

6. Крон, Г. Тензорный анализ сетей / Г. Крон. М. : Сов. радио, 1978.

7. Крон, Г. Исследование сложных систем по частям. Диакоптика / Г. Крон. М. : Наука, 1972.

8. Веревкина, Е. В. Тензорная методология исследования нагрузки в информационных сетях : монография / Е. В. Веревкина, О. А. Корякина, М. Н. Петров ; под ред.

проф. М. Н. Петрова ; НИИ систем упр., волновых процессов и технологий. Красноярск, 2004.

9. Петров, М. Н. Тензорная методология в информационных сетях : монография / М. Н. Петров, Е. В. Веревкина, М. О. Захарченко ; под ред. проф. М. Н. Петрова; НИИ систем упр., волновых процессов и технологий. Красноярск, 2001.

K. V. Kolesov

DIACOPTICS METHOD APPLICATION FOR SOFTWARE RELIABILITY ANALYSIS

The opportunity of the diacoptics method application for the algorithms reliability of the various software and managing systems research is considered. The value of this method is extremely actual now in connection with a sharp increase in complexity of algorithms structures and considerable financial expenses for the software development.

Key words: reliability, diacoptics, tensor analysis, software.

УДК 681.3

М. В. Карасева, П. В. Зеленков

РЕАЛИЗАЦИЯ МОДУЛЯ ПОИСКА ИНФОРМАЦИИ НА ОСНОВЕ ИСПОЛЬЗОВАНИЯ МУЛЬТИЛИНГВИСТИЧЕСКИХ ТЕЗАУРУСОВ

Предложен модуль поиска, ранжирования и определения уровня релевантности документов путем использования метапоисковых мультилингвистических алгоритмов обработки информации и управления.

Ключевые слова: ранжирование, уровень релевантности, метапоиск.

В настоящее время разработано множество моделей и алгоритмов для представления информации в распределенных информационных системах. Частным случаем подобных систем являются информационно-управляющие системы, корпоративные информационные системы и интенсивно развивающиеся системы поддержки принятия решения. Однако большинство моделей распределенных систем строятся на основе одноязычного представления информации или учитывают многоязычность неявно [1].

Сейчас активно ведутся работы по созданию новых алгоритмов поиска, ранжирования и определения релевантности информации как в глобальной сети Интернет, так и во всевозможных локальных информационно-управляющих системах. Одним из перспективных направлений при разработке новых моделей и алгоритмов обработки информации является применение предметных словарей, или тезаурусов. Тезаурус – это максимально полный объем лексики, организованной по тематическому (семантическому) принципу с отражением определенного набора базовых семантических отношений, являющихся полным систематизированным набором данных о какой-либо области знаний, который позволяет человеку или вычислительной машине в ней ориентироваться. Необходимо отметить, что в современных системах подобные словари-тезаурусы очень редко встречаются представленными в мультилингвистической частотной реализации. Авторами

статьи в рамках предлагаемых ими моделей применяются тезаурусы, выполненные на основе мультилингвистической технологии для проведения поисковой процедуры в информационных системах.

Данный подход направлен в первую очередь на решение проблемы многоязычного представления информации в информационно-управляющих системах. В современных условиях даже небольшие корпоративные информационные системы работают в мультилингвистическом режиме. Нередко управленческому персоналу в процессе подготовки и принятия решения требуется своевременное предоставление документов, принадлежащих различным языковым множествам. Значительно возрастают требования к оперативности и качеству функционирования систем поиска мультилингвистической информации, используемых лицом, принимающим решение, в различных ситуациях.

В рамках предлагаемого авторами модуля основная работа по процессу поиска, ранжирования и определению уровня релевантности производится путем использования метапоисковых мультилингвистических алгоритмов обработки информации и управления [2]. Для этого сначала следует определить параметры процесса поиска. К ним относятся функции выбора предметной области и настройки языковых множеств, в рамках которых необходимо производить поиск.

Кроме того, необходимо отдельно показать возможность работы со строкой поиска информации как в Интернете, так и в корпоративной сети. Согласно предлагаемому авторами подходу, работа с поисковой строкой может проводиться в двух режимах:

- режиме ручного ввода строки поиска;
- режиме автоматизированного формирования строки поиска.

При вводе строки поиска в ручном режиме система проверяет наличие введенных термов в частотном мультилингвистическом тезаурусе. В случае отсутствия термина в словаре пользователю предлагается ввести поисковую строку с изменением термов в строке поиска.

Процесс формирования запроса по заданной предметной области при вводе строки поиска в автоматическом режиме [3], рассмотрим более подробно. Модуль поиска информации основан на применении частотных мультилингвистических тезаурусов, которые повышают качество определения релевантности документов по запросам. Эти тезаурусы позволяют выделить направленность документа, вплоть до нахождения предметной области, к которой принадлежит документ. Качество определения уровня релевантности в предлагаемом решении соответствует уровню релевантности в каталожных системах ручной индексации [1]. Исходя из частотных характеристик терминов можно по заданному алгоритму сформировать поисковую строку, корректируемую пользователем или дополняемую им при необходимости.

Следует отметить, что в современных корпоративных информационных системах может храниться и мультилингвистическая информация. Однако пользователь поискового модуля не может знать всех языков, представленных

в сети. Поэтому нужно учитывать конкретность указания языковых множеств, необходимых пользователю.

После завершения процесса формирования поисковой строки и указания языков, в которых будет проходить поиск информации, необходимо приступить непосредственно к процедуре поиска [3]. В результате происходит последовательный опрос всех информационных корпоративных ресурсов и формируется массив ссылок на интересующие пользователя документы, а также проводится разбиение всего множества ссылок по принципу принадлежности к языковому множеству.

Кроме того, пользователь может увидеть следующую дополнительную информацию, которая учитывается при ранжировании документов и определении уровня релевантности каждого найденного документа:

- заголовок документа;
- объем документа;
- количество найденных термов в документе.

Это является первым шагом обработки процедуры поиска.

На втором шаге происходит определение уровня релевантности и ранжирование мультилингвистического массива ссылок. Здесь пользователю предоставляется дополнительная информация уже другого рода:

- об уровне релевантности найденного документа;
- общем весе релевантных термов в документе.

Третий шаг – это непосредственный просмотр найденных документов. На данном шаге можно не только просмотреть документ, но и получить о нем расширенную информацию, представленную, например, в виде таблицы, составленной при завершении процесса обработки англоязычного документа:

Термин	Частота	Вес термина в документе	Вес термина в тезаурусе
...	...	...	...
activity	8	0,002	0,000 008 17
process	3	0,000 9	0,000 010 8
search	1	0,000 01	0,000 001 76
...	...	...	...

Эти характеристики важны при расчете уровня релевантности и ранжирования найденных документов [2]. Рассмотрим структуру данной таблицы:

- термин – список терминов, которые встретились в тезаурусе и документе;
- частота показывает, сколько раз данный термин встретился в документе;
- вес термина в документе рассчитывается относительно частоты термина и общего количества термов в документе;
- вес термина в тезаурусе рассчитывается как отношение частотной характеристики термина в словаре к общей суммарной частотной характеристике всех термов словаря.

Предлагаемый авторами модуль поиска и обработки информации в корпоративных системах поддержки принятия решений полностью удовлетворяет требованиям к системам подобного уровня и позволяет решить проблему организации, хранения и обработки информации в современных распределенных мультилингвистических корпоративных системах поддержки принятия решений.

Реализация метапоисковых принципов способствует охвату индексов самых популярных поисковых веб-служб, при этом сокращается объем нерелевантных ссылок в результатах поиска, существенно повышается качество обработки запросов пользователя и уменьшается объем трафика при формировании собственной базы релевантных документов.

Кроме того, представленные мультилингвистические модели позволяют составить более гибкие мультилингвистические ответы даже на одноязычные запросы по сравнению с простой распределенно-информационной системой, с учетом неопределенности описания как мультилингвистических документов, так и запросов.

Библиографический список

1. Зеленков, П. В. Мультилингвистическая модель распределенной системы на основе тезауруса / П. В. Зеленков, И. В. Ковалев, М. В. Карасева, С. В. Рогов // Вестник СибГАУ. 2008. № 1 (18). С. 26–27.

2. Метапоисковая мультилингвистическая система поиска узкоспециализированной информации / И. Н. Карцан, П. В. Зеленков, Д. А. Рагзин и др. М., 2007. Зарег. во Всерос. научн.-техн. информ. центре, № 50200701673, рег. № ОФАП 8891.

3. Зеленков, П. В. Проблема развития метапоисковых технологий / П. В. Зеленков, Т. А. Ковалева // Вестник НИИ СУВПТ : сб. науч. тр.: НИИ систем упр., волновых процессов и технологий. Вып. 14. Красноярск, 2004. С. 95–103.

M. V. Karaseva, P. V. Zelenkov

REALIZATION OF THE DATA SEARCH MODULE ON THE BASES OF MULTILINGUAL THESAURUSES

The model of searching, ranking and relevance level of documents determination is introduced. It is done with the help of meta-search multilingual algorithms of data processing and control.

Keywords: ranking, relevance level, meta-search.

УДК 004.052.3

П. В. Ковалев, А. Н. Лайков, С. Н. Гриценко

ОПРЕДЕЛЕНИЕ НАДЕЖНОСТИ МУЛЬТИВЕРСИОННОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ АНАЛИЗА СЕТЕЙ

Предложена методика определения надежности мультиверсионного программного обеспечения, которая позволяет использовать алгоритмы и методы анализа сетей для определения надежности программного обеспечения, разработанного с использованием мультиверсионного подхода.

Ключевые слова: надежность, мультиверсионность, сети, методы оценки, методы анализа сетей.

На заре компьютерной эры первостепенной задачей являлось развитие аппаратных компьютерных средств. В первую очередь это было обусловлено высокой стоимостью обработки и хранения данных. Однако впоследствии успехи микроэлектроники привели к резкому увеличению производительности компьютеров при значительном снижении их стоимости. поэтому основной задачей в 1990-х гг. и начале XXI в. стало совершенствование качества компьютерных приложений, возможности которых целиком определяются программным обеспечением (ПО).

Современный персональный компьютер теперь имеет производительность большой ЭВМ 1980-х гг. Сняты практически все аппаратные ограничения на решение задач, а оставшиеся ограничения приходятся на долю ПО.

Чрезвычайную актуальность приобрели следующие проблемы:

- аппаратная сложность опережает умение строить ПО, использующее потенциальные возможности аппаратуры;
- умение строить новые программы отстает от требований к новым программам;
- реализации возможностей по эксплуатации существующих программ угрожает низкое качество их разработки.

Ключом к решению этих проблем является грамотная организация процесса создания ПО и реализация технологических принципов промышленного конструирования программных систем (ПС) [1]. Кроме того, существует ряд способов повышения надежности программ-

ного обеспечения, одним из которых является введение избыточности.

Избыточность как метод повышения надежности ПО.

За последние несколько лет тема мультиверсионного программного обеспечения не раз затрагивалась в различных диссертационных работах, например в [2]. Согласно этой работе, сбой программных систем может повлечь за собой большие потери и иметь весьма серьезные последствия. Поскольку абсолютная уверенность в безупречности программных средств достигается редко, то для выполнения требований к надежности проекта применяются методы повышения отказоустойчивости программного обеспечения. Программная отказоустойчивость достигается благодаря использованию алгоритмов программирования и методов разработки ПО, которые повышают вероятности того, что конечная реализация проекта приведет к правильному и (или) безопасному результату. Так как правильность и безопасность – концепции системного уровня, то потребность и степень использования программной отказоустойчивости непосредственно зависит от предназначения приложения и полного проектирования систем [2].

Избыточность, применяемая для обеспечения надежности функционирования комплекса программ (КП), используется прежде всего для контроля и селекции искажений вычислительного процесса или данных и для выработки мер по снижению последствий этих аномалий. Основная задача состоит в ограничении или исключении возможности аварийных последствий, соответствующих отказу системы в процессе функционирования.