

ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ В ЗАДАЧАХ РАЦИОНАЛЬНОГО РАСКРОЯ И УПАКОВКИ

Балабанов В.Н., Скобцов Ю.А., Фонов А.М.

Донецкий национальный технический университет, г. Донецк
кафедра автоматизированных систем управления

E-mail: akavrt@gmail.com, skobtsov@kita.dgtu.donetsk.ua, stas@nedra.dn.ua

УДК 681.518

Abstract

Balabanov V.N., Skobtsov Y.A., Fonotov A.M. Dataware issues in cutting stock and packing problems. *There are several typical problems facing researcher in area of operational research. In this paper we consider the efficient way to tackle some of them. It starts off with a brief overview of NestingXML format that is used to describe cutting and packing problems. Then we discuss testing with benchmark data sets. We finally conclude with some advices on publishing over the Internet results and other related materials obtained from individual researches in order to increase scientific collaboration and informational exchange in this specific area.*

Информационное обеспечение в задачах раскроя и упаковки. Существует ряд задач, которые относятся к области исследования операций. В статье предлагается эффективный метод решения одной из них. В начале статьи приводится краткий обзор встраиваемого XML формата для описания задач раскроя и упаковки. Далее обсуждается тестирование на основе наборов данных из бенчмарков. Статья завершается рекомендациями по публикации в интернет результатов и сопутствующих материалов полученных от различных исследователей с целью улучшения научного взаимодействия по данной проблеме.

Інформаційне забезпечення в задачах розкрою й упаковки. Існує ряд задач, які відносяться до області дослідження операцій. У статті пропонується ефективний метод рішення однієї з них. На початку статті приводиться короткий огляд XML формату, який застосовується для опису задач розкрою й упаковки. Далі обговорюється тестування на основі наборів даних з бенчмарков. Стаття завершується рекомендаціями з публікації в інтернет результатів і супутніх матеріалів, які отримані від різних дослідників, з метою поліпшення наукової взаємодії по даній проблемі.

Общая постановка проблемы. Оптимизационные задачи комбинаторной природы, подобные задачам рационального раскроя и упаковки, характерны для многих отраслей промышленности. Значительный экономический потенциал и практическая применимость стимулируют активные исследования по разработке и апробации методов решения таких задач.

Впервые задача о рациональном раскрое была сформулирована Канторовичем в работе «Математические методы организации и планирования производства» [1]. Рассмотренная проблема заключалась в нахождении такого способа раскроя материала на

комплекты заготовок, при котором отходы производства были бы минимальными. Постановка задачи, формализованная Канторовичем, выглядит следующим образом.

Комплект состоит из заготовок n различных типов. Количество заготовок каждого типа, которое необходимо раскроить, предварительно задано и равняется b_j . Пусть a_{ij} — число заготовок вида j , получаемых при раскрое одной части материала i -тым способом. Допускается возможность свободного выбора частей материала из нескольких типов, c_i — ценность части, использованной в i -том раскрое. Обозначим через x_i употребимость (в среднем на один комплект заготовок) i -того раскроя. Если допустить, что все возможные частные раскрои перечислены, то задача составления плана раскроя заключается в нахождении неизвестных x_i , удовлетворяющих следующим требованиям (1) – (3):

$$\sum_{i=1}^N a_{ij} x_i \geq b_j \quad (j = 1, \dots, n) \quad (1)$$

$$x_i \geq 0 \quad (i = 1, \dots, N) \quad (2)$$

$$\min \sum_{i=1}^N c_i x_i \quad (3)$$

Подобная формулировка представляет собой одну из типовых задач линейного программирования. В дальнейшем были предложены и другие модели, основанные на использовании графов, ациклических сетей и т.д. [2]. В обзорных работах 70-ых годов была выявлена устойчивая взаимосвязь между задачами упаковки и раскроя, обоснована целесообразность их совместного рассмотрения как NP-полных задач комбинаторной оптимизации. Первая значительная работа, посвященная вопросам систематизации задач раскроя и упаковки, была опубликована Dycckhoff в 1990 году [3]. Разработанная система классификации обладала некоторыми недостатками, поэтому в недавней работе была предложена улучшенная типология [4], основанная на применении многоуровневого подхода в классификации задач раскроя и упаковки.

Информационное обеспечение в задачах рационального раскроя и упаковки.

Практически в каждом исследовании, предполагающем разработку собственных программных средств и их последующее использование для проведения различных экспериментов, необходимо ответить на ряд типовых вопросов, продиктованных специфическими особенностями жизненного цикла, характерного для большинства научных проектов, связанных с операционными исследованиями, среди прочего включающими задачи рационального раскроя и упаковки.

Следует выбрать существующие или разработать собственные входные и выходные форматы для хранения и передачи информации; найти или сгенерировать тестовые наборы для проведения численных экспериментов и сравнительного анализа эффективности реализованного алгоритма решения задачи; уделить внимание публикации не только результатов, но и обеспечению возможности свободного доступа к исходным кодам и соответствующей документации.

Казалось бы, решением большинства из этих проблем следует озаботиться лишь по мере их возникновения, сосредоточившись на выполнении текущих задач, непосредственно связанных с проблемно-ориентированной спецификой частной задачи раскроя или упаковки. Однако, как подсказывает опыт, недооценка важности первичного проектирования и предварительного планирования впоследствии может значительно замедлить прогресс в работе над всем исследованием. Подобные негативные явления характеризуются необходимостью написания дополнительного кода для адаптации к новым форматам данных, изменением структуры или способа взаимодействия программных модулей, нарушающим их работоспособность, и т.д. Каждая из трех вышеупомянутых проблем применительно к задачам рационального раскроя и упаковки требует более детального рассмотрения.

Входные и выходные форматы. В рамках каждого научного проекта исследователь-разработчик самостоятельно решает, каким образом он будет хранить входные данные,

описывающие конкретную проблему, результаты промежуточных вычислений (если таковые требуется хранить) и решения. Обычно для этих целей используются текстовые файлы различной структуры, которые в случае необходимости снабжаются краткими пояснениями и размещаются в Интернете. Следовательно, входные и выходные данные хранятся в виде совокупности файлов, зачастую плохо документированных, снабженных отрывочными пояснениями и распространяемых чаще всего в виде отдельного архива. Для последующего использования подобных наборов данных приходится прилагать значительные усилия, каждый раз заново изучать структуры файлов, писать дополнительный код, адаптирую собственные программные модули к новым форматам представления данных.

Очевидным решением озвученных проблем является применение в рамках различных исследований единого стандартного формата, снабженного общедоступной документацией и учитывающего базовую логическую структуру задач раскроя и упаковки. Подобный формат, названный авторами NestingXML (название отражает тот факт, что созданный формат является одним из диалектов метаязыка eXtensible Markup Language, сокр. XML), был представлен в недавней публикации [5]. Создание, разработка новых версий и последующая популяризация формата осуществляются при поддержке EURO Special Interest Group on Cutting and Packing (ESICUP) [6]. Язык XML, как основа для разрабатываемого формата, был выбран не случайно. Наличие большого количества редакторов и библиотек, предназначенных для работы с XML, обеспечивает удобство и эффективность его применения, а также исключает низкоуровневый разбор синтаксических конструкций и позволяет работать с данными напрямую.

Следует отметить, что создатели формата NestingXML уделили значительное внимание не только разработке синтаксических элементов, учитывающих специфику задач раскроя и упаковки, но также и решению проблем, связанных с отдельным хранением входных и выходных данных. Структура файлов этого формата изначально предполагает совместное хранение в одном файле *любых* данных, в том числе и результатов промежуточных вычислений.

Структурно файл формата NestingXML разделен на пять частей (см. рис.1). Первая секция включает элементы *name*, *author*, *date* и *description*, содержащие описательную информацию о конкретном наборе данных. Следующая секция предназначена для формального описания задачи и содержит элементы *boards* и *lot*, которые соответствуют наборам больших и малых объектов (раскраиваемый материал или пространство, доступное для размещения упаковываемых предметов, и отрезаемые или упаковываемые части, соответственно).

<pre><?xml version="1.0" encoding="UTF-8" ?> <nesting version="1.0"></pre>	
<pre><name>example</name> <author>John Doe</author> <date>01/01/09</date> ...</pre>	<p>краткое описание (метаданные)</p>
<pre><problem> <boards> <piece id="board1" quantity="1"> ... </piece> ... </boards> <lot> <piece id="piece1" quantity="1"> ... </piece> ... </lot> </problem></pre>	<p>формальное описание задачи раскроя или упаковки</p>
<pre><nfps> ... </nfps> <ifps> ... </ifps></pre>	<p>результаты промежуточных вычислений</p>
<pre><polygons> <polygon id="poly0" nVertices="8"> ... </polygon> ... </polygons></pre>	<p>геометрические данные</p>
<pre><solutions> <solution> ... </solution> </solutions></pre>	<p>решения задачи</p>
<pre></nesting></pre>	

Рисунок 1 — Структура файла формата NestingXML

Каждый из объектов описывается при помощи полигонов, тогда как геометрическая информация отделена от формального описания и вынесена в отдельный раздел, которому соответствует элемент *polygons*. Распределение перечней объектов и геометрических описаний по разным секциям введено во избежание дублирования данных, соответствующих элементам с одинаковой геометрической формой. Вместо повторного указания данных используется ссылка на одну и ту же запись соответствующего полигона. Подробное описание всех синтаксических элементов может быть найдено в соответствующей документации [5].

Формат NestingXML предлагает только базовую функциональность, которая может не учитывать специфику частной задачи раскроя или упаковки. Возможно самостоятельное расширение формата за счет добавления собственных синтаксических конструкций. Например, может быть введена система приоритетов, присваиваемых раскраиваемым или упаковываемым объектам. Описание геометрических объектов неправильной формы в рамках формата NestingXML осуществляется посредством аппроксимации кривых соответствующими полигонами с заданной степенью точности. Текущая версия формата не предполагает непосредственного описания кривых, возможно, подобная функциональность, будет реализована в дальнейшем.

Разумеется, формат NestingXML не лишен некоторых недостатков. Совместное хранение входных, промежуточных и выходных данных в перспективе может привести к значительному увеличению размеров файла. Перечень программного обеспечения, специально разработанного для работы с файлами этого формата, пока что ограничен лишь несколькими утилитами, позволяющими экспортировать содержимое файлов в некоторые популярные форматы. Графический редактор, который позволит визуализировать процесс работы с решениями задач раскроя и упаковки, по-прежнему находится в стадии разработки.

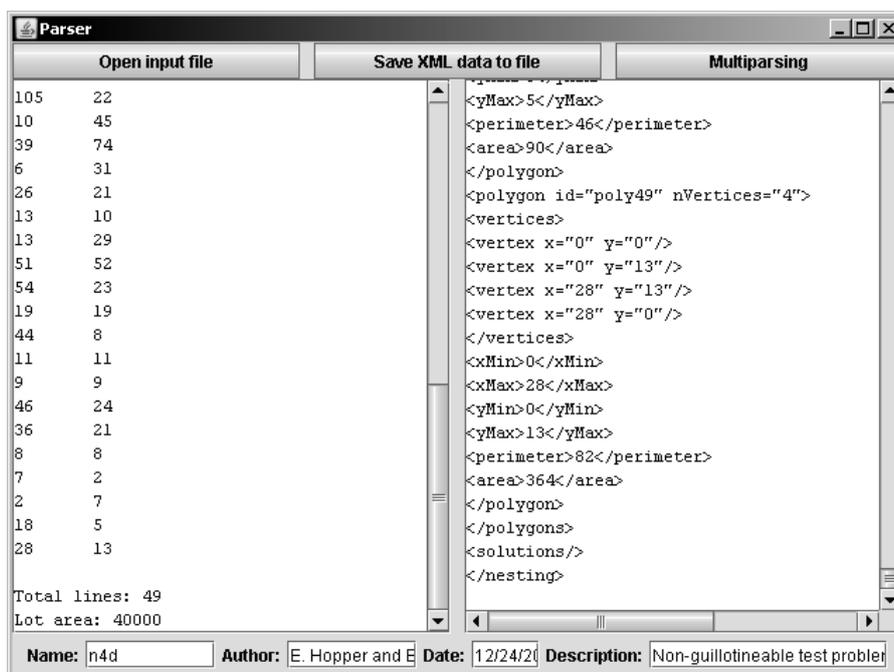


Рисунок 2 — Интерфейс парсера, реализованный с использованием средств Swing

Фактически, каждый исследователь должен самостоятельно реализовывать программную поддержку формата NestingXML в своей системе, основываясь на текущей версии стандарта и доступной xsd-схеме, позволяющей проверить корректность сохраняемых данных и структуры файла. В качестве примера подобной реализации был создан парсер (см. рис.2), позволяющий конвертировать тестовые наборы данных из [7], представленные в виде текстовых файлов, в файлы формата NestingXML. При написании парсера использовался язык программирования Java и библиотека Swing, позволяющая создавать кросс-платформенные приложения с графическим пользовательским интерфейсом.

К сожалению, перечень тестовых наборов данных, доступных в формате NestingXML, все еще значительно ограничен, что в некоторой степени препятствует дальнейшей популяризации формата. Разумеется, эта проблема носит лишь временный характер. Преимущества, предлагаемые этим форматом, позволяют рекомендовать его использование в качестве необходимого условия проведения любого нового исследования, связанного с решением задач рационального раскроя и упаковки.

Подбор тестовых данных. Практически каждая научная работа предполагает оценку эффективности нового или модификации уже существующего метода решения задачи, разработанного в рамках конкретного исследования. Во многих работах приводятся численные эксперименты, основанные на сравнительном анализе качества получаемых решений; вычислительной сложности, оцениваемой как временем решения отдельных

тестовых задач, так и количеством итераций; реже изучается стабильность работы алгоритма для нескольких последовательных запусков и т.д. Применимость и переносимость результатов подобных вычислительных экспериментов в силу ряда объективных причин значительно ограничена. Конкретные показатели могут значительно зависеть как от используемых тестовых задач, так и вычислительных сред, применяемых для их численного решения.

Унификация вычислительных сред, используемых в различных исследованиях, является задачей трудновыполнимой и практически нереализуемой. В тоже время применение одинаковых наборов тестовых данных позволяет в дальнейшем использовать показатель качества получаемых решений не только в рамках одной научной работы, но и при сравнении эффективности работы алгоритмов и их модификаций, разработанных в рамках различных исследований. Наряду с реальными тестовыми наборами зачастую используются синтетические искусственно сгенерированные данные, которые могут реализовывать предварительно заданные свойства (например, известно оптимальное решение, преобладает определенное соотношение длин сторон раскраиваемых деталей или упаковываемых предметов и т.д.).

Для того чтобы иметь возможность оценить эффективность работы алгоритмов не только на каких-то собственных наборах данных (реальных или синтетических, не важно), но и на эталонном общепризнанном тесте, исследователь, в первую очередь, должен этим эталонным тестом располагать. Популярность, а значит и последующее признание в качестве эталонных, тех или иных тестовых наборов зависит, в первую очередь, от их доступности, что предполагает их обязательную публикацию в традиционной, печатной, или же электронной форме. Если для наборов небольших размерностей публикация в бумажном виде не представляет каких-либо трудностей, то для обмена задачами большой размерности более подходящей является электронная форма представления. Наиболее оптимальным вариантом является использование как собственных, так и эталонных тестовых наборов данных.

В случае если источником тестовых данных является реальное производство, и не существует юридических, коммерческих или каких-либо других препятствий для открытой публикации, следует вкратце описать специфические особенности тестового набора, в точности воспроизвести его содержимое и обеспечить свободный доступ, например, посредством размещения материалов в Интернете. Если же используются синтетические данные, то следует описать особенности их генерации, и при необходимости также разместить соответствующие модули в Интернете. Возможные способы такой публикации будут рассмотрены далее. К сожалению, многие авторы и поныне пренебрегают подробным описанием применяемых тестовых данных, ограничиваясь лишь обсуждением полученных результатов. Оценить эффективность того или иного метода, предложенного в рамках подобного исследования, довольно затруднительно.

Развитие Интернета способствовало появлению различных библиотек с тестовыми данными. Наиболее известной из таких библиотек является OR-Library [8, 9], которая обеспечивает свободный доступ к обширной коллекции наборов тестовых данных из различных отраслей операционных исследований, в том числе и задач рационального раскроя и упаковки. Кроме того, следует упомянуть о существующем на сайте ESICUP отдельном разделе [6], предназначенном для обмена тестовыми наборами.

Классифицировав исследуемую в рамках конкретного исследования частную задачу раскроя или упаковки в соответствии с предложенной в [4] типологией, следует подобрать соответствующие ей тестовые данные, которые, скорее всего, будут представлены в виде обычных текстовых файлов. Необходимо создать программное средство или же вручную конвертировать эти тестовые наборы в формат NestingXML, в дальнейшем используя только

файлы одного формата. Подобный подход значительно упрощает последующую публикацию результатов исследования и обмен полученными данными на соответствующих специализированных ресурсах в Интернете.

Обеспечение свободного доступа к результатам исследований. Под публикацией результатов исследования следует понимать не только текстовый и графический анализ в виде материалов, представленных в печатной или электронной форме, но также обеспечение доступа как к используемым тестовым наборам и получаемым решениям, так и к исходным кодам, реализующим применяемый оптимизационный аппарат.

В настоящее время основными препятствиями для свободного распространения тестовых наборов данных в операционных исследованиях являются несоответствие используемых форматов хранения подобных наборов, недооценка важности публикации полученных решений и недостаточные функциональные возможности существующих тематических интернет-ресурсов, предназначенных для централизованного хранения соответствующей информации, что характеризуется значительным ее рассредоточением, усложняя поиск и обмен необходимыми данными.

Для решения первой из этих проблем в исследованиях, связанных с задачами рационального раскроя и упаковки, следует использовать обсуждавшийся ранее формат хранения данных NestingXML. Он позволяет не только унифицировать форму представления и повысить читаемость данных за счет включения в структуру файла стандартных семантических элементов, но также и объединить в одном файле тестовый набор данных с множеством полученных решений. Выбрав формат NestingXML еще на этапе проектирования, исследователь тем самым избавляет себя от необходимости осуществлять дополнительную конвертацию данных при их последующей публикации.

Существующие в Интернете ресурсы, предназначенные для хранения и обмена тестовыми данными, плохо соответствуют современной методологии сетевого взаимодействия, которая реализует принцип привлечения пользователей к наполнению и многократной верификации данных. Например, упомянутый ранее ресурс OR-Library обновляется только одним человеком — его автором, что значительно замедляет публикацию новых тестовых наборов и исправление возможных ошибок. Добавление тестовых наборов в базу данных ESICUP осуществляется в соответствие со схожим принципом: данные сначала отправляются по электронной почте и только после проверки вручную размещаются на сайте. Применение подобной схемы негативно отражается на актуальности хранимой информации.

Значительный успех проектов, подобных Wikipedia [10], доказал эффективность использования коллективного подхода, когда пользователи получают возможность добавления и модификации информации, а контроль за ее достоверностью возлагается на сообщество активных участников проекта. Для реализации этих идей применительно к задачам исследования операций были созданы ресурсы COmputational INfrastructure for Operations Research (COIN-OR) [11] и libOR (по аналогии с OR-Library) [12].

Проект COIN-OR основан в 2003 году при поддержке корпорации IBM и предполагает разработку программного обеспечения с открытыми исходными кодами, распространяемого на условиях лицензии Common Public License (CPL) [11]. В рамках этого проекта уже созданы программные реализации некоторых эвристических и метаэвристических методов, применяемых в операционных исследованиях. Любой заинтересованный исследователь может воспользоваться этими исходными кодами или принять непосредственное участие в работе проекта, присоединившись к процессу разработки уже имеющихся или созданию новых решателей. Реализуя тот или иной метод в рамках собственного исследования, имеет смысл предусмотреть структуру программного обеспечения таким образом, чтобы выделить вычислительную часть в единый логически

завершенный модуль, исходные коды которого впоследствии могут дать начало отдельному проекту в рамках COIN-OR.

В отличие от предыдущего проекта libOR представляет собой тематический портал, предлагающий расширенные возможности по хранению и обмену данными [12]. Логически структура хранимой информации разделена на три различных уровня: формальные описания задач, тестовые наборы и соответствующие им решения. Пользователи могут независимо друг от друга добавлять информацию, когда один исследователь загружает формальное описание задачи, другой — тестовые наборы данных для этой же задачи, третий — полученные им решения. Данные предлагается хранить в различных форматах, разрабатываемых на основе XML и учитывающих характерные особенности каждого класса задач. Следует отметить, что формат NestingXML не соответствует трехуровневой структуре libOR, так как предполагает совместное хранение тестовых наборов и полученных решений в одном файле. Решение разработчиков NestingXML не отделять тестовые наборы от получаемых решений выглядит более удачным, так как исследователь, располагающий данными в таком формате, всегда сможет установить соответствие между тестовыми наборами и решениями. Возможно, авторам libOR следует перейти к двухуровневой структуре, объединив тестовые наборы и соответствующие им решения на одном уровне. Управление порталом и верификацию данных по аналогии с Wikipedia предполагается возложить на группу модераторов, выбираемых из числа активных пользователей. В настоящее время существует прототип портала, который реализует базовую функциональность и действует в тестовом режиме.

Возможно, в скором времени появятся специализированные интернет-ресурсы нового поколения, позволяющие не только хранить исходные коды, тестовые наборы и решения, но также размещать презентационные материалы, вести тематический блог и т.д. Альтернативой использования подобных порталов является создание собственного сайта для размещения всех необходимых данных, рабочих материалов и сопутствующей документации. Из недостатков, присущих этому варианту, следует упомянуть необходимость наличия определенных знаний для создания и поддержания работоспособности сайта, трудозатраты, напрямую не связанные с исследовательской деятельностью, а также нежелательное рассредоточение данных в Интернете, затрудняющее поиск и обмен научной информацией.

Выводы. Значительный интерес, проявляемый учеными к задачам рационального раскрытия и упаковки на протяжении последних десятилетий, не только способствовал разработке новых методов их решения, но также выявил некоторые очевидные проблемы, связанные с обменом информацией и обеспечением преемственности в научных исследованиях. Потенциал электронных средств коммуникации, в том числе и Интернета, применительно к задачам рационального раскрытия и упаковки не реализован в полной мере.

В статье предложена типовая структура информационного обеспечения, которая предполагает реализацию в рамках каждого исследования следующих мер:

1. Для хранения данных следует использовать стандартный формат NestingXML, расширив его в случае необходимости собственными синтаксическими элементами.

2. При оценке эффективности программной реализации того или иного метода решения задачи рекомендуется использовать как собственные, так и эталонные тестовые наборы данных.

3. Если эталонные тестовые наборы представлены в нестандартном формате, их необходимо конвертировать в формат NestingXML, тем самым способствуя дальнейшей популяризации стандартного формата.

4. Публикация результатов исследования не должна ограничиваться текстовым и графическим анализом в виде материалов, представленных в печатной или электронной форме. Следует обеспечить свободный доступ не только к используемым тестовым наборам данных и полученным решениям, но и к исходным кодам созданных программных средств.

Реализация этого короткого перечня практически осуществимых мер позволит значительно активизировать информационный обмен между заинтересованными учеными, что, в свою очередь, положительно отразится на научной новизне и общей продуктивности новых исследований, связанных с решением задач рационального раскроя и упаковки.

Литература

1. Канторович Л.В. Математические методы в организации и планировании производства / Канторович Л.В. — Л.: Изд-во ЛГУ, 1939.
2. G. Desaulniers. Column generation / G. Desaulniers, J. Desrosiers, M. M. Solomon. — New-York: Springer Science+Business Media, Inc., 2005. — 358 p.
3. H. Dyckhoff. A typology of cutting and packing problems / H. Dyckhoff // European Journal of Operational Research. — 1990. — Vol. 44, No. 2. — PP. 145–159.
4. G. Wäscher. An improved typology of cutting and packing problems / G. Wäscher, H. Haußner, H. Schumann // European Journal of Operational Research. — 2007. — Vol. 183, No. 3. — PP. 1109–1130.
5. R. Martins. Solving Combinatorial Problems : An XML-based Software Development Infrastructure / R. Martins, M. Carravilla, C. Ribeiro // Technical report, XATA. — 2006.
6. EURO Special Interest Group on Cutting and Packing [Электронный ресурс]. — Режим доступа к странице : <http://paginas.fe.up.pt/~esicup/>.
7. E. Hopper. Two-dimensional packing utilising evolutionary algorithms and other meta-heuristic methods : Ph. D-Thesis / E. Hopper ; Cardiff University, UK. — 2000.
8. J.E. Beasley. OR-Library : distributing test problems by electronic mail / J.E. Beasley // Journal of the Operational Research Society. — 1990. — Vol. 41, No. 11. — PP. 1069–1072.
9. OR-Library [Электронный ресурс]. — Режим доступа к странице: <http://people.brunel.ac.uk/~mastjib/jeb/info.html>.
10. Wikipedia [Электронный ресурс]. — Режим доступа к странице: <http://www.wikipedia.org/>.
11. R. Lougee-Heimer. The Common Optimization INterface for Operations Research : Promoting open-source software in the operations research community / R. Lougee-Heimer // IBM J. Res. & Dev. — 2003. — Vol. 47, No. 1. — PP. 57–66.
12. K. Ven. Stimulating information sharing, collaboration and learning in operations research with libOR / K. Ven, K. Sörensen, J. Verelst, M. Sevaux // International Journal on Digital Libraries. — 2008. — Vol. 8, No. 2. — PP. 79–90.

Рецензент

зав. каф. РТЗИ ДонНТУ,
д.т.н, проф. Чичикало Н.И.