

УДК 004.434+004.91

**М.А. Жигалова, А.О. Сухов****ПРЕДМЕТНО-ОРИЕНТИРОВАННЫЙ ЯЗЫК ОПИСАНИЯ ДОКУМЕНТОВ,  
ИСПОЛЪЗУЕМЫХ ПРИ ПРОЕКТИРОВАНИИ ИНФОРМАЦИОННЫХ  
СИСТЕМ\***

*В процессе разработки и сопровождения программных продуктов создается и используется большое число проектных документов. Для автоматизации их обработки и контроля упомянутых в документах требований к разрабатываемому продукту требуется создание инструментальных средств поддержки работы аналитика при проектировании информационных систем. Для того чтобы существовала возможность выполнять контроль каждого вида требований в отдельности необходимо выделить в проектном документе разделы, описывающие конкретный вид требований к системе. Для выполнения подобного анализа необходимы средства, позволяющие описывать структуру анализируемого документа. В качестве таких средств могут быть использованы визуальные предметно-ориентированные языки, позволяющие в терминах предметной области в графической нотации выполнить описание структуры документа. В работе рассмотрен подход к созданию такого предметно-ориентированного языка. Язык имеет два уровня: первый уровень позволяет определить общую совокупность документов и связи между ними, а второй – структуру каждого проектного документа в отдельности. Разработанный предметно-ориентированный язык может быть интегрирован в систему поддержки работы аналитика при проектировании информационных систем. Это позволит, с одной стороны, на основе загруженного в систему набора проектных документов выполнить их анализ и разбор, представив разделы документов в виде отдельных элементов модели. А с другой стороны, с использованием разработанного языка аналитик может описать каждый раздел проектного документа в отдельности, а затем сгенерировать на их основе единое текстовое описание. Язык имеет простую графическую нотацию, поэтому работать с ним могут как IT-специалисты, так и заказчики, не являющиеся профессиональными программистами.*

*Проектная документация; предметно-ориентированные языки; моделирование; DSM-платформа.*

**M.A. Zhigalova, A.O. Sukhov****DOMAIN-SPECIFIC LANGUAGE FOR DESCRIBING DOCUMENTS USED IN  
INFORMATION SYSTEMS DESIGN**

*In the process of software products development and their maintenance a large number of project documents is created and used. To automate their processing and control of the product requirements referred in the documents, developing tools supporting the work of the analyst in the design of information systems is required. In order to carry out control of each type of requirements individually, it is needed to define the project document sections that describe a particular type of system requirements. To perform such analysis, tools for describing the structure of the analyzed document are required. Visual domain-specific languages can be used as these tools, allowing to perform a description of the document structure in terminology of the domain using graphical interface. The approach to creating such domain-specific language is considered. The language has two levels: the first level allows to determine the total set of documents and the relationships between them, and the second – the structure of each project document separately. The developed domain-specific language can be integrated into a system of support of analyst working in the information systems design process. On the one hand, it will allow based on a set of loaded into the system design documentation to perform their analysis and parsing, presenting each part of the documentation in the form of individual element of the model. On the other hand, using developed language the analyst*

\* Работа выполнена при финансовой поддержке Научного фонда НИУ ВШЭ (проект № 14-09-0228) и РФФИ (проекты № 14-07-31330-мол\_а и № 14-07-31273-мол\_а).

*can describe each part of the design documentation separately, and then generate on their basis uniform text description. The language has a simple graphical notation, therefore it can be used by as IT-specialists and clients who are not professional programmers.*

*Design documentation; domain-specific languages; modeling; DSM-platform.*

**Введение.** Каждый этап создания информационных систем неразрывно связан с разработкой проектных документов. Необходимость подготовки технической документации на каждом этапе выполнения работ обуславливается несколькими причинами. Прежде всего, это упрощает процесс хранения и передачи технических сведений о системе, позволяет формализовать и унифицировать взаимодействие разработчика и заказчика. Кроме того, техническая документация выполняет нормативную функцию, т.е. фиксирует взаимные обязательства участников разработки, что помогает избежать недоразумений и злоупотреблений на этапе сдачи-приемки [7, 11].

Виды и комплектность документов стандартизованы и закреплены в соответствующих ГОСТ. Однако, в силу того, что все технические документы структурно очень похожи (все они состоят из разделов и подразделов, могут включать дополнительные документы, схемы, таблицы и др.), для их описания может быть разработан специальный язык, позволяющий определять структуру документа и связи между различными документами. Это позволит автоматизировать процесс анализа исходного набора проектных документов и генерации новых. Так могут быть разработаны инструментальные средства, позволяющие извлечь из проектной документации требования, предъявляемые к разрабатываемой системе, а затем контролировать их выполнение в процессе реализации системы. С другой стороны, процессы написания проектных документов достаточно трудоемкая работа, требующая знания четкой структуры создаваемого документа. Однако данный процесс также можно автоматизировать. Средства автоматизации генерации проектной документации позволят на основе описания различных разделов документа, заданных в удобном визуальном пользовательском интерфейсе, сгенерировать единый проектный документ, который в дальнейшем может быть доработан «вручную».

**Язык описания проектной документации.** В настоящее время существуют различные языки, применяемые для описания структуры документов: XML, HTML, SGML и др. [1, 9, 12]. Однако большинство из них имеет текстовую нотацию, представляющую документ в виде дерева тегов. Такая нотация затрудняет интерпретацию структуры и не столь наглядна по сравнению с графическим представлением. Графическая нотация позволяет в удобной визуальной форме определить структуру конкретного документа, а также изобразить явные и опосредованные связи между документами.

Для описания документации, применяемой в процессе проектирования информационных систем (ИС), может быть разработан визуальный предметно-ориентированный язык. *Предметно-ориентированный язык* (Domain-Specific Language, DSL) – это язык моделирования, предназначенный для решения определенного класса задач в конкретной предметной области. В отличие от языков моделирования общего назначения, DSL более выразителен, прост в применении и понятен различным категориям специалистов, поскольку язык оперирует привычными для них терминами предметной области. Именно поэтому в настоящее время разработано большое число предметно-ориентированных языков, применяемых для создания систем разного назначения: систем искусственного интеллекта, распределенных систем, мобильных приложений, систем имитационного моделирования и др. [8]. Для поддержки процесса разработки и сопровождения DSL используется специальный вид программного обеспечения, получивший название *DSM-платформа*. Система *MetaLanguage* – DSM-платформа, предназначенная для создания визуальных предметно-ориентированных языков, построения моделей с использованием этих языков и преобразования созданных моделей в иные текстовые и графические нотации [14, 15].

Для создания DSL необходимо описать его метамодель – модель языка моделирования и задать пиктограммы, используемые для визуализации различных конструкций языка. После определения языка пользователь может приступить к созданию моделей с его использованием.

Благодаря тому, что DSM-платформа MetaLanguage в процессе своей работы выполняет интерпретацию описания метамодели вместо генерации исходного кода редактора DSL, существует возможность внесения изменений в описание языка моделирования в динамике, без перезапуска системы и модификации исходного кода. Это делает разрабатываемый язык достаточно гибким и легко настраиваемым на предметную область.

Поскольку при описании проектных документов важно не только определить их структуру, но и задать связи между ними, то разработаем двухуровневый предметно-ориентированный язык описания документов, используемых при проектировании ИС. Первый уровень языка предназначен для описания всего множества документов и взаимосвязи между отдельными документами, а второй – для описания структуры конкретного документа [6].

**Язык описания связей между документами.** При разработке языка описания связей между документами необходимо учитывать, как могут быть связаны различные документы между собой, со стандартами, на основе которых они разработаны, и на какие законодательные акты и стандарты могут ссылаться эти документы.

Базовыми конструкциями разрабатываемого языка моделирования являются:

- ◆ Сущность «Документ», которая описывает проектный документ и имеет следующие атрибуты: «Название», «Автор», «Описание».
- ◆ Сущность «Стандарт», которая описывает стандарт, в соответствии с которым разработан документ или на который ссылается документ. Атрибутами сущности являются: «Номер», «Название», «Описание».
- ◆ Сущность «Законодательный акт» определяет законодательный акт, на который ссылается документ. Данная сущность имеет следующие атрибуты: «Номер», «Название», «Дата принятия», «Описание».

Для унифицированного описания перечисленных сущностей определим абстрактную сущность «Абстрактный документ», которая имеет следующие атрибуты: «Название», «Описание». Абстрактная сущность является родительской для сущностей «Документ», «Стандарт», «Законодательный акт».

После описания сущностей метамодели следует определить отношения между ними. Метамодель разрабатываемого языка содержит следующие отношения ассоциации (рис. 1):

- ◆ «Ссылается» – это однонаправленное отношение, которое соединяет сущность «Документ» с сущностями «Стандарт» и «Законодательный акт», а также сущности «Документ», «Стандарт», «Законодательный акт» с самими собой. Данное отношение позволяет указать, что некоторый документ ссылается на конкретный стандарт, законодательный акт и/или другой документ.
- ◆ «Создан на основе» – это однонаправленное отношение, которое может соединять сущность «Документ» с сущностями «Стандарт» и «Законодательный акт», а также с самой собой. Данное отношение позволяет указать что некоторый документ создан на основе конкретного стандарта, законодательного акта и/или другого документа.

Опишем с использованием разработанного языка модель, содержащую информацию о следующих документах: техническое задание на создание автоматизированной системы, технико-экономическое обоснование и технический проект. Структура и содержание технико-экономического обоснования определены в ГОСТ 24.202-80 «Требования к содержанию документа "Технико-экономическое обоснование создания АСУ"» [3]. Техническое задание составляется на основе

ГОСТ 34.602-89 «Техническое задание на создание автоматизированной системы» [5]. Информация для написания технического проекта приводится в руководящем документе по стандартизации РД 50-34.698-90 «Автоматизированные системы. Требования к содержанию документов» [13], который содержит ссылки на ГОСТ 34.602-89 и ГОСТ 34.201-89 «Виды, комплектность и обозначения документов при создании автоматизированных систем» [4]. Разработанная модель представлена на рис. 2.

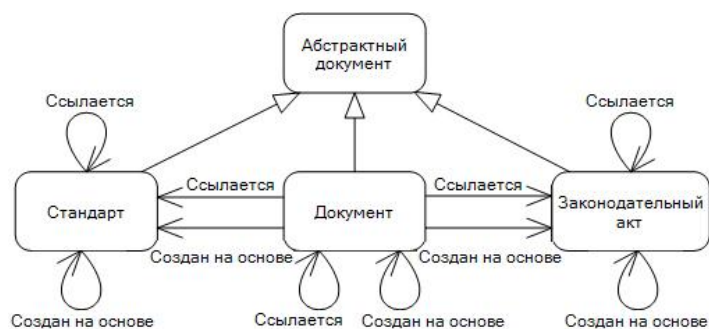


Рис. 1. Метамоделю языка описания связей между документами

**Язык описания структуры проектного документа.** Поскольку большинство проектных документов разрабатывается в соответствии со стандартами, содержащими требования к оформлению и содержанию документации, то для построения метамодели предметно-ориентированного языка, описывающего структуру проектных документов, были проанализированы следующие стандарты ГОСТ 24.202-80, ГОСТ 34.602-89, ГОСТ РД 50-34.698-90.

Анализ стандартов позволил выделить следующие основные элементы документации, применяемой в процессе проектирования ИС:

- ♦ документ, входящий в состав проектной документации (например, техническое задание, технико-экономическое обоснование и др.);
- ♦ вложенный документ, т.е. документ, не являющийся проектным, но который необходим для составления других документов (например, пояснительная записка, ведомость технического проекта и др.);

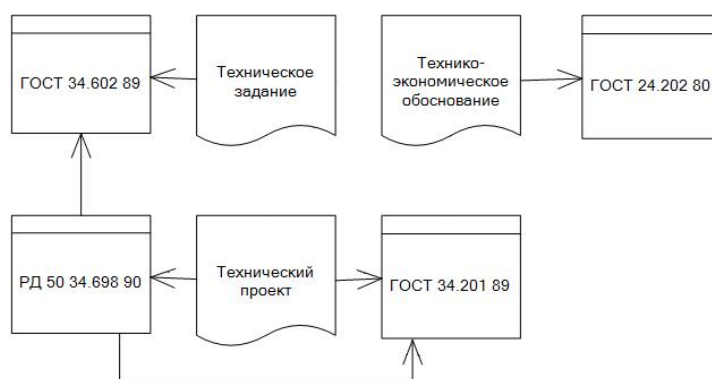


Рис. 2. Связи между техническими документами и государственными стандартами

- ♦ структурная часть документа – раздел проектного документа;
- ♦ элемент раздела документа (например, таблица, схема и др.).

Основываясь на перечисленных элементах проектной документации, опишем сущности метамодели языка (см. рис. 3).



Рис. 3. Метамодель языка описания структуры проектного документа

Абстрактная сущность «Абстрактный документ», имеющая атрибуты «Название», «Описание», является родительской для сущностей «Документ» и «Вложенный документ». Дочерние сущности собственных атрибутов не имеют.

Для унифицированного описания сущностей, соответствующих структурным частям документа, была создана абстрактная сущность «Структурная часть» с атрибутами «Название», «Порядковый номер», «Описание». Дочерними сущностями для нее являются «Раздел» и «Оptionальный раздел», которые собственных атрибутов не имеют. Сущность «Оptionальный раздел» позволяет описать раздел проектного документа, который является необязательным, но может присутствовать в некоторых документах, например, приложение к техническому заданию.

Сущность «Элемент документа» с атрибутом «Имя» является абстрактной родительской сущностью для сущностей «Схема» и «Таблица». Сущность «Схема» («Таблица») позволяет определить некоторую схему/изображение (таблицу), которую должен включать некоторый документ.

Определим отношения между сущностями. Метамодель содержит отношения наследования между всеми дочерними сущностями и их родителями, а также отношение агрегации «Включает», которое позволяет указать что некоторая сущность является частью другой сущности, так «Документ» может включать «Вложенный документ»; оба вида документа содержат «Структурные части»; структурные части, в свою очередь, включают элементы – «Таблица» и/или «Схема». Кроме того, данное отношение позволяет сущностям «Документ», «Вложенный документ», «Раздел» и «Оptionальный раздел» включать экземпляры той же сущности.

На рис. 4-5 приведено описание структуры технического задания на разработку автоматизированной системы и технико-экономического обоснования, выполненные в нотации разработанного языка описания структуры проектного документа.



Рис. 4. Описание структуры технического задания на разработку автоматизированной системы

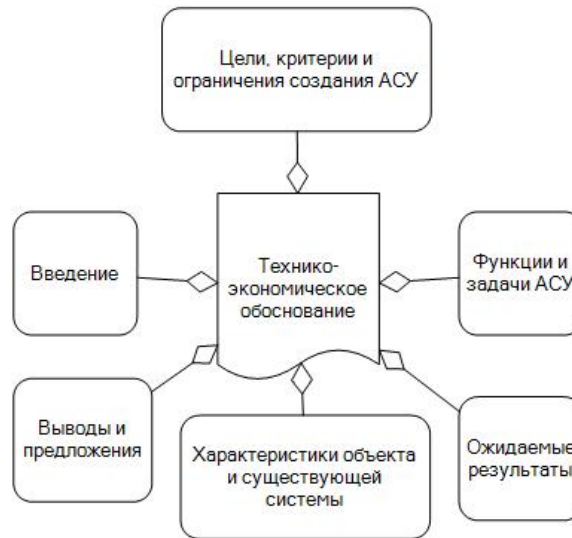


Рис. 5. Описание структуры технико-экономического обоснования

**Заключение.** Разработанный предметно-ориентированный язык описания проектной документации имеет два уровня. Первый уровень языка позволяет описывать набор документов и связи между ними, второй уровень – структуру конкретного документа.

Созданный DSL может быть интегрирован в систему поддержки работы аналитика при проектировании ИС [2, 10]. Это позволит, с одной стороны, на основе загруженного в систему набора проектных документов выполнить их анализ и

разбор, представив разделы документов в виде отдельных элементов модели. А с другой стороны, с использованием разработанного языка аналитик может описать каждый раздел проектного документа в отдельности, а затем сгенерировать на их основе единое текстовое описание. Язык имеет простую графическую нотацию, поэтому работать с системой смогут как IT-специалисты, так и заказчики, не являющиеся профессиональными программистами.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Агалакова Е.З., Ланин В.В.* Предметно-ориентированный язык описания структуры и содержания электронных документов // Сборник статей Международной научно-практической конференции «Технологии разработки информационных систем ТРИС-2014». – Таганрог: Издательство ЮФУ, 2014. – С. 70-75.
2. *Агалакова Е.З., Ланин В.В., Щучалова Ю.С.* Онтология документов, используемых при проектировании информационных систем // Вестник молодых ученых ПГНИУ: Сб. науч. тр. – Пермь: Пермский государственный национальный исследовательский университет, 2014. – Вып. 4. – С. 253-258.
3. ГОСТ 24.202-80. Требования к содержанию документа "Технико-экономическое обоснование создания АСУ". – Введ. 1981-01-01.
4. ГОСТ 34.201-89. Виды, комплектность и обозначения документов при создании автоматизированных систем. – Введ. 1990-01-01.
5. ГОСТ 34.602-89. Техническое задание на создание автоматизированной системы. – Введ. 1990-01-01.
6. *Жигалова М.А., Сухов А.О.* Валидация проектной документации на основе предметно-ориентированного языка // Вестник молодых ученых ПГНИУ: Сб. науч. тр. – Пермь: Пермский государственный национальный исследовательский университет, 2014. – Вып. 4. – С. 224-228.
7. *Заболеева-Зотова А.В., Орлова Ю.А.* Автоматизация процедур семантического анализа текста технического задания // Известия Волгоградского государственного технического университета. – 2007. – Т. 9, № 3. – С. 52-55.
8. *Замятина Е.Б., Лядова Л.Н., Сухов А.О.* Мультиязыковое моделирование с использованием DSM-платформы MetaLanguage // Информатизация и связь. – 2013. – № 5. – С. 11-14.
9. *Клышинский Э.С.* Перспективные методы обработки проектной документации // Электронные библиотеки: Перспективные методы и технологии, электронные коллекции: Труды XII Всероссийской научной конференции RCDI'2010. – Казань: Казанский университет, 2010. – С. 129-134.
10. *Ланин В.В.* Интеллектуальная система поддержки работы аналитика при проектировании информационной системы // Научно-технический вестник Поволжья. – 2014. – № 6. – С. 197-199.
11. *Орлова Ю.А.* Методика анализа текста технического задания // Известия Тульского государственного университета. Технические науки. – 2011. – № 3. – С. 213-220.
12. *Панов С.А.* Формальный язык описания структуры документов и его интерпретация в формат метода компонентных цепей // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2014. – № 1 (31). – С. 197-200.
13. РД 50-34.698-90. Автоматизированные системы. Требования к содержанию документов. – Введ. 1992-01-01.
14. *Сухов А.О.* Инструментальные средства создания визуальных предметно-ориентированных языков моделирования // Фундаментальные исследования. – 2013. – № 4 (Ч. 4). – С. 848-852.
15. *Sukhov A.O., Lyadova L.N.* An Approach to Development of Visual Modeling Toolkits // Advances in Information Science and Applications. Volumes I & II. Proceedings of the 18th International Conference on Computers (part of CSCC'14). – Vol. I. – Santorini Island: CSCC, 2014. – P. 61-66.

## REFERENCES

1. Agalakova E.Z., Lanin V.V. Predmetno-orientirovanny yazyk opisaniya struktury i soderzhaniya elektronnykh dokumentov [Domain-specific language for describing structure and content of electronic documents], *Sbornik statey Mezhdunarodnoy nauchno-prakticheskoy konferentsii «Tekhnologii razrabotki informatsionnykh sistem TRIS 2014»* [Collection of articles of International scientific-practical conference "Technologies of information systems development TRIS 2014"]. Taganrog: Izdatel'stvo YuFU, 2014, pp. 70-75.
2. Agalakova E.Z., Lanin V.V., Shuchalova Yu.S. Ontologiya dokumentov, ispol'zuemykh pri proektirovanii informatsionnykh sistem [The ontology of documents used in the design of information systems], *Vestnik molodykh uchenykh PGNIU: Sb. nauch. tr.* [Bulletin of young scientists of RUSSIA: Collection of yuchih works]. Perm': Permskiy gosudarstvennyy natsional'nyy issledovatel'skiy universitet, 2014, Issue 4, pp. 253-258.
3. GOST 24.202 80. Trebovaniya k soderzhaniyu dokumenta "Tekhniko-ekonomicheskoe obosnovanie sozdaniya ASU" [State Standard 24.202 80. Requirements to the content of the document "Feasibility study of creating automated"]. Vved. 1981–01–01.
4. GOST 34.201 89. Vidy, komplektnost' i oboznacheniya dokumentov pri sozdanii avtomatizirovannykh sistem [State Standard 34.201 89. Types, completeness and designation of documents for creation of automated systems]. Vved. 1990–01–01.
5. GOST 34.602 89. Tekhnicheskoe zadanie na sozдание avtomatizirovannoy sistemy [State Standard 34.602 89. The technical project on creation of the automated system]. Vved. 1990–01–01.
6. Zhigalova M.A., Sukhov A.O. Validatsiya proektnoy dokumentatsii na osnove predmetno-orientirovannogo yazyka [Validation of the project documentation on the basis of object-oriented language], *Vestnik molodykh uchenykh PGNIU: Sb. nauch. tr.* [Bulletin of young scientists of RUSSIA: Collection of scientific papers]. Perm': Permskiy gosudarstvennyy natsional'nyy issledovatel'skiy universitet, 2014, Issue 4, pp. 224-228.
7. Zaboloeva-Zotova A.V., Orlova Yu.A. Avtomatizatsiya protsedur semanticheskogo analiza teksta tekhnicheskogo zadaniya [Automation of procedures of semantic analysis of the text of the technical specifications], *Izvestiya Volgogradskogo gosudarstvennogo tekhnicheskogo universiteta* [News of the Volgograd State Technical University], 2007, Vol. 9, No. 3, pp. 52-55.
8. Zamyatina E.B., Lyadova L.N., Sukhov A.O. Mul'tiyazykovoe modelirovanie s ispol'zovaniem DSM-platformy MetaLanguage [Multi-language modeling using the DSM-platform MetaLanguage], *Informatizatsiya i svyaz'* [Informatization and Communication], 2013, No. 5, pp. 11-14.
9. Klyshinskiy E.S. Perspektivnye metody obrabotki proektnoy dokumentatsii [Advanced methods of processing of project documentation], *Elektronnye biblioteki: Perspektivnye metody i tekhnologii, elektronnye koleksii: Trudy XII Vserossiyskoy nauchnoy konferentsii RCDI'2010* [Digital libraries: advanced methods and technologies, digital collections: proceedings of the XII all-Russian scientific conference RCDI'2010]. Kazan': Kazanskiy universitet, 2010, pp. 129-134.
10. Lanin V.V. Intellektual'naya sistema podderzhki raboty analitika pri proektirovanii informatsionnoy sistemy [Intelligent system support analyst in the design of information systems], *Nauchno-tekhnicheskiiy vestnik Povolzh'ya* [Scientific and Technical Volga region Bulletin], 2014, No. 6, pp. 197-199.
11. Orlova Yu.A. Metodika analiza teksta tekhnicheskogo zadaniya [The method of analysis of the text of the technical specifications] *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki* [Izvestiya of the Tula State University. Technical Science], 2011, No. 3, pp. 213-220.
12. Panov S.A. Formal'nyy yazyk opisaniya struktury dokumentov i ego interpretatsiya v format metoda komponentnykh tsepey [A formal language for describing the structure of documents and his interpretation of the format method component circuits], *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki* [Reports of Tomsk state University of control systems and Radioelectronics], 2014, No. 1 (31), pp. 197-200.
13. RD 50 34.698 90. Avtomatizirovannyye sistemy. Trebovaniya k soderzhaniyu dokumentov [The automated control system. Requirements for the content of documents]. Vved. 1992–01–01.
14. Sukhov A.O. Instrumental'nye sredstva sozdaniya vizual'nykh predmetno-orientirovannykh yazykov modelirovaniya [Tools for creating visual domain-specific modeling languages] *Fundamental'nye issledovaniya* [Fundamental Research], 2013, No. 4 (Part 4), pp. 848-852.
15. Sukhov A.O., Lyadova L.N. An Approach to Development of Visual Modeling Toolkits, *Advances in Information Science and Applications*. Vol. I & II. Proceedings of the 18th International Conference on Computers (part of CICC'14), Vol. I. Santorini Island: CICC, 2014, pp. 61-66.

Статью рекомендовал к опубликованию д.т.н. О.Г. Пенский.



**Жигалова Мария Александровна** – Национальный исследовательский университет «Высшая школа экономики»; e-mail: mashuzden@mail.ru; 614070, г. Пермь, бул. Гагарина, 37а; тел. +73422545608; кафедра информационных технологий в бизнесе; студентка.

**Сухов Александр Олегович** – e-mail: ASuhov@hse.ru; кафедра информационных технологий в бизнесе; к.ф.-м.н., старший преподаватель.

**Zhigalova Maria Aleksandrovna** – National Research University “Higher School of Economics”; e-mail: mashuzden@mail.ru; 37a, Gagarina street, Perm, 614070, Russia; phone: +73422825372; the department of information technologies in business; student.

**Sukhov Alexander Olegovich** – e-mail: ASuhov@hse.ru; the department of information technologies in business; cand. of phys.-math. sc.; senior lecturer.

УДК 004.652

**А.В. Чернов, В.И. Янц, Е.В. Карпенко**

### **ПРИМЕНЕНИЕ RADIX ДЕРЕВЬЕВ ДЛЯ ИНДЕКСАЦИИ СЛАБОСТРУКТУРИРОВАННЫХ ДАННЫХ\***

*В системах управления слабоструктурированными данными остро стоит проблема осуществления поиска в больших наборах данных. При осуществлении поиска выполняется значительно число дисковых операций, т.к. для работы с любыми видами данных требуется извлечь их из постоянного хранилища, при этом, на сегодняшний день, дисковая подсистема является самым медленным элементом компьютера. Как следствие, рост количества дисковых операций неминуемо приводит к быстрому исчерпанию вычислительных ресурсов СУБД. Для решения данной проблемы предлагается использовать Radix деревья с модифицированной многослойной структурой для построения поискового индекса. Каждый слой в такой структуре представляет собой дополнительное Radix дерево. Внутри слоя, ветви дерева помещены в структуры соответствующие размеру дискового блока. Переход между блоками осуществляется при помощи двух видов ссылок – маркированных и немаркированных. Ссылки связывают слои и позволяют переходить либо к конкретному, результирующему элементу в слое, либо к блоку данных содержащему ссылку на результирующий элемент дерева в следующем слое. Такая структура позволит выполнять не более 1-й дисковой операции на каждый слой индекса. При этом количество слоев в полученной структуре будет зависеть от объема данных и степени разнородности и лишь в редких случаях превысит число 3. Данная многоуровневая структура легко поддается кэшированию и позволяет сократить количество дисковых операций до одной на один поисковый запрос. Полученная структура позволяет нивелировать недостатки несбалансированных Radix деревьев и получить быстрый и компактный поисковый индекс для применения в широком круге систем управления слабоструктурированными данными, обеспечивая высокую скорость работы с наборами неоднородных, длинных и сложных строк.*

*Слабоструктурированные данные; индекс; поиск; базы данных, нагруженные деревья, radix деревья, сжатие данных; системы управления базами данных.*

**A.V. Chernov, V.I. Yants, E.V. Karpenko**

### **ADAPTATION OF RADIX TREES FOR INDEXING SEMISTRUCTURED DATA**

*In semi-structured data management systems there is acute problem of searching in large data sets. During search implementation, considerable number of disk operations is being performed, which inevitably leads to a rapid depletion of computing resources database with the increase of queries number. To solve this problem, we propose to use Radix tries with a modi-*

\* Работа выполнена при финансовой поддержке РФФИ, проекты 13-01-00325 А, 15-01-03067 А.