

УДК 004

СТРУКТУРИРОВАНИЕ КОНТЕНТА ИНФОРМАЦИОННОГО ПРОСТРАНСТВА НА ОСНОВЕ СИСТЕМНОЙ МОДЕЛИ РЕШАЕМОЙ ЗАДАЧИ С ИСПОЛЬЗОВАНИЕМ АППАРАТА ТЕОРИИ КАТЕГОРИЙ

С. Ф. БАБАК¹, Г. Г. КУЛИКОВ², А. А. БАРМИН³, Г. В. СТАРЦЕВ⁴

²gennadyg_98@yahoo.com, ³abarmin@outlook.com, ⁴startsev@gmail.com

ФГБОУ ВПО «Уфимский государственный авиационный технический университет» (УГАТУ)

Поступила в редакцию 10 апреля 2014 г.

Аннотация. Накопление больших объемов данных требует новых инструментов их обработки. Современные системы поддержки принятия решений позволяют в автоматизированном режиме снабжать лиц, принимающих решения, необходимой информацией. Использование информационно-поисковых систем позволяет идентифицировать и обеспечить прослеживаемость как структурированных, так и неструктурированных информационных ресурсов в рамках информационного пространства организации. Использование аппарата теории категорий для представления контента информационного пространства позволяет оперировать разнородными объектами с использованием общих механизмов и инструментов. Использование данного подхода позволяет выработать общие правила и реализовать их для работы с произвольными типами контента.

Ключевые слова: контент; информационное пространство; теория категорий; идентификация; прослеживаемость.

ВВЕДЕНИЕ

Сегодня мы все являемся свидетелями лавинообразного роста объемов информации как в глобальной сети Интернет, так и в корпоративных интранет-сетях. В информационных хранилищах, распределенных сетях собраны терабайты данных, накопленных в результате работы корпоративных автоматизированных информационных систем. Эти данные можно рассматривать, с одной стороны, как сетевую среду для информационного поиска, с другой стороны, – как источник данных для исследований и поддержки принятия решений.

Использование систем поддержки принятия решений позволяет в автоматизированном режиме снабжать людей, принимающих решения в сложных условиях, необходимой информацией для полного и объективного анализа деятельности.

Существуют различные типы СППР. В зависимости от уровня процессов управленческих решений выделяют следующие типы СППР:

Индивидуальная СППР обслуживает отдельно взятое лицо, принимающее решение, –

руководителя объединения, предприятия, организации.

Групповая СППР ориентирована на обслуживание группы лиц, взаимодействующих между собой при решении какой-либо проблемы. Поддержка процесса выработки групповых решений осуществляется за счет устранения коммуникационных барьеров между членами группы, применения количественных методов анализа решений группой лиц, рациональной организацией самих процедур работы группы.

Организационные и межорганизационные СППР применяются при анализе сложных проблем комплексного, междисциплинарного характера, для решения которых нужны знания и опыт в самых разнообразных областях.

Обобщенная архитектура системы поддержки принятия решений, предложенная Г. М. Маракасом включает следующие элементы [1]:

1. Система управления данными (the data base management system – DBMS).
2. Система управления моделями (the model management system – MMS).
3. Машина вывода (the knowledge engine – KE).

4. Интерфейс пользователя (the user interface – UI).

5. Пользователь (the user).

Для поддержки принятия решений с помощью информационных технологий, включая анализ и выработку альтернатив, в системах поддержки принятия решений используются следующие методы:

1. Информационный поиск;
2. Интеллектуальный анализ данных;
3. Извлечение (поиск) знаний в базах данных;
4. Рассуждение на основе прецедентов;
5. Имитационное моделирование;
6. Генетические алгоритмы;
7. Искусственные нейронные сети;
8. Методы искусственного интеллекта.

Математическая теория категорий изучает отношения между математическими объектами, определенных, в свою очередь, их моделями.

Реальные объекты, отражаемые в форме структурных моделей, с определенной адекватностью и точностью отражают их свойства и отношения. Использование математического аппарата теории категорий позволяет объединять совокупности категорий, порожденных моделями в категории более высокого уровня.

В данной статье рассматривается информационный поиск как метод поддержки принятия решений. Также в статье рассматривается объединение категорий в структурные модели информационных систем с целью уникальной идентификации их элементов и прослеживаемости по связям.

ИСПОЛЬЗОВАНИЕ АППАРАТА ТЕОРИИ КАТЕГОРИЙ ДЛЯ СТРУКТУРИРОВАНИЯ ИНФОРМАЦИОННОГО ПРОСТРАНСТВА

Свойства системных моделей бизнес-процессов, объединенных в категорию

Рассмотрим диаграмму бизнес-процесса в нотации структурного моделирования IDEF0. IDEF0 представляет собой методологию функционального моделирования и графическую нотацию, предназначенную для формализации и описания бизнес-процессов. Таким образом, графическая нотация определяет алфавит, а правила построения модели в данной нотации – грамматику. С точки зрения грамматики Хомского, синтаксис IDEF можно отнести к контекстно-свободным грамматикам. Диаграмма в но-

тации IDEF0 представляет собой две структуры: первая – ориентированный двудольный граф, представляющий формальную структуру модели и глоссарий, определяющий ее семантическое описание.

Исследуем свойство отношения «вход-выход» между элементами диаграммы. Положим, что вершины графа являются элементами категории, а ребра – морфизмами, тогда, если объекты диаграммы декомпозируются, то можем определить следующие свойства диаграммы:

1. Ассоциативное – $Hom_c(A_{11}, A_{12}) \equiv A_1$, то есть возможна декомпозиция и композиция функциональных блоков;

2. Некоммутативное – $Hom_c(A_{12}, A_{11}) \neq A_1$, то есть блоки упорядочены в одном направлении.

Вершина графа модели в нотации IDEF0 имеет уникальный иерархический идентификатор, тогда все элементы этой диаграммы являются идентифицируемыми и прослеживаемы, задача поиска сводится к выбору нужного элемента по известному пути.

Введем в рассмотрение вторую диаграмму в нотации IDEF0, построенную по тем же правилам. Следовательно, исследуя данную диаграмму по свойству «вход-выход» можно также говорить о возможности идентификации и прослеживания всех ее блоков. Таким образом, мы имеем две категории, построенные по сходным правилам и с одинаковыми свойствами. Тогда справедливо утверждение, что две данные диаграммы можно объединить в более общую категорию с такими же свойствами.

Методология структурного анализа и моделирования SADT содержит в своем составе набор нотаций для описания предметной области в различных аспектах: функциональном, информационном, динамическом, семантическом и других. В этом случае указанные свойства категории при их объединении в системную модель также будут сохранены.

Описанные выше положения реализуются с использованием, например, программных средств ERWin Process Modeler (BPWin, начиная с версии 7.3, вошел в состав CA ERWin Modeling Suite). Программное средство позволяет моделировать бизнес-процессы в нотации IDEF0, IDEF1X, IDEF3 и DFD. На рис. 1 приведен демонстрационный пример декомпозиции функционального блока диаграммы в нотации IDEF0 с помощью диаграммы в нотации IDEF3. Также декомпозиция может быть выполнена в отдельном файле диаграммы и подключена с помощью стрелки вызова.

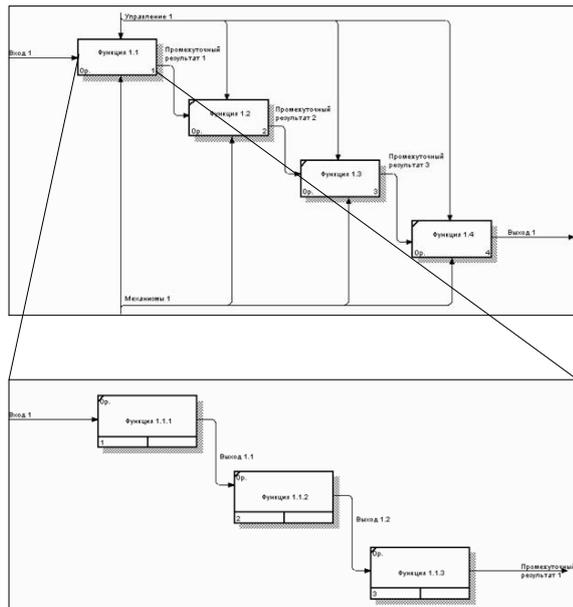


Рис. 1. Декомпозиция функционального блока IDEF0 с помощью диаграммы в нотации IDEF3

Свойства категорий в интернет-проектах для обеспечения идентификации объектов и прослеживаемости их связей

Современные сложные системы также функционируют, сохраняя свойства категории. Рассмотрим в качестве примера свободную общедоступную мультимедийную многоязычную интернет-энциклопедию «Википедия». Википедия функционирует как общедоступный для просмотра и редактирования интернет-ресурс, основной целью которого является сбор и систематизация энциклопедических данных. Данные в энциклопедии представлены в форме статей, навигация между статьями выполняется с использованием встроенной информационно-поисковой системы и гиперссылок, связывающих статьи.

Представим структуру энциклопедии в форме ориентированного графа, где вершинами выступают статьи, а ребрами – гиперссылки между ними. Тем не менее в данном графе, построенном на пространстве интернет-энциклопедии, можно выделить композицию за счет использования категорий. Категория в контексте википедии является одним из ключевых способов организации информации, наряду с пространствами данных и порталами и имеют гомоморфную структуру (см. рис. 2). В связи с этим можем говорить о декомпозиции и композиции категорий. Большинство связей между статьями и категориями однонаправлены, сле-

довательно, большая часть пространства википедии некоммутативна.

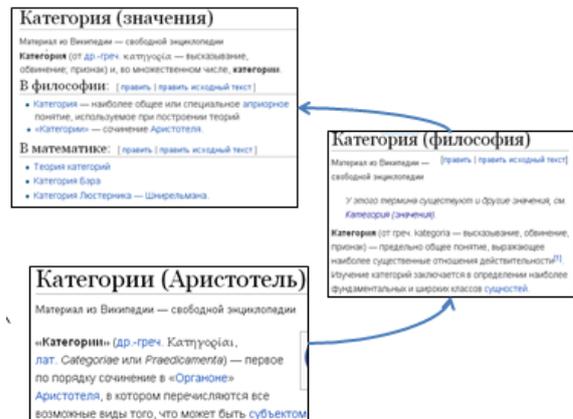


Рис. 2. Свойства категорий в интернет-энциклопедии

Различные пространства данных (множества страниц, начинающиеся с определенной приставки) Википедии имеют сходную структуру, за счет чего объединены в общий проект и существуют согласованно.

Все объекты интернет-энциклопедии различимы и прослеживаемы за счет того, что имеют уникальные имена в контексте соответствующего пространства. Тем не менее, ввиду большой размерности, поиск в такой структуре может осуществляться лишь путем многовариантного поискового запроса.

Использование принципов объектно-ориентированного программирования для формирования категории высшего порядка

Объектно-ориентированное программирование – парадигма программирования, основными концепциями которой является представление предметной области в виде совокупности взаимодействующих объектов и классов. Основными принципами объектно-ориентированного программирования являются:

- абстракция – способ выделения набора значимых характеристик объекта, исключая из рассмотрения незначительные;
- наследование – свойство системы, позволяющее описывать новый класс на базе уже существующего с частично или полностью замененной функциональностью;
- инкапсуляция – свойство системы, позволяющее объединить методы и данные, и скрыть детали их реализации;

- полиморфизм – это свойство системы использовать объекты с одинаковым интерфейсом без информации о типе и внутренней структуре объекта.

В рамках исследования, проводимого на кафедре АСУ УГАТУ, был разработан подход, позволяющий объединить данные из разнородных источников на основе общих правил. Объединение данных заключается в использовании промежуточного слоя, который обеспечивает единообразный интерфейс для доступа к данным, находящимся в источниках с различными способами хранения за счет использования полиморфизма.

Диаграмма классов приведена на рис. 3.

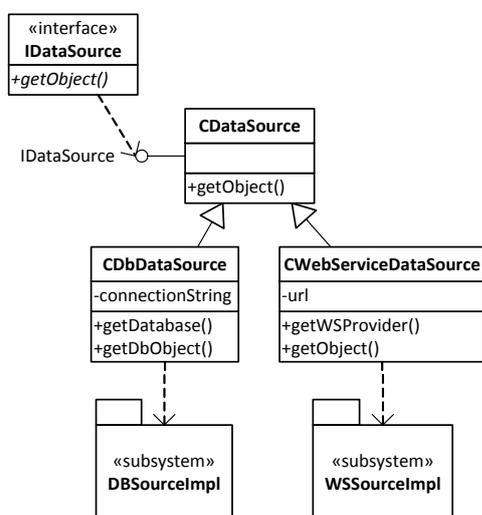


Рис. 3. Диаграмма классов объектов доступа к данным

Использование механизма инкапсуляции, полиморфизма и внедрения зависимостей во время исполнения позволяет обеспечить прозрачный доступ к данным, находящимся в различных источниках. Таким образом, достигается преобразование структуры категории, представленной в конкретном хранилище данных (например, в реляционной базе данных или пространстве локальной википедии) к единообразной для всех источников данных структуре. В результате обеспечивается идентификация и прослеживаемость объектов по общим правилам.

ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИОННО-ПОИСКОВОЙ СИСТЕМЫ ДЛЯ ИНФОРМАЦИОННОЙ ПОДДЕРЖКИ НА ВЕБ-ПОРТАЛЕ ВЫПУСКАЮЩЕЙ КАФЕДРЫ ТЕХНИЧЕСКОГО УНИВЕРСИТЕТА

Существующие решения по созданию единых информационных пространств часто используют информационные технологии типовых корпоративных порталов (например, Microsoft SharePoint, IBM WebSphere Portal и другие) или систем совместной работы (например, Microsoft Exchange, IBM Domino, MegaPlan, Evernote и др.). Очень часто такие решения дополняются информационно-поисковыми системами, к таким системам можно отнести Microsoft FAST Search Server для Microsoft SharePoint, IBM Content Analytics with Enterprise Search для решений IBM, Oracle Secure Enterprise Search для Oracle WebCenter Content. Также на рынке представлены и самостоятельные информационно-поисковые системы общего назначения, которые могут быть интегрированы в информационное пространство организации: Apache Solr, Sphinx, Google Search Appliance, Яндекс.Сервер и другие.

Использование информационно-поисковых систем общего назначения позволяет адаптировать их для использования в различных проблемных областях без необходимости вмешательства во внутренние механизмы функционирования. Однако в данном случае необходима дополнительная подготовка поисковых запросов и снабжение их дополнительными параметрами для структурирования информационного пространства в необходимом аспекте.

Автоматизированный метод применения информационно-поисковых систем в структурированном информационном пространстве

В соответствии с подходом, изложенным выше, при построении информационных моделей различных предметных областей существует следующий порядок идентификации структурированной части модели контента: строится формальная математическая атрибутивная модель; на основании данной модели определяется структура баз данных; на основе ключевых атрибутов производится уникальная идентификация записей. Для семантического обеспечения используется словарь данных или глоссарий информационной модели.

Для решения задач построения интеллектуальных алгоритмов управления возникает необходимость представления описания в киберне-

тическом аспекте, так как данные содержат кибернетические свойства в неявной форме. Для определения структуры системы как объекта управления (проектирования) необходимо произвести кластеризацию атрибутов на четыре класса: входные, выходные, управляющие атрибуты и атрибуты механизма исполнения в соответствии с методологией системного моделирования [2, 3]. Следующей задачей является формирование обучающей статистики, которая в неявной разнородной форме представлена в хранилищах данных и которые необходимы для параметрической идентификации модели. Данные статистики требуется представлять в аналитической форме в виде конечных последовательностей знаков и цифр, взятых из некоторого конечного алфавита, в виде зависимостей, осуществляющих преобразования входных данных V в выходные Y посредством разделения V на части v_1, v_2, \dots, v_N , выполнения над которыми некоторой последовательности операций (функций или процессов) $F = (f_1, f_2, \dots, f_N)$ для получения результатов y_1, y_2, \dots, y_N и формирования на их основе выходных данных. Определение и построение данной подстановки аналитическими средствами встречает сложности: как правило, эти зависимости разнородные, недетерминированные, не являются стационарными, не обладают свойствами эргодичности и т. д. Поэтому прямое применение аналитических и статистических методов идентификации параметров кибернетической модели малоэффективно. Применение универсальных нейросетевых методов в данном случае сопряжено со сложностью организации хранилищ данных, извлечением из них указанных статистик в явной форме. Отсутствует строгая теория по выбору типа и архитектуры нейронной сети, что приводит к необходимости применять алгоритмы самоорганизации, которые также работают достаточно медленно, а в ряде случаев неприемлемо из-за длительности времени обучения [4].

С другой стороны, интеллектуальная поддержка управления различными сложными системами, как правило, основана на опыте применения традиционных средства анализа, оперирующих данными, которые представлены в форме реляционных баз данных. Для идентификации данных в этом случае, как правило, используются ключевые атрибуты модели данных. Их использование обеспечивает однозначную различимость всех записей базы данных.

При таком подходе легко решается задача информационного поиска, так как есть четкое соответствие между таблицами реляционной

базы данных и сущностями предметной области. В современных языках программирования широкое распространение получила технология объектно-реляционного связывания, которая обеспечивает отображение сущностей предметной области в объекты конкретного языка программирования. Использование объектно-ориентированного подхода к разработке программного обеспечения является наиболее предпочтительным, так как классы наиболее полно отображают сущности предметной области. Многие сущности предметной области предназначены для долговременного хранения. Широкое использование реляционных баз данных для хранения объектно-ориентированных данных приводит к семантическому разрыву, который выражен в необходимости постоянного преобразования реляционных данных в объектно-ориентированную форму и обратно. Это не только накладывает ограничения, но и существенно снижает производительность [5].

Задача поиска объекта по идентификатору в реляционной базе данных сводится к формированию структурированного запроса, в котором присутствует совокупность ключевых атрибутов и их значений. Результат выполнения запроса будет содержать требуемые данные.

Использование моделей бизнес-процессов для формирования структурированных запросов к реляционным базам данных позволяет определить необходимые таблицы и ключевые атрибуты для выполнения запроса за счет однозначного соответствия между сущностями информационной модели и таблицами реляционной базы данных.

В настоящее время широкое распространение получили нереляционные базы данных типа «ключ-значение». Такие базы системы управления базами данных позволяют определять и изменять модель данных в процессе исполнения приложения, что обеспечивает быструю адаптацию схемы данных под изменяющийся бизнес-процесс. Тем не менее использование нереляционных баз данных также обеспечивает идентификацию данных за счет наличия глобальных ключевых полей и уникальных в пределах всей базы данных ключей.

Автоматизированный метод применения информационно- поисковых систем в слабоструктурированном информационном пространстве

Значительная доля данных, обрабатываемых в автоматизированных информационных системах, не имеет заранее определенной структуры

либо структура данных меняется быстрее, чем программное обеспечение успевает реагировать на изменения. В этом случае необходим механизм, обеспечивающий возможность интеллектуального структурирования контента информационного пространства на основе данных, находящихся в нем.

Для сохранения целостности свойств информационной системы в различных контекстах, включая семантический, необходимо их постоянное согласование, что обеспечивается определенными соотношениями между формализованной и абстрактной частями информационной системы. Речь идет об описании свойств отношений между математическими объектами, не зависящими от внутренней структуры объектов, т.е. использование теории категорий. В частности, теория категорий неявно предполагает многоуровневое рассмотрение предметной области. Кроме того, использование функций в качестве базового примитива позволяет единообразно трактовать конечные и бесконечные области значений [6, 7].

Предлагается выполнять идентификацию информационных ресурсов на основе системных моделей в слабоструктурированном информационном пространстве по описанному далее алгоритму (см. рис. 4).

Первым шагом алгоритма является определение системы измерений для системной модели. Определение системы измерений для системной модели заключается в формировании многомерного базиса на основе терминов глоссария. Термины глоссария кластеризуются

в пять классов: термины, описывающие входные данные, выходные данные, управляющие воздействия, механизмы и исполнителей, а также термины, описывающие функции. Термины каждого из пяти классов формируют одну шкалу или одно измерение базиса, в результате чего каждая вершина графа или его ребро может быть идентифицировано в этом многомерном базисе. Базис может быть представлен в форме многомерного OLAP-куба.

С точки зрения информационной технологии поиска данных процесс представляет собой индексирование глоссария модели бизнес-процесса.

Следующим шагом является определение границ информационного пространства, включающего прецеденты в скрытой форме, схожих с системной моделью в определенной выше системе измерений. На данном этапе на основе априорных данных о хранилищах, используемых в информационном пространстве, ограничивается область, в которой может находиться контент, необходимый для исполнения бизнес-процесса. Программные и аппаратные подсистемы информационного пространства имеют заранее определенную цель функционирования и предназначены для сбора данных о предметной области в определенном аспекте. В случае необходимости использования данных одной подсистемы в другой выполняется их интеграция путем объединения словарей данных. Таким образом, заранее известен набор функций, выполняемых каждой из подсистем.

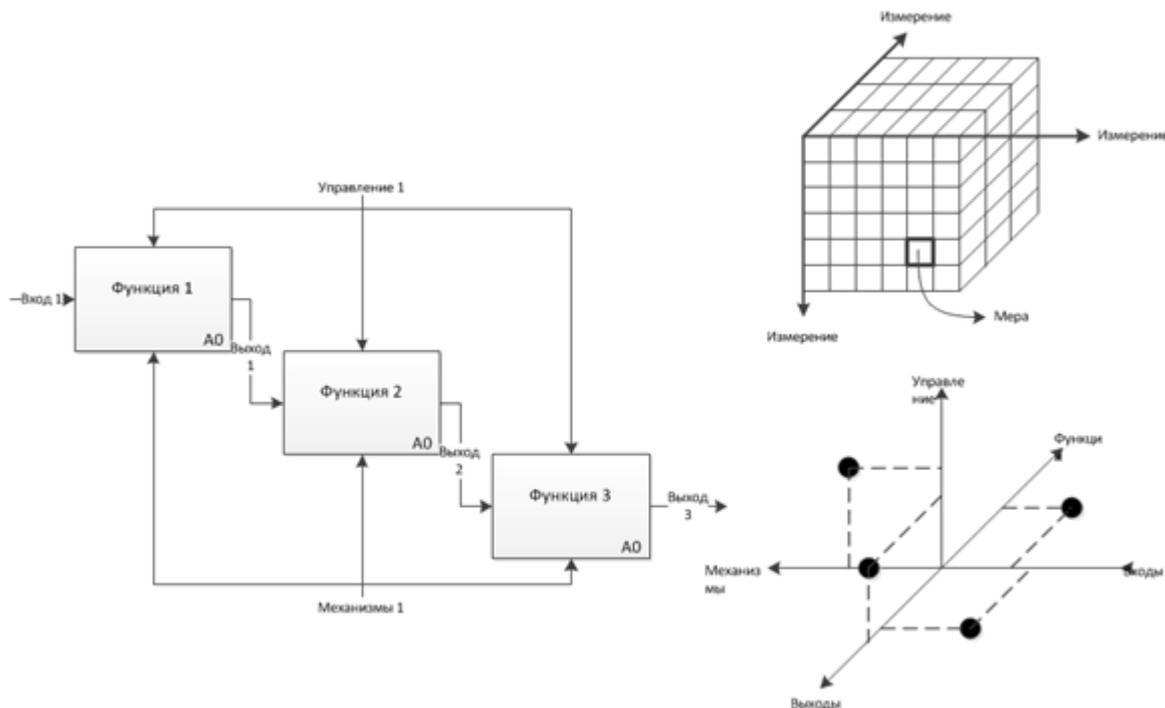


Рис. 4. Формирование многомерного базиса на основе системной модели

На основе этих данных определяются границы информационного пространства – совокупность подсистем, потенциально содержащих данные о бизнес-процессе. Определение подсистем выполняется на основе установления соответствия между их функциями и хранимыми данными и глоссарием модели бизнес-процесса.

Третьим шагом является структурирование области информационного пространства по тождественному базису для системы измерений системной модели решаемой задачи. На этом шаге выполняется структурирование контента области информационного пространства в соответствии с базисом, построенным ранее для системной модели решаемой задачи. В результате такого структурирования будет установлено отношение «многие-ко-многим» между классами терминов глоссария системной модели и объектами области информационного пространства. Стоит отметить, что в область структурирования попадают как относящиеся к решаемой задаче данные, так и не имеющие к ней отношения. Термины, отсутствующие в базисе, но присутствующие в объектах информационного пространства, добавляются для идентификации. Таким образом, формируется расширенный базис на основе контента информационного пространства.

С точки зрения информационной технологии данный процесс представляет собой индексирование контента области информационного пространства. Типовым решением для индексирования контента является формирование инвертированного индекса. Данная процедура выполняется по следующему алгоритму:

1. Выделение текстовой составляющей объектов контента;
2. Разбиение текстов на отдельные термины в соответствии с языком и правилами построения предложений, приведение терминов к канонической форме;
3. Формирование списка словопозиций для каждого термина – указателей на положение термина в объектах контента.

В результате выполнения данного алгоритма будет сформирована разряженная таблица, содержащая в качестве строк объекты контента, а в качестве столбцов термины. На пересечении строк и столбцов находится список словопозиций. Так как таблица является разряженной, предлагается исключить из нее все пустые элементы, в результате будет сформирован инвертированный индекс [8]. Данная структура также может быть представлена в форме многомерного куба, где измерениями являются объекты

контента и термины, находящиеся в них, а мерой выступает список словопозиций.

Следующим шагом является последовательное определение нечетких прецедентов по критерию семантической близости параметрам системной модели. На данном шаге выполняется последовательный просмотр индексированных на предыдущем шаге объектов области информационного пространства и сравнение их с моделью бизнес-процесса. Семантическая близость определяется путем вычисления расстояния Хэмминга. Расстояние Хэмминга показывает число позиций, в которых соответствующие символы двух слов различны. В данном случае сравниваются значения шкал для каждого из объектов и значений этих же шкал для модели бизнес-процесса. Объекты, имеющие большее число совпадений являются более релевантными по отношению к рассматриваемому бизнес-процессу. Полученные в результате сравнения оценки сортируются в порядке убывания семантической близости. В результате выполнения данной операции будут выделены объекты контента области информационного пространства, связанные с выполнением конкретного бизнес-процесса. Определение порогового значения близости выполняется экспертным путем.

Следующий шаг – анализ полученной параметризованной системной модели в расширенном пространстве с тождественной структурой. В результате выполнения предыдущего шага отбираются объекты информационного пространства, семантически близкие к рассматриваемому бизнес-процессу. Теперь эти объекты могут быть использованы в качестве дополнительных данных для исполнения бизнес-процесса, например, в качестве входных или справочных данных. Таким образом, обеспечивается снабжение необходимыми ресурсами для выполнения бизнес-процесса.

В результате снабжения модели бизнес-процесса конкретными ресурсами, необходимыми для его выполнения, возможно автоматизированное формирование параметризованных бизнес-процессов и их последующее выполнение с использованием BPMN-технологий.

Технология использования информационно-поисковых систем совместно с BPMN-системами

Для реализации бизнес-процессов, протекающих в информационном пространстве организации, требуется использование универсальной среды разработки и описания моделей бизнес-процессов, а также сред исполнения описанных и разработанных моделей.

В качестве такого средства выступают BPMN-системы, они позволяют в автоматизированном режиме моделировать и исполнять в распределенной среде бизнес-процессы, представленные в форме модели на одном из языков моделирования.

Многие BPMN-системы содержат встроенные механизмы поиска, но в процессе деятельности их возможностей может оказаться недостаточно ввиду сложности исполняемых бизнес-процессов, поэтому возникает необходимость использования информационно-поисковой системы общего назначения.

Одним из инструментов автоматизации деятельности кафедры является веб-портал. Он является единой точкой доступа ко всем информационным ресурсам кафедры и предоставляет инструменты многоаспектного многомерного поиска в доступных источниках данных. Веб-портал автоматизирует выполнение организационных бизнес-процессов кафедры. Бизнес-процессы представлены в форме задач, модель обрабатываемых данных описывается совокупностью метаданных, информационная модель которых приведена на рис. 5.

Таким образом, поисковый запрос формируется на основе запроса, заданного пользователем, и метаданных, полученных из модели бизнес-процесса.

ОЦЕНКА ЭФФЕКТИВНОСТИ ИСПОЛЬЗОВАНИЯ ИНФОРМАЦИОННО-ПОИСКОВЫХ СИСТЕМ В КАЧЕСТВЕ СРЕДСТВ ИНФОРМАЦИОННОЙ ПОДДЕРЖКИ НА ВЫПУСКАЮЩЕЙ КАФЕДРЕ

Целью автоматизации является повышение эффективности информационной поддержки управления учебным процессом на всех этапах его жизненного цикла.

В качестве одного из ключевых показателей эффективности автоматизированной информа-

ционной поддержки управления учебным процессом можно выделить показатель оперативности (временная эффективность согласно ГОСТ 28195-99).



Рис. 5. Модель метаданных для описания моделей бизнес-процессов

Для примера проведен анализ использования предложенных моделей и метода обработки и анализа. Данный анализ для процесса заполнения карточек документов и поиска данных внутри информационного пространства кафедры показал, что достигается сокращение временных затрат в 2–2,5 раза.

По оценкам экспертов применение предложенных моделей для получения, обработки и анализа данных, необходимых для управления учебным процессом, позволило организовать оперативный поиск и снабжение лиц, принимающих решения, необходимой информацией (процесс поиска вручную последовательно в нескольких системах не мог обеспечить приемлемую оперативность его выполнения).

Кроме того, использование моделей бизнес-процессов в качестве критериев позволяет значительно повысить точность и полноту поиска за счет последовательного сужения области поиска заранее заданными параметрами.

Результаты оценки эффективности приведены в табл. 1.

Таблица 1

Оценка эффективности использования информационно-поисковых систем

Действие	Без использования средств информационного поиска	С использованием средств информационного поиска	Изменение
Время заполнения карточек документов, мин.	4	3	-25%
Поиск в информационном пространстве, мин.	15	2	-86,67%
Среднее число обращений к системе	501352	526061	+4,93%
Среднее число обращений к системе информационного поиска	1079	4172	+286,62%

Таким образом, прямой эффект выражен в сокращении «ручных» операций (снижение времени на поиск, синхронизацию и интеграцию информации, необходимой для управления учебным процессом, ее анализ), что достигается за счет применения методов и средств интеллектуального информационного поиска и формирования единого структурированного информационного пространства.

Также сокращение ручных операций значительно повышает точность и полноту предоставляемых данных, что имеет существенное значение для поддержки принятия управленческих решений.

ЗАКЛЮЧЕНИЕ

Автоматизированная информационная поддержка пользователей заключается в использовании программных средств, которые позволяют в автоматизированном режиме снабжать лиц, принимающих решения, необходимой информацией. Многие автоматизированные информационные системы содержат в своем составе подсистемы информационного поиска. Тем не менее их возможностей может оказаться недостаточно, когда выполнение одного бизнес-процесса требует информации из нескольких систем. В этом случае целесообразно использование информационно-поисковых систем общего назначения. Для повышения эффективности поиска предлагается использовать модели выполняемых бизнес-процессов в качестве дополнительных сужающих критериев поиска, что позволяет значительно повысить его точность и полноту.

Представление контента предметной области в форме категории дает возможность рассматривать и оперировать его объектами независимо от их внутренней структуры. Такое представление позволяет выявлять общие закономерности в поведении разнородных систем.

СПИСОК ЛИТЕРАТУРЫ

1. **Marakas G. M.** Decision support systems in the twenty-first century. Upper Saddle River, N. J.: Prentice Hall, 1999. [G. M. Marakas, *Decision support systems in the twenty-first century*. Upper Saddle River: N. J.: Prentice Hall, 1999.]

2. **Аврамчук Е. Ф.** Технология системного моделирования. М.: Машиностроение, 1988. 521 с. [E. F. Avramchuck, *System modeling technology*, (in Russian). Moscow: Machinostroenie, 1988.]

3. **Буч Г.** Объектно-ориентированный анализ и проектирование с примерами приложений. М.: Вильямс, 2008. 721 с. [G. Buch, *Object-oriented analy-*

sis and modeling with examples, (in Russian). Moscow: Williams, 2008.]

4. **Омату С.** Нейроуправление и его приложения. Кн. 2. М.: ИРПЖР, 2000. 272 с. [S. Omatu, *Supplemental neurocontrol*, Book 2, (in Russian). Moscow: IRPJR, 2000.]

5. **Фаулер М.** Архитектура корпоративных приложений. М.: Вильямс, 2006. 544 с. [M. Fauler, *Enterprise applications design patterns*, (in Russian). Moscow: Williams, 2006.]

6. **Куликов Г. Г., Сулейманова А. М., Шилина М.А., Старцев Г. В.** Системная модель электронного обучения студентов в среде интегрированного информационного пространства вуза // Вестник ВЭГУ. 2011. № 5 (55). С. 23–30. [G. G. Kylikov, A. M. Suleymanova, M. A. Shilina, G. V. Startsev, "E-learning system model in integrated university information space," (in Russian), *Vestnik VEGU*, no. 5 (55), pp. 23-30, 2011.]

7. **Куликов Г. Г., Антонов В. В., Антонов Д. В.** Теоретико-множественная модель информационной системы для многомерного аналитического анализа, отвечающая требованиям хранилищ данных // Вестник УГАТУ. 2012. Т. 16. № 6 (51). С. 189-201. [G. G. Kylikov, V. V. Antonov, D. V. Antonov, "Set-theoretic model of information system for multidimensional analysis based on data warehouses requirements," (in Russian), *Vestnik UGATU*, vol. 16, no. 6 (51), pp. 189-201, 2012.]

8. **Маннинг К., Рагхаван П., Шютце Х.** Введение в информационный поиск. М.: Вильямс, 2011. 528 с. [K. Manning, P. Raghavan, H. Shutce, *Introducing information retrieval*, (in Russian). Moscow: Williams, 2011.]

ОБ АВТОРАХ

БАБАК Сергей Федерович, доц. каф. АСУ. Дипл. инж. (УАИ, 1970). Канд. техн. наук (УАИ, 1978).

КУЛИКОВ Геннадий Григорьевич, проф., зав. АСУ. Дипл. инж. по автоматиз. машиностроения (УАИ, 1971). Д-р техн. наук по системн. анализу, автоматическ. упр. и тепл. двигателям (УАИ, 1989). Иссл. в обл. АСУ и упр. силовыми установками ЛА.

БАРМИН Александр Александрович, асп. каф. АСУ. Дипл. информатик-экономист (УГАТУ, 2011). Готовит дис. в обл. инф.-поисков. систем.

СТАРЦЕВ Геннадий Владимирович, доц. той же каф. М-р техн. и технол. по инф.-упр. системам (УГАТУ, 2003). Канд. техн. наук по АСУ (УГАТУ, 2006). Иссл. в области проектир. инф. систем на основе веб-технологий.

METADATA

Title: Structuring of information space based on models of business-processes and categories theory

Authors: S. F. Babak¹, G. G. Kylikov², A. A. Barmin³, G. V. Startsev⁴

Affiliation:

Ufa State Aviation Technical University (UGATU), Russia.

Email: ²gennadyg_98@yahoo.com, ³abarmin@outlook.com, ⁴startsev@gmail.com.

Language: Russian.

Source: Vestnik UGATU (scientific journal of Ufa State Aviation Technical University), vol. 18, no. 4 (65), pp. 105-114, 2014. ISSN 2225-2789 (Online), ISSN 1992-6502 (Print).

Abstract: Accumulation of large amounts of data requires new processing tools. Modern decision support systems provide decision-makers necessary information in automated way. Using information retrieval systems allows to identify and to ensure the traceability of both structured and unstructured information resources in information space of organization. Use of the theory of categories for presenting content information space allows to operate heterogeneous objects using common tools and instruments. Using this approach allows to develop common rules and implement them to work with arbitrary content types.

Key words: content; information space; category theory; identification; traceability.

About authors:

BABAK, Sergey Fedorovich, Dept. of Automated Systems, Engineer, PhD.

KULIKOV, Gennady Grivorievich, Prof., Dept. of Automated Systems. Dipl. Ingeneer automator (UAI, 1987). Dr. of Tech. Sci. (UGATU, 1989).

BARMIN, Aleksandr Aleksandrovich, Postgrad. (PhD) Student, Dept. of Automated Systems. Informatics and economist(UGATU, 2011).

STARTSEV, Gennady Vladimirovich, Dept. of Automated Systems. Master of technic and technology (USATU, 2003). PhD (USATU, 2006).