)$ при $x \neq y$.

Рассмотрим случай, когда семантика данных предполагает, что данные представляют собой несколько строк упорядоченных символов. С практической точки зрения проблема определения подобия представляет собой комбинацию двух подзадач: а) какой тип сходства наиболее релевантен сравниваемым данным? и б) как задать строку запроса Q для поиска с подходящей сложностью (стоимостью), чтобы найти все строки в наборе данных, чьи расстояния до строки Q не больше, чем заданный порог. Ниже рассмотрим два класса метрик расстояния/подобия – лексикографические и семантические, которые можно использовать в модели для оценки удобства использования ИП.

4.3. Лексикографические метрики

Лексикографическая метрика выражает схожесть (которую нам надо оценить, чтобы понять, насколько ясен для пользователя может быть переход по данной гиперссылке) как некоторую функцию числа или доли похожих слов или лексем. Лексикографические метрики просто сравнивают знаки (элементы текста) и их последовательности в сравниваемых словах. Например, это может понадобиться для оценки подобия между терминами в онтологии, изображенной на рис. 3.

В качестве примеров можно указать метрики Жаккара [7] и Левенштейна [8]. Расстояние Жаккара (Jaccard) $JD(x, y) = \frac{|x \Delta y|}{|x \cup y|}$, где Δ обозначает

симметричную разность между двумя наборами элементов множеств x, y (то есть x и y рассматриваются как неупорядоченный набор символов из данного алфавита с числом элементов $|x|, |y|$). Соответственно определены метрики подобия Жаккара

$$J_S(x, y) = 1 - JD(x, y) = \frac{|x \cap y|}{|x \cup y|}. \text{ При этом элементами}$$

множеств могут быть знаки кода ASCII, слова текста или так называемые n -граммы слов, под которыми понимают последовательность из n знаков [7].

Например, если мы рассматриваем множество как собрание отдельных знаков (букв, пробелов, знаков препинания), то выделяемые из фразы «to_be_or_not_to_be» 2-граммы выглядят, как: to, o, _b, be, e, _o, or, r, _n, no, ot, t, _t, to, o, _b, be, где знаком _ (нижнее подчеркивание) отображается пробел. Если же элемент – слово английского языка, то игнорируя пробелы, получаем следующие 2-граммы, для наглядности заключенные в скобки: (to be) (be or) (or not) (not to) (to be).

Эту метрику можно интерпретировать как вероятность того, что случайное отображение с помощью хэш-функции h_i (разные отображения для разных i) не содержит случайных коллизий, то есть вероятность $Pr(h(x_i) = h(y_i)) = J_S(x_i, y_i) + (1 - J_S(x_i, y_i))/2^k$ – это вероятность того, что случайная перестановка разных подмножеств текста (в частности, подстрок) не отображается в одинаковый индекс таблицы, k – это число битов, отображаемых хэш-функцией h_i [7].

Таким образом, рассматривая гиперссылки x = «Бакалавриат/специалитет» (при входе по гиперссылке «Гид по специальностям» на страницу «Контрольные цифры приема 2020/2021 учебного года») и y = «Бакалавриат и специалитет» (левая навигационная панель веб-структуры курсов) как последовательности букв русского алфавита, находим, что $J_S(x, y) = 0.94$, то есть имеет место исключительно высокое сходство, при существенно отличном содержании результатов перехода по этим ссылкам: «Бакалавриат и специалитет» → «Программы обучения, число мест» → «Контрольные цифры приема 2020/2021 учебного года по программам бакалавриата и программам специалитета очной формы обучения» (содержащие информацию «Специальности, число мест»), и дальше для получения искомой информации в полном объеме надо идти по ссылке «Гид по специальностям», на который можно было бы прийти сразу с домашней страницы.

Если предположить, что величина $J_S(x, y)$ действительно отражает психологическое восприятие пользователем сходства, то этот результат означает крайне неудачный выбор названия гиперссылок, поскольку дает надежду пользователю на семантическую близость информации, получаемой по этим ссылкам. Это, однако, не соответствует действительности, поскольку может приводить к замедлению получения информации пользователем (конфликт содержания переходов по двум достаточно схожим ссылкам).

Наоборот, тексты заголовков гиперссылок «Гид по специальностям» и «Программы обучения, число мест» имеют почти нулевое значение $J_S(x, y)$, но претендуют на то, чтобы предоставить пользователю семантически-схожую информацию. При этом, чтобы найти соответствующую гиперссылку на одной из страниц, надо перейти на другую страницу (дуга E_{72} на рис. 2) и там уже искать релевантную информацию, что говорит о нерациональности такого поиска.

Помимо меры сходства Жаккара для анализа схожести терминов в онтологиях используется Расстояние Левенштейна или Расстояние Редактирования (Edit Distance, ED). Под расстоянием редактирования понимается число символов, которое может потребоваться при преобразовании типа «вставить\удалить\заменить» строки x в y или наоборот.

Например, для строк $str1$ = «INTENTION» и $str2$ = «EXECUTION» нетрудно видеть, что минимальное расстояние редактирования между $str1$ и $str2$ оказывается равным 5. Действительно, выполняя все операции сравнения с $str1$, мы должны сделать по одной операции удаления I, замены N на E, T на X, введения символа C, и замены T на U.

Данная мера отвечает всем условиям (1) и поэтому является метрикой расстояния. Она используется, например, в [9] для вычисления внутреннего сходства между терминологиями (Domain ontology linking), чтобы связать термины, используемые в ссылках, с соответствующими знаниями в некоторой базе знаний или формализованными знаниями пользователя. При этом свойства ED позволяют, например, сравнивать термины с учетом суффиксов и префиксов, отличие в которых может препятствовать связыванию термина и сущности.

Итак, чтобы оценить (формализовать), насколько данный интерфейс позволяет пользователю легко найти нужные ссылки (информацию), мы можем оценивать семантическую схожесть терминов (фраз, представляющих ссылки) внутри онтологий, описывающих соответствующие области запросов (в нашем случае – сведений для поступающих).

4.4. Структурно-семантический подход к оценке сходства текстовых объектов

Структурно-семантический подход к оценке сходства текстовых объектов основан на сравнении расстояний в некоторой метрике семантики объектов рассматриваемых областей. Например, если можно определить некоторую меру сходства (отвечающую аксиомам расстояния (1)) между онтологией «Уровень образования» (рис. 3) и используемых в ней тем «Направление подготовки/специальность» и «Профильные дисциплины» и таких тем, как «Проходной балл в прошлом», «Число мест на данную специальность», которые можно представить, например, в онтологии «Условия поступления», это может помочь абитуриенту определить наиболее подходящую для него специальность, как компромисс между интересом к специальности и желанием стать студентом в этом году.

Конкретизируем меру *расстояния/сходства* для подхода, рассматриваемого в настоящем подразделе. Для любой пары из двух вершин V_i и V_j на рис. 3 (i, j – номера узлов, принимающие значения между 1 и общим числом узлов онтологии) известный метод с интуитивно понятной оценкой их сходства [6, 9] заключается в вычислении расстояния между узлами, равного количеству узлов в иерархии данной онтологии, которые надо пройти («кликнуть») чтобы попасть в узел V_j из V_i . При этом, чем короче расстояние, тем выше сходство. В случае если существует несколько

путей между узлами, может использоваться самое короткое или среднее расстояние из всех путей.

Этот подход обычно называют методом семантического расстояния [6], поскольку он дает меру расстояния между двумя терминами, определяющими семантику в данной онтологии. Расстояние может быть легко преобразовано в меру подобия и нетрудно показать, что эта мера удовлетворяет аксиомам расстояния (1) и поэтому является метрикой.

Можно указать следующие два фактора, определяющих расстояние (хотя есть и целый ряд других [6]):

- 1) плотность в графе онтологий: чем выше плотность, тем ближе расстояния между узлами;
- 2) глубина узлов: чем глубже расположены узлы, тем более очевидна разница между узлами.

Учитывая иерархическую структуру онтологии [3], семантическое сходство между двумя терминами в целом определяется как функция расстояния между вершинами в графе, соответствующего иерархической структуре онтологии (рис. 3).

Например, после выбора концепта «Программа обучения», включающего в себя «детей» (последователей) гиперссылки «Программы обучения, число мест», как новые ссылки, так и статьи с описанием программ обучения, у пользователя отсутствует уверенность в том, ответы на какие вопросы он может получить, перейдя по этой ссылке по концепту «Гид по специальностям».

Мы можем интерпретировать эти затруднения как отсутствие подобия между сформулированной целью пользователя (концептом) «Программа обучения» и ссылками, которые доступны пользователю на данном шаге поиска. Формально это подобие выражается принятой в современных подходах к анализу информационно-поисковых систем мерой «онтологического подобия», выражающей семантическую удаленность (расстояние) одних концептов, представленных вершинами графа на рис. 3 от других, которые рассматриваются как *тематические события* (Thematic event, TE) в данной онтологии. Такое TE могло бы быть просто названием гиперссылки, например, «Гид по специальностям», либо названием гиперссылки, дополненным аннотацией, например, «Специальности и изучаемые профильные дисциплины». Это упростило бы пользователю поиск. Очевидно, что с точки зрения описываемой меры, это означает снижение удаленности данной ссылки от необходимой пользователю информации о профильных дисциплинах. Полезность такой аннотации для нашего конкретного примера очевидна.

При этом следует отметить, что одна из существенных проблем оценки подобия состоит в том, что для тематически родственных слов (фраз) могут быть разные написания. Например, «Студент» и «Профессор» – это родственные термины, которые не похожи друг на друга. Все сходные понятия связаны между собой. Обратное не всегда верно.

4.5. Некоторые рекомендации по использованию мер подобия для улучшения юзабилити ИП

Следует отметить, что в рассматриваемом примере лексикографически схожие ссылки через промежуточные семантически и лексикографически схожие гиперссылки второго уровня (комбинации слов «бакалавриат» и «специалитет») ведут к абсолютно несхожим текстам статей, в которых пользователь будет искать ссылку на нужную информацию. Это говорит о низком уровне юзабилити такого ИП, поскольку пользователь как минимум трижды будет останавливаться из-за необходимости поиска приемлемой навигации.

Иными словами, нарушается следующее важное требование к навигации сайта – текст гиперссылки должен точно отражать содержание адресуемой страницы. Например, переход по гиперссылке «Программы обучения, число мест» (бакалавриат и специалитет) действительно позволяет получить информацию о числе мест. Однако для получения подробной информации о программе обучения придется идти на другую ветку по гиперссылке «Гид по специальностям», по которой пользователь мог бы получить ту же информацию. Поэтому заголовки гиперссылки на статьи должны иметь регулярные синтаксические (и семантические) структуры, которые дают подсказки для переходов.

Для выбора гиперссылки, на которую следует перейти для получения ответа на интересующий вопрос, желательно иметь текстовую подсказку. Например, гиперссылку «Гид по специальностям» следует снабдить аннотацией о списке изучаемых предметов, увеличив тем самым значение рассмотренной метрики онтологического сходства.

Отметим также, что помимо метрики расстояния для оценки юзабилити ИП можно использовать тестовые шкальные оценки знакомства с онтологиями и возможными трудностями восприятия (scores) [9].

Подчеркнем, что мы рассматриваем не задачу автоматического извлечения онтологий из данных [5], а задачу оценки близости используемых элементов (терминов) внутри онтологии или между ранее построенными онтологиями с возможностью их ручной доработки для таких простых примеров, как в данной статье.

4.6. Средства построения онтологий

Для рассматриваемого сайта построение онтологии очевидным образом выполняется вручную. Чтобы использовать онтологический подход в разработке ИП сайтов повышенной сложности, необходимо использовать специальные инструменты, именуемые редакторами онтологий. Среди указанных инструментальных средств можно, в первую очередь, выделить [10]: *Protégé 4.3*, *SWOOP*, *Apollo*,

используемые для проектирования, редактирования и анализа онтологий.

Основная функция любого редактора онтологий состоит в поддержке процесса формализации знаний и представлении онтологии как спецификации (точного и полного описания), и, соответственно, редакторы онтологий предоставляют языковые средства описания онтологии как формальной модели.

Набор возможных действий, поддерживаемый редактором, обычно включает создание, редактирование, удаление понятий, отношений и прочих структурных элементов онтологии, редактирование таксономии. Инструментальные средства нужны для того, чтобы не только вводить и редактировать онтологическую информацию, но и анализировать ее, выполняя типичные операции над онтологиями, например: выравнивание (alignment) онтологий – установка различного вида соответствий между двумя онтологиями для того, чтобы они могли использовать информацию друг друга; отображение (mapping) одной онтологии на другую – нахождение семантических связей между подобными элементами разных онтологий; объединение (merging) онтологий – операция, которая по двум онтологиям генерирует третью, объединяющую информацию из первых двух.

Например, редактор онтологий *Protégé* – это свободно распространяемая Java-программа, предназначенная для построения (создания, редактирования и просмотра) онтологий той или иной прикладной области. Она включает редактор онтологий, позволяющий проектировать онтологии, разворачивая иерархическую структуру абстрактных и конкретных классов и слотов. Данный инструмент поддерживает использование языка OWL и позволяет генерировать HTML-документы, отображающие структуру онтологий.

С точки зрения нашей задачи, главное назначение редактора – создание базы знаний о возможных связях концептов, по которой можно оценивать психологически-обоснованное семантическое расстояние между ссылками.

5. ВЕРОЯТНОСТНАЯ МОДЕЛЬ ЮЗАБИЛИТИ ИП

Для изучения вариантов навигации в соответствии с возможностями ИП применяется модель цепи Маркова первого порядка ко всем путям навигации. Эта модель может представить динамику общения пользователей с сайтом, а именно, вероятности переходов к конкретным ссылкам и прогноз времени перехода к тем или иным страницам или полного выполнения задачи, по сути, интегрируя опыт работы по тестируемой группе, по которой вычисляют матрицу вероятностей переходов цепи.

Процесс навигации по сайту можно рассматривать как конечную цепь Маркова с дискретным пространством состояний и дискретным параметром времени. Для этого будем считать каждую гиперссылку как одно состояние, а один клик – как параметр времени, и учитывать только предыдущую страницу (ссылку), прежде чем перейти на текущую страницу, рассматриваемую, как следующее состояние цепи. Иными словами, мы считаем, что выбор дальнейших действий пользователя определяется только информацией, полученной им в текущий момент.

Следовательно, следующее уравнение для вероятностей переходов между указанными состояниями справедливо, если X_n – это состояние страницы, появляющейся при n -м клике:

$$\begin{aligned} \text{Prob}\{X_{n+1}=j | X_0=i_0, X_1=i_1, \dots, X_n=i_n\} = \\ = \text{Prob}\{X_{n+1}=j | X_n=i_n\}. \end{aligned}$$

Если дугам графа (рис. 2), представляющим навигацию по сайту, поставить в соответствие направления бакалавриата с подходящими предметами, проходными баллами и ценой, и присвоить им вероятности переходов, то этот граф можно рассматривать как граф соответствующей цепи Маркова.

Мы рассмотрим только простой случай однородной цепи Маркова, где вероятности перехода не зависят от n . В этом случае вероятности переходов p_{ij} , можно оценить по данным журнала обращений к сайту при работе некоторой группы тестирующих этого сайта. Иными словами, схема перемещения посетителя на сайте рассматривается как последовательность переходов из одного состояния в другое с добавлением дополнительного состояния, чтобы обозначить случай, когда посетитель прекращает поиск по данному сайту. Будем рассматривать это состояние как поглощающее.

Вероятности перехода цепи Маркова представлены $m + 1$ квадратной матрицей, где m – число состояний (число гиперссылок и текстов, по которым может проходить пользователь при решении своей текущей задачи), элементы которой $p_{ij}(n)$, $i, j = 1, \dots, m + 1$ таковы, что $p_{ij}(n) = P(X_{n+1} = j | X_n = i)$ и

$$\sum_{i,j=1}^m p_{ij} = 1.$$

Вероятности цепи $p_{ij} = n_{ij}/N$, где n_{ij} – число переходов от страницы с индексом i к странице с индексом j (рис. 2) при выполнении тестовых заданий группой тестирующих [1], N – общее число кликов при работе данной группы тестирующих.

Пользователь должен вернуться к выбору темы, если он не может найти целевую статью (страницу,

ссылку) после выбора темы. Пользователь достигает своей цели, не возвращаясь к выбору темы, если выбранная тема является искомым путем к целевой статье, превращающая выбранную тему в поглощающее состояние для цепи Маркова (найдя ответ, пользователь прекращает данный сеанс общения с сайтом).

Поскольку, как отмечалось выше, индекс подобия Жаккара J_S имеет вероятностную интерпретацию, принципиально возможно установить вероятность перехода от страницы к странице на основе полученных значений J_S с некоторым весом каждого перехода, то есть вероятности можно установить априорно, по мерам сходства. Таким образом, может быть получена предварительная оценка юзабилити ИП. Используя эту оценку, можно в ряде случаев сократить время анализа ИП путем исключения заведомо неприемлемых вариантов реализации ИП. Очевидно, что для сложных сайтов желательно использовать те или иные средства автоматизации оценки расстояний (подобия) [11].

Цепь Маркова может быть использована для прогнозирования среднего числа кликов или времени поиска (времени пребывания на сайте для решения рассматриваемой задачи). Обе эти характеристики могут быть важными показателями удобства использования ИП. Если предположить, что с вероятностью, равной 1, пользователь начинает поиск с домашней страницы, количество кликов и соответственно оценка времени пребывания на сайте могут быть вычислены по известному уравнению [12] как среднее число шагов до попадания цепи Маркова в поглощающее состояние.

Это уравнение позволяет прогнозировать среднее число кликов, необходимых пользователю, посещающему сайт, чтобы достигнуть состояния поглощения (обнаружения целевой статьи или ее заголовка), исходя из нахождения на домашней странице. Таким образом, данная модель позволяет дать некоторую вероятностную оценку удобства использования ИП веб-приложений.

6. НЕСКОЛЬКО ЗАМЕЧАНИЙ О ТЕХНИКЕ ОЦЕНКИ ВЕРОЯТНОСТЕЙ ПЕРЕХОДОВ И СЕМАНТИЧЕСКИХ СВЯЗЕЙ

Для оценки вероятностей переходов в цепи Маркова можно использовать инструменты веб-аналитики, например, сервис Яндекс.Метрика^{2,3}, предназначенный для сбора статистики посещений, с включенным в него инструментальным средством

Вебвизор, одна из полезных функций которого – запись всех действий пользователя на сайте⁴. При этом Вебвизор может воспроизвести все действия, совершенные пользователем за конкретный сеанс, из чего можно сделать вывод, какие цели он преследовал на сайте.

Система веб-аналитики фиксирует абсолютно все щелчки левой кнопкой мыши по кнопкам, элементам панелей вертикального и горизонтального меню, ссылкам. После нажатия на кнопку «Карта кликов»⁵ на экране отобразятся данные о кликах, собранные системой сбора статистики. В частности, по данным карты кликов сайта Приемной комиссии МИРЭА – Российского технологического университета⁶ было получено, что только 11% желающих поступить на бакалавриат воспользовались гиперссылкой «Гид по специальностям» с домашней страницы, в то время как подавляющее большинство пользователей к данной гиперссылке обращаются со страницы сайта, открытой по гиперссылке «Программы обучения, число мест». Это на наш взгляд обусловлено тем, что отсутствует явная смысловая связь данной ссылки с информацией о конкретной специальности. Поэтому можно предположить, что пользователям неочевидно назначение гиперссылки «Гид по специальностям», как способа выбора подходящего направления обучения, в то время как большой популярностью пользуется гиперссылка «Программы обучения, число мест».

ЗАКЛЮЧЕНИЕ

Оценка и улучшение юзабилити ИП поисковой системы только на основе тестовой эксплуатации прототипа на этапе проектирования весьма затруднительна, прежде всего, потому, что на этом этапе бывает трудно найти достаточно представительные группы пользователей различных категорий (интересов, одного уровня компьютерной подготовки и т.д.).

В данной статье на примере простого университетского сайта показано, какие математические подходы можно применять для оценки важных характеристик юзабилити ИП при проектировании и разработке сайтов. Этими характеристиками являются меры сходства понятий, отображаемых элементами интерфейса пользователя данной предметной области.

⁴ URL: <https://webliberty.ru/vebvizor-sayt-glazami-polzovatelya/>, дата обращения 20.08.2020. [URL: <https://webliberty.ru/vebvizor-sayt-glazami-polzovatelya/>, Accessed August 20, 2020.]

⁵ URL: <https://akiwa.ru/blog/karta-klikov-v-yandeks-metrike/>, дата обращения 20.08.2020. [URL: <https://akiwa.ru/blog/karta-klikov-v-yandeks-metrike/>, Accessed August 20, 2020.]

⁶ URL: <https://priem.mirea.ru/>, дата обращения 20.08.2020. [URL: <https://priem.mirea.ru/>, Accessed August 20, 2020.]

² URL: <https://blog.sibirix.ru/2016/07/06/ga-vs-yam/>, дата обращения 20.08.2020. [URL: <https://blog.sibirix.ru/2016/07/06/ga-vs-yam/>, Accessed August 20, 2020.]

³ URL: <https://znet.ru/instrumenty/yandeks-metrika/>, дата обращения 20.08.2020. [URL: <https://znet.ru/instrumenty/yandeks-metrika/>, Accessed August 20, 2020.]

Предлагаемый в данной работе подход на наш взгляд является тем более актуальным, поскольку в последнее время возникает все больше задач, связанных с обеспечением юзабилити в многопользовательских и сетевых системах, в программных средствах поддержки социальных сетей. В частности, появились специфические задачи, связанные с безопасностью и секретностью, с администрированием SDN-сетей. Ввиду сложности организации тестирования юзабилити таких систем, актуальной проблемой становится задача их моделирования и предсказания поведения. Это, безусловно, требует разработки математических моделей.

Отметим, что методы, основанные на тестах юзабилити ИП, предоставляют непосредственную информацию о том, как пользователи взаимодействуют

с данным веб-сайтом, и о том, каковы их реальные проблемы с конкретным ИП [13]. При этом юзабилити ИП имеет смысл оценивать на моделях цепей Маркова или онтологий для сайтов, у которых пространство анализа целей (например, сайт, рассматриваемый в этой статье) достаточно велико, и удобство использования сайта практически невозможно оценить, используя индивидуальные пользовательские тесты.

Благодарности

Работа выполнена при частичной финансовой поддержке РФФИ (проекты 18-07-00669 и 18-29-03100).

Вклад авторов. Все авторы в равной степени внесли свой вклад в исследовательскую работу.

Authors' contribution. All authors equally contributed to the research work.

СПИСОК ЛИТЕРАТУРЫ

1. Басок Б.М., Рожанская А.Н., Френкель С.Л. О тестировании удобства использования веб-приложений. *Российский технологический журнал*. 2019;7(6):9–24. <https://doi.org/10.32362/2500-316X-2019-7-6-9-24>
2. Kirajima A., Kariya N., Takagi H. and Zhang Y. Evaluation of Website Usability Using Markov Chains and Latent Semantic Analysis. *IEICE Trans., Com.* 2005;E88–B(4):1467–1475. <https://doi.org/10.1093/ietcom/e88-b.4.1467>
3. Shahzad S., Ontology-based User Interface Development: User Experience Elements Pattern. *Journal of Universal Computer Science*. 2011;17(7):1078–1088. <https://doi.org/10.3217/jucs-017-07-1078>
4. Antoniou G., Franconia E., van Harmelen F. Introduction to semantic web ontology languages. In: *Reasoning Web. Lecture Notes in Computer Science*, Eisinger N., Maluszynski J. (Eds.) Vol. 3564. Berlin: Heidelberg, Springer; 2005, p. 1–21. https://doi.org/10.1007/11526988_1
5. Islam M., Islam A. Ontology mapping and semantics of web interface signs. *Cent. Comput. Inf. Sci.* 2016;6:20. <https://doi.org/10.1186/s13673-016-0077-y>
6. Gan M., Dou X., Jiang R. From ontology to semantic similarity: calculation of ontology-based semantic similarity. *The Scientific World Journal*. 2013;2013: Article ID 793091. <http://dx.doi.org/10.1155/2013/793091>
7. Leskovec J., Rajaraman A., Ullman J.D. *Mining of massive datasets*. California: Cambridge University Press; 2014. ISBN: 978-1-107-07723-2
8. Levenshtein V.I. Binary codes capable of correcting deletions, insertions, and reversals. *Dokl. Akad. Nauk SSSR*. 1965;163(4):845–848.
9. Liu V., Lang B., Zepeng G. Calculating semantic similarity between academic articles using topic event and ontology. URL: <https://arxiv.org/abs/1711.11508>
10. Kapoor B., Sharma S. A comparative study ontology building tools for semantic web applications. *International journal of Web & Semantic Technology*. 2010;1(3). <http://dx.doi.org/10.5121/ijwest.2010.1301>

REFERENCES

1. Basok B.M., Rozhanskaya A.N., Frenkel' S.L. On web-applications usability testing. *Rossiiskii tekhnologicheskii zhurnal = Russian Technological Journal*. 2019;7(6):9–24 (in Russ.). <https://doi.org/10.32362/2500-316X-2019-7-6-9-24>
2. Kirajima A., Kariya N., Takagi H. and Zhang Y. Evaluation of website usability using Markov chains and latent semantic analysis. *IEICE Trans., Com.* 2005;E88–B(4):1467–1475. <https://doi.org/10.1093/ietcom/e88-b.4.1467>
3. Shahzad S. Ontology-based User Interface Development: User Experience Elements Pattern. *Journal of Universal Computer Science*. 2011;17(7):1078–1088. <https://doi.org/10.3217/jucs-017-07-1078>
4. Antoniou G., Franconia E., van Harmelen F. Introduction to semantic web ontology languages. In: *Reasoning Web. Lecture Notes in Computer Science*, Eisinger N., Maluszynski J. (Eds.) Vol. 3564. Berlin: Heidelberg, Springer; 2005, p. 1–21. https://doi.org/10.1007/11526988_1
5. Islam M., Islam A. Ontology mapping and semantics of web interface signs. *Cent. Comput. Inf. Sci.* 2016;6:20. <https://doi.org/10.1186/s13673-016-0077-y>
6. Gan M., Dou X., Jiang R. From ontology to semantic similarity: calculation of ontology-based semantic similarity. *The Scientific World Journal*. 2013;2013: Article ID 793091. <http://dx.doi.org/10.1155/2013/793091>
7. Leskovec J., Rajaraman A., Ullman J.D. *Mining of massive datasets*. California: Cambridge University Press; 2014. ISBN: 978-1-107-07723-2
8. Levenshtein V.I. Binary codes capable of correcting deletions, insertions, and reversals. *Dokl. Akad. Nauk SSSR*. 1965;163(4):845–848.
9. Liu V., Lang B., Zepeng G. Calculating semantic similarity between academic articles using topic event and ontology. URL: <https://arxiv.org/abs/1711.11508>
10. Kapoor B., Sharma S. A comparative study ontology building tools for semantic web applications. *International journal of Web & Semantic Technology*. 2010;1(3). <http://dx.doi.org/10.5121/ijwest.2010.1301>

11. Strsim 0.0.3. A library implementing different string similarity and distance measures. URL: <https://pypi.org/project/strsim/>
12. Mitrofanova A. Absorbing states in Markov chains. Mean time to absorption. Wright-Fisher Model. Moran Model. URL: <https://cs.nyu.edu/mishra/COURSES/09.HPGP/scribe2>
13. Yang Liu, Xiangji Huang, Aijun An. Personalized recommendation with adaptive mixture of Markov models. *J. of The American SOC. for Inf, Sc, and Tech.* 2007;58(12):1851–1870. <http://dx.doi.org/10.1002/asi.20631>
11. Strsim 0.0.3. A library implementing different string similarity and distance measures. URL: <https://pypi.org/project/strsim/>
12. Mitrofanova A. Absorbing states in Markov chains. Mean time to absorption. Wright-Fisher Model. Moran Model. URL: <https://cs.nyu.edu/mishra/COURSES/09.HPGP/scribe2>
13. Yang Liu, Xiangji Huang, Aijun An. Personalized recommendation with adaptive mixture of Markov models. *J. of The American SOC. for Inf, Sc, and Tech.* 2007;58(12):1851–1870. <http://dx.doi.org/10.1002/asi.20631>

Об авторах

Басок Борис Моисеевич, к.т.н., доцент кафедры Математического обеспечения и стандартизации информационных технологий Института информационных технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: basok@mirea.ru. Scopus Author ID: 6602074818, ResearcherID: X-6744-2019, <https://orcid.org/0000-0002-8935-4248>

Френкель Сергей Лазаревич, к.т.н., старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН) (119333, Россия, Москва, ул. Вавилова, д. 44-2). E-mail: fsergei51@gmail.com. Scopus Author ID: 36865996500, <https://orcid.org/0000-0002-2150-7236>

About the authors

Boris M. Basok, Cand. Sci. (Eng.), Associate Professor, Department of Mathematical Software and Standardization of Information Technology, Institute of Information Technology, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: basok@mirea.ru. Scopus Author ID: 6602074818, Researcher ID: X-6744-2019, <https://orcid.org/0000-0002-8935-4248>

Sergey L. Frenkel, Cand. Sci. (Eng.), Senior Researcher, Federal Research Center Computer Science and Control, Russian Academy of Sciences (44-2, Vavilova ul., Moscow, 119333 Russia). E-mail: fsergei51@gmail.com. Scopus Author ID: 36865996500, <https://orcid.org/0000-0002-2150-7236>