

4. Decalogue of good practices in institutional web positioning [Электронный ресурс] Режим доступа: \www/ URL: http://webometrics.info/en/Best_Practices – 18.11.2012 г. – Загл. с экрана.
5. Карпенко, О. М. Интернет-сайты российских вузов: динамика развития по оценке международного рейтинга университетов WEBOMETRICS в 2007-2009 гг. [Текст] /О.М. Карпенко, М.Д. Бершадская, Ю.А. Вознесенская //Инновации в образовании – 2010 – N1 - с. 65-90.
6. Харченко В.П., Иванкевич О.В., Вахнован В.Ю. Аналіз заходів для визначення місця університету в міжнародних рейтингових системах // Вісник НАУ 2011 – №2 http://www.nbuuv.gov.ua/portal/natural/vnau/2011_2/2-1.pdf.
7. Цалко М.П., Жук С.С. Использование международных рейтингов университетов в качестве инструмента маркетинга образовательных и исследовательских услуг //V научн.-практ. конф. студ. “Национальная экономика Республики Беларусь: проблемы и перспективы развития” Минск: БГЭУ, 2012 http://www.ekonomika.by/downloads/Thesis_Tsalko_Guk.doc.
8. Дикке В.С., Якунина Е.Н. Представительский сайт ТГУ и проблема повышения международного рейтинга университета //Труды XVI Всеросс. н. –метод. конф. ”Телематика ‘2009”. Т.1, Серия А. – СПб: Ред.-изд. отдел СПб гос. ун-та инф. технологий, механики и оптики, 2009. – с. 230-231
9. Schonfeld Uri, Shivakumar Narayanan. Sitemaps: Above and Beyond the Crawl of Duty // WWW 2009 Proceedings, Track: XML and Web Data / Session: XML Extraction and Crawling, MADRID, 2009, pp. 991-1000.

Abstract

The article considers the possibilities of improvement of the position taken by the national universities in the international Webometrics rating. The principal aim of the work is to determine the priority actions, which will not require a significant investment and will provide good prospects for long-term growth of performances.

It was shown that the first phase should focus on the increase of the value of presence. A set of practical recommendations was represented to arrange relevant business processes in the document and information management activities of the university. Their implementation and practical application should be supported by technical capabilities of management of multiple sites on sub-domains and by authorization of a high number of users with efficient publication of their documents.

The results of the work can be applied while forming the policy of development of websites of national universities for foreign students with the aim to attract more students

Keywords: rating, Webometrics, internet, site, indexing

В даній роботі проведено дослідження проблеми, що склалася під час проведення децентралізованої автоматизації організації, було розглянуто існуючі методи рішення даних проблем. Запропонована інформаційна технологія інтеграції БД

Ключові слова: інформаційна технологія, база даних, інтеграція, інформаційна система, система управління базами даних

В данной работе исследована проблема, сложившаяся в ходе децентрализованной автоматизации организации, рассмотрены возможные пути её решения. Предложена информационная технология интеграции БД

Ключевые слова: информационная технология, база данных, интеграция, информационная система, система управления базами данных

УДК 004.652.8

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ИНТЕГРАЦИИ БАЗ ДАНЫХ

Н. В. Черненко

Аспирант

Кафедра информационных управляющих систем
Харьковский национальный университет

радиоэлектроники

пр. Ленина, 14, г. Харьков, Украина, 61166

Контактный тел.: (057)702-17-36,093-795-31-65

E-mail: iyc@kture.kharkov.ua

1. Введение

В результате активного развития информационных технологий за последние 20 лет было предложено множество подходов, методологий, архитектур и инструментальных средств реализации проектов инфор-

мационных систем. Одни устоявшиеся «канонические» и современные архитектуры сменялись другими. С таким активным, часто неуправляемым процессом развития информационных технологий формировались и развивались множество организаций и предприятий, руководство которых понимало, что необходимо

внедрять информационные технологии для оптимизации основных бизнес-процессов, но мало компаний могло привлечь услуги качественного IT-консалтинга. Становление стеков технологий, архитектур, практик применения информационных технологий (ИТ) и, одновременно с этим, отсутствие квалифицированного подхода к автоматизации деятельности объекта управления привело к тому, что сегодня в большинстве организаций функционирует спектр информационных систем (ИС), реализованных с применением множества разнообразных, а часто разнородных информационных технологий [1].

По мере функционирования на одном объекте управления нескольких информационных систем происходит накопление информации в их базах данных, что приводит к снижению коэффициента полезного действия от автоматизации. Это может сказаться на формировании корректных управляющих воздействий на основании неактуальной, а зачастую, и противоречивой информации, которая выявляются на стыке предметных областей унаследованных информационных систем. Данный факт вызван тем, что механизм обеспечения целостности данных реализуется на уровне системы управления базами данных. Сложной задачей является поддержание целостности данных на «стыке» систем, особенно по мере роста объемов хранимых данных.

Как показывает исследование данной проблемы, решением является интеграция баз данных информационных систем [2,3]. Особое место отводится базам данных в виду того, что информационные системы направлены на удовлетворение потребностей пользователей в информации, а непосредственным инструментом хранения и работы с данными являются базы данных.

Задачам интеграции уделено значительное внимание в научных трудах [2-4]. Начиная с 70-х годов, велись работы по межмодельной структурной интеграции. Часть трудов посвящены вопросам интеграции данных [3,4], некоторые – вопросам интеграции ИС [2]. Существуют также работы, посвященные интеграции баз данных, в которых рассматривается как интеграция структурной составляющей, так и интеграция данных [5]. Существуют также работы, позволяющие вовлечь в процесс интеграции не только реляционные, но и плоские БД [6], а также дополнительные критерии для оценки возможности интеграции БД на основании функциональной обеспеченности СУБД [7].

Активное развитие ИТ, практик проектирования ИС и программирования, выполнения резервного копирования ориентированы на создание ИС с нуля, однако они не представляют инструментария и методологии для исчерпывающего представления как процесса реинжиниринга унаследованных ИС в целом, так и для отдельных комплексов, образующих ИС.

Существует множество разнообразных методов, практик и т. д. по интеграции БД, описывающих данный процесс на различных уровнях абстракции, часто не связанных между собой.

2. Постановка задачи исследования

Целью проводимых исследований является разработка информационной технологии интеграции БД,

которая позволила бы описать процесс интеграции, начиная от анализа исходных БД и, заканчивая, формированием результата.

Как отмечалось выше, существуют методы интеграции реляционных БД [5], плоских и реляционных БД [6], критерии оценки целесообразности интеграции БД [7], позволяющие минимизировать риски (в первую очередь, финансовые) получения некорректного результата интеграции БД унаследованных ИС. Однако не раскрыта очередность применения тех или иных методов в контексте единого процесса интеграции БД.

Для формального описания процесса интеграции баз данных предлагается информационная технология интеграции баз данных, которая основывается на применении представленных ранее методах [5-6] и критерия оценки целесообразности интеграции БД [7]. Данная технология позволяет осуществлять интеграцию реляционных, объектно-реляционных БД и БД, хранящихся в отдельном файле.

3. Разработка информационной технологии интеграции БД

Для описания прикладной технологии предлагается выделить следующие её функциональные блоки:

- 1) блок 1 «Анализ начальных условий и оценки целесообразности интеграции БД»;
- 2) блок 2 «Проверка условий, определяющих состав дальнейших работ»;
- 3) блок 3 «Первичное преобразование данных» (в случае необходимости интеграции плоских БД, хранящихся в отдельных файлах);
- 4) блок 4 «Интеграция БД».

В блоке 1 на начальных этапах реализации информационной технологии интеграции БД в осуществляет блок анализа начальных условий и оценки целесообразности интеграции БД. Как показано на схеме ниже (рис. 1), реализация данного блока состоит из двух этапов:

- 1) анализ начальных условий;
- 2) оценка целесообразности интеграции БД на основе функциональной обеспеченности СУБД.

Этап 1 – «Анализ начальных условий».

В качестве начальных условий необходимо рассмотреть следующие аспекты БД унаследованных ИС: тип модели данных, используемая БД, степень согласованности БД в целом, степень нормализации БД, степень пересечения предметных областей унаследованных ИС.

Эти начальные условия являются определяющими при выборе конкретной модификации разработанного метода синтеза интегрированной реляционной БД [5].

Этап 2 – «Оценка целесообразности интеграции БД на основании функциональной обеспеченности СУБД».

Данный этап необходим для минимизации рисков, связанных с интеграцией БД. Как было отмечено выше, реинжиниринг ИС в целом, и внутримашинного информационного обеспечения в частности, является дорогостоящим мероприятием. Для этого необходимо привлечение высококвалифицированных специалистов в данной области. Поэтому для минимизации

временных, трудовых затрат и рисков, связанных с процессом интеграции, предлагается на ранних этапах данного процесса определить вероятность получения требуемого результата, основываясь на тех функциях систем управления базами данных, которые в определенной мере использованы при реализации функциональности ИС. В работе [7] был предложен интегральный критерий оценки целесообразности интеграции нескольких баз данных на основе функциональной обеспеченности СУБД. В качестве результата выполненных работ данного этапа может выступать сообщение о нецелесообразности интеграции БД с указанием тех показателей, которые наиболее критично влияют на получение успешного результата либо сообщение об ориентировочных затратах на процесс интеграции. Более подробно процесс оценки целесообразности будет раскрыт далее при разработке псевдокода, который позволит описать данную информационную технологию интеграции БД более детально.

В блоке 2 определяется состав дальнейших работ по интеграции БД. Если в ходе проверки условия соответствия модели данных выясняется, что модель данных не подходит для интеграции, то формируется сообщение о невозможности интеграции из-за несоответствия типа модели данных.

В блоке 3 реализуются функции двух этапов, которые позволяют преобразовать исходные БД к виду, которому должны соответствовать исходные БД, чтобы можно было бы произвести их интеграцию.

Первый этап – «Преобразование к виду текстового файла».

Следует отметить, что ключевыми условиями, которые определяют состав работ, являются вид модели данных БД, а также необходимость интеграции реляционных БД и плоских БД, хранящихся в отдельном файле. В случае необходимости такой интеграции необходимо проанализировать формат представления данных в плоских БД и при необходимости осуществить преобразование их к виду текстового файла, используя разработанный в работе [6] алгоритм перехода между представлениями данных в плоских БД.

Второй этап – «Преобразование плоской БД к виду универсального отношения реляционной БД».

Он предусматривает приведение данных из плоской БД к виду универсального отношения реляционной модели данных. Для этого выполняются работы, регламентируемые разработанным в работе [6] методом приведения плоских БД, хранящихся в отдельных файлах к виду реляционного отношения реляционной модели данных. В ходе выполнения работ согласно

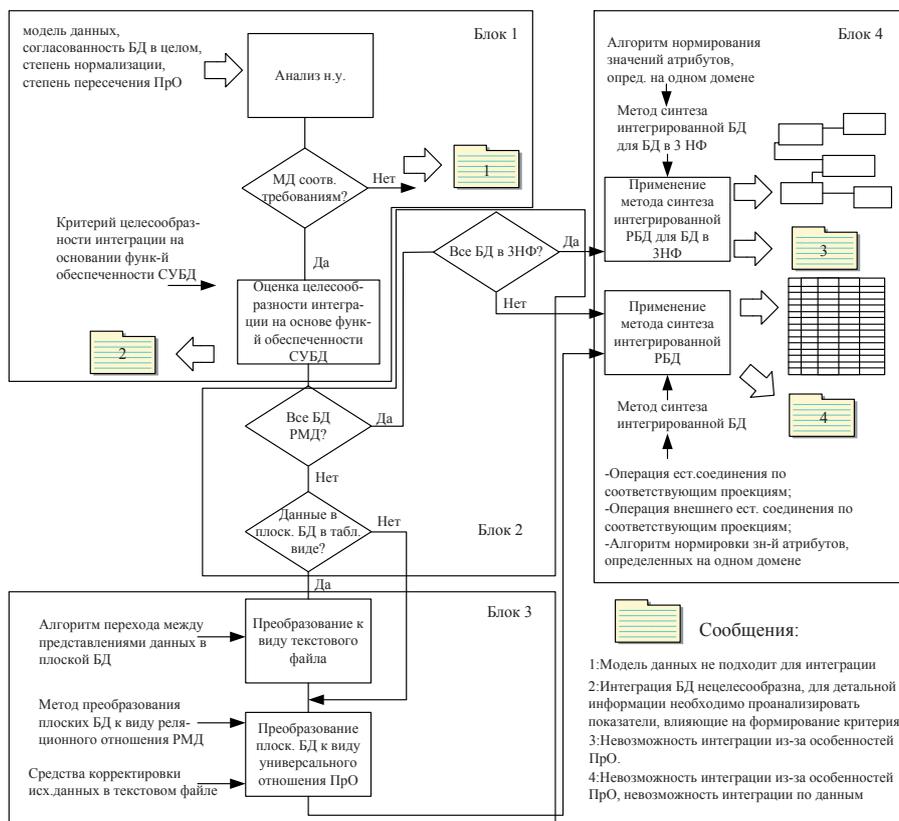


Рис. 1. Схема прикладной информационной технологии интеграции БД

данному методу исполнителю необходимо применить средства корректировки исходных данных в текстовом файле. Более подробно данные средства будут описаны при практической реализации предлагаемой информационной технологии.

В блоке 4 осуществляется применение метода синтеза интегрированной реляционной БД или его усовершенствований [5] в зависимости от начальных условий.

Метод синтеза интегрированной реляционной БД (РБД) предполагает возможность интеграции реляционных БД, независимо от того, в какой степени нормализации они находятся. В ходе проведенных исследований было определено, что существуют случаи, когда на объекте автоматизации функционируют не только реляционные БД. Для их интеграции между собой и с реляционными БД в работе [6] предложен усовершенствованный метод за счет введения преобразования исходных данных плоских БД к виду универсального отношения реляционной модели данных (РМД) на его ранних этапах. Также было произведено усовершенствование базовой модификации метода синтеза интегрированной РБД для случаев, когда на ОУ функционируют ИС, базы данных которых нормализованы до 3 нормальной формы. Данная модификация позволяет произвести интеграцию БД без приведения их к виду универсального отношения, что минимизирует стоимость и сроки данного процесса.

В зависимости от полученных результатов в ходе выполнения работ в данном блоке необходимо проводить интеграцию, применяя одну из разработанных модификаций метода синтеза интегрированной реляционной:

– метод синтеза интегрированной реляционной БД;

– метод синтеза интегрированной реляционной БД для БД в 3 НФ.

Для этого введем следующие равнозначные взаимисключающие этапы:

Этап 1 – применение метода синтеза интегрированной реляционной БД;

Этап 2 – применение метода синтеза интегрированной реляционной БД для БД в 3НФ.

На этапе 1 для случаев, когда производится интеграция РБД, имеющих разную степень нормализации, плоских БД, хранящихся в отдельных файлах, и объектно-реляционных БД следует применять базовую модификацию метода синтеза интегрированной реляционной БД.

В результате проведения работ на данном этапе получим следующие результаты: формирование сообщений о невозможности интеграции БД из-за коренного различия предметных областей или из-за отсутствия необходимого набора общих данных; формирование результирующего общего универсального отношения, которое соответствует глобальной предметной области, покрывающей все локальные предметные области унаследованных информационных систем.

Этап 2 – «Применение метода синтеза интегрированной реляционной БД для БД в 3НФ».

Работы в рамках данного этапа выполняются, если в ходе этапа «Анализ начальных условий» было выявлено, что БД всех унаследованных ИС находятся в 3 НФ. Данный этап предполагает применение усо-

вершенствованного метода синтеза интегрированной реляционной БД из множества отдельных реляционных баз данных (3 НФ) без приведения их к виду универсального отношения. Данное усовершенствование предполагает использование алгоритма нормирования значений атрибутов, определенных на одном домене. В результате выполнения данного этапа могут быть получены следующие два результата: формирование сообщения о невозможности интеграции из-за коренного различия предметных областей; сформированная схема интегрированной реляционной БД, нормализованной до третьей нормальной формы (наполнена непротиворечивыми данными, что позволяет ввести ее в эксплуатацию без дополнительных преобразований).

4. Выводы

В статье предложена информационная технология интеграции БД, позволяющая: установить логические взаимные связи между уже разработанными методами интеграции БД; сформировать наглядное представление их применения, а за счет использования критерия оценки целесообразности интеграции БД на основе функциональной обеспеченности СУБД снизить вероятность формирования неудовлетворительного результата после окончания проведения затратных работ. Данная информационная технология в дальнейшем может быть реализована в качестве инструментального средства интеграции БД и использоваться проектными организациями в своей деятельности.

Литература

1. Лоскутная автоматизация, или как управлять «зоопарком» программ – [Электронный ресурс] – <http://www.bytemag.ru/articles/detail.php?ID=14862> – 20.05.2012 – Загл. с экрана.
2. P. Ziegler Three decades of data integration – all problems solved? [Электронный ресурс] - Режим доступа: www.researchportal.ch/unizh/p6033.html - 18.10.2011- Загл. с экрана.
3. Christoph Koch Dissertation Data integration against. Multiple Evolving Autonomous Schemata. [Электронный ресурс] - Режим доступа: www.csd.uoc.gr/~hy562/Papers/thesis_final.pdf - 18.10. 2011 – Загл. с экрана.
4. Кулик, Ю.А. Методика создания контекстно-независимой системы интеграции данных [Текст] / Ю.А.Кулик, А.С.Ковалев // Радио-электронные и компьютерные системы, - 2011. - №1, - С.46-51.
5. Левыкин, В.М. Метод синтеза единой структурной составляющей реляционной модели данных [Текст] / В.М. Левыкин, В.А. Филатов, Н.В. Черненко // Управляющие системы и машины, - 2011. - №4, - С.3-8.
6. Левыкин, В.М. Метод интеграции реляционных и плоских БД [Текст] / В.М.Левыкин, Н.В. Черненко // Київ: Збірник наукових праць: «Системи управління, навігації та зв'язку», - № 1 (21) - 2012. - С.122-126.
7. Левыкин, В.М., Интегральный критерий оценки целесообразности интеграции нескольких баз данных на основе функциональной обеспеченности СУБД [Текст] / В.М.Левыкин, Н.В. Черненко, А.В. Петриченко // Харків: Науково-технічний журнал: «Радіоелектронні і комп'ютерні системи» № 4 (52) - 2011. - С.99-103.

Abstract

The article considers the problem that has developed in the course of decentralized organization automation, and its possible solutions. To eliminate the defects it was suggested to integrate the databases of heterogeneous information systems into a single database of organization. A study of existing methodologies, approaches and practices of the integration of databases took place. The article suggests the formal description of the information technology of database integration, which allows you to select some solutions from a variety of approaches to database integration, to determine their interaction and their application to solve the problems of integration of databases of information systems. This information technology can be implemented as a tool, with its further application in the operation of the design organizations for the re-engineering of information systems

Keywords: *information technology, database, integration, information system, database management system*