

На правах рукописи



004617725

A handwritten signature in black ink, consisting of several overlapping, stylized loops and lines, positioned to the right of the barcode and its number.

Янюшкин Вадим Вадимович

**МАТЕМАТИЧЕСКИЕ МОДЕЛИ ОПТИМИЗАЦИИ
РАСПРЕДЕЛЁННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ
ТРЕНАЖЁРНО-МОДЕЛИРУЮЩИХ КОМПЛЕКСОВ**

Специальность 05.13.18 – Математическое моделирование, численные
методы и комплексы программ

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Новочеркасск – 2010

Работа выполнена на кафедре «Автоматизированные системы управления» ГОУ ВПО «Южно-Российский государственный технический университет» (Новочеркасский политехнический институт)» и ООО «Центр тренажеростроения и подготовки персонала» (г. Москва)

Научный руководитель	доктор технических наук, профессор Михайлов Анатолий Александрович
Официальные оппоненты	доктор технических наук, профессор Курейчик Виктор Михайлович
	кандидат технических наук, Захаров Владимир Львович
Ведущая организация	Закрытое акционерное общество «Научно-исследовательский институт «Центрпрограммсистем» (г. Тверь)

Защита диссертации состоится «22» октября 2010 г. в 10 часов на заседании диссертационного совета Д 212.304.02 при ГОУ ВПО «Южно-Российский государственный технический университет (Новочеркасский политехнический институт)» по адресу: 346428, г. Новочеркасск, Ростовской обл., ул. Просвещения, 132, (гл. корпус, ауд. 107)

С диссертацией можно ознакомиться в научно-технической библиотеке «Южно-Российского государственного технического университета (Новочеркасского политехнического института)». С текстом автореферата можно ознакомиться на сайте ЮРГТУ (НПИ) www.npi-tu.ru

Автореферат разослан «7»09. 2010г.

Ученый секретарь диссертационного совета
канд. тех. наук, профессор



Иванченко А.Н.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Разработка тренажерно-моделирующих комплексов (ТМК) различного назначения подразумевает проектирование архитектуры вычислительной системы, которая строится на базе набора вычислительных узлов – персональных компьютеров и высокопроизводительных серверов. Особое место занимает разработка набора взаимосвязанных моделей, позволяющих воспроизводить или имитировать те или иные процессы и явления в тренажере. Каждая модель является потребителем некоторого набора исходных данных на входе и источником набора данных на выходе. Совокупность наборов входных и выходных данных моделей называют модельным миром. Это понятие означает множество данных, циркулирующих в системе между различными моделями. В системах реального времени принято различать понятия «жесткого» и «мягкого» реального времени, при этом данные модельного мира, потребляемые различными моделями в процессе своего функционирования, также должны отвечать условиям системы реального времени. При проектировании ТМК, как правило, жестко связывают архитектуру системы и специальное программное обеспечение, помещая на вычислительных узлах набор моделей, функционирующих в соответствии со своей логикой. Создание различных тренажеров, в том числе систем комплексирования нескольких автономных тренажеров, значительно усложняет характер информационных потоков между отдельными моделями. Важной является задача размещения исходных данных модельного мира на всех узлах ТМК в соответствии с их входными потребностями, а также задача нахождения такого варианта размещения, при котором будет обеспечена максимальная производительность ТМК при ограничениях, накладываемых на ресурсы вычислительной системы.

Наиболее близкой задачей является проблема проектирования распределенных баз данных и нахождение методик оптимальной репликации, но эти задачи не учитывают многие аспекты специфики построения тренажеров. Общие подходы к оценке ресурсов в тренажерных системах отражены в работах Шукшунова В.Е., Потоцкого С.И., Кобзева В.В., Шилова К.Ю. В работах Крестьянинова В.Б. задача повышения производительности тренажерной системы сводится к задаче рационального распределения компонентов специального программного обеспечения.

При проектировании ТМК задачу размещения данных в распределенной системе решают двумя способами: интуитивное распределение исходных данных и создание множества дополнительных каналов связи для доставки и изменения составных частей модельного мира вычислительным узлам ТМК; использование специализированных стандартов распределённого моделирования, таких как High Level Architecture (HLA). Недостаток первого подхода заключается в потенциальной возможности размещения данных, которое может привести к снижению производительности тренажера, а также обязательном привлечении экспертов и специалистов на этапах проектирования системы. К недостаткам второго подхода можно отнести сложность внедрения и использования стандарта HLA, перевода структуры системы в соответствие специфики разработок тренажеров высокоуровневой архитектуры.

Проектирование, разработка и введение в строй в настоящее время новых образцов подводных лодок, вооружения, средств обнаружения и разработка соответствующих методик проведения подготовки, где одной из основных задач повышения

уровня боевой подготовки является интенсификация тренировок по специальности всех категорий личного состава кораблей и частей флота с использованием морских тренажерных комплексов, является приоритетным направлением развития области тренажеростроения. Введение уникальной функциональности на основе математических методов теории принятия решений повышает эффективность и конкурентоспособность программных продуктов на развивающемся рынке данных услуг.

Таким образом, актуальность работы складывается из следующего:

1. Необходимости проектирования и создания программно-аппаратных тренажерных комплексов, которые требуют поддержания необходимой функциональности размещенных на узлах вычислительной сети моделей.

2. Отсутствия математических постановок и методов, обеспечивающих размещение модельного мира в системе в условиях жесткой привязки моделей к вычислительным узлам тренажера, которые бы позволили минимизировать временные характеристики при получении необходимых входных данных, а также сократить совокупный объем хранения информационных массивов в распределенной информационной системе.

3. Необходимости применения современных подходов проектирования распределенных информационных систем, концепции сервисно-ориентированной архитектуры (SOA) и методов построения аппаратно-технической платформы.

Диссертационная работа выполнена в рамках: комплексной целевой программы по развитию средств обучения и подготовки Вооруженных Сил Российской Федерации на период до 2020 г., где одним из важных факторов реализации является создание единой системы учебно-тренировочных средств и комплексных тренажеров подготовки; научного направления ЮРГТУ (НПИ) «Интеллектуальные тренажерно-обучающие комплексы, тренажеры, системы виртуальной реальности, виртуальные лаборатории – основа инновационных образовательных программ в технических университетах»; госбюджетной темы 7.05 «Разработка теории, методов оптимальной функциональности и программно-технической платформы корпоративных информационных систем» (утверждено решениями ученого совета от 25.04.2001 и 15.05.2003).

Целью диссертационной работы является построение математических моделей оптимизации структуры информационного обеспечения распределенной системы тренажерно-моделирующих комплексов на этапах проектирования и функционирования за счет оптимизации размещения данных, которая позволит увеличить скорость доступа в рамках функционирования моделей тренажера, а также снизить совокупный объем хранимой информации в системе путем рационального распределения данных модельного мира.

Для достижения этой цели решаются следующие задачи: анализ существующих архитектурных решений и математических моделей оптимизации распределенных систем; формализация и построение математической модели оптимизации размещения модельного мира, учитывающей специфику этапов функционирования ТМК; анализ современных концепций проектирования распределенных информационных систем и применение их при разработке ТМК; разработка моделей размещения данных в информационной системе; построение программного комплекса размещения модельного мира и проведение экспериментального исследования эффективности предложенных алгоритмов и моделей.

Методы исследований и достоверность результатов. В работе использованы методы теории принятия решений, имитационного моделирования, а также теории вероятностей и генетических алгоритмов (ГА). Достоверность результатов подтверждается корректным применением элементов теории принятия решений, планирования экспериментов, сопоставлением полученных экспериментальных результатов с имитационным моделированием, непротиворечивостью предложенных математических моделей и методов поиска решения, а также положительной оценкой внедрения результатов в разрабатываемые ТМК.

Объектом исследования являются современные архитектурные решения ТМК, принципы их построения, концепция центров обработки данных, облачных вычислений и многоуровневых информационных систем.

Предметом исследования являются наборы данных, циркулирующие в ТМК, математические модели их описывающие, специфика функционирования процессов обучения и подготовки персонала в современных реализациях морских ТМК, а также влияние этих факторов на особенности размещения и использования данных.

Научная новизна. В диссертации получены следующие новые научные и практические результаты: формализованная математическая модель представления распределенного информационного пространства, учитывающая иерархическое представление системы и взаимосвязи между вычислительными узлами, моделями и наборами данных; математические модели оптимизации размещения данных на основе архитектуры одноуровневой локальной вычислительной сети (стандартная схема информационной системы ТМК), в представлении облачных вычислений и на основе кластеров, которые позволяют, учитывая специфику архитектуры построения системы, проводить процедуру оценки затрат на поддержку модельного мира и оптимизации характеристик за счет минимизации занимаемого объема памяти; методы поиска решений на основе применения ГА, позволяющие учитывать особенности функционирования и ограничения систем реального времени; программные компоненты уровня информационного обеспечения распределенной информационной системы на основе SOA и реализующей данный подход технологии WCF (Windows Communication Foundation).

Основные положения, выносимые на защиту:

1. Постановка задачи оптимизации размещения модельного мира в распределенной системе ТМК, этапы ее решения. Новизна первого научного результата заключается в представлении распределенного информационного пространства, учитывающего иерархическое построение системы и взаимосвязи между вычислительными узлами, моделями и наборами данных.

2. Комплекс моделей на основе применения различных стратегий размещения данных, ГА и модифицированных постановок задач. Новизна второго научного результата состоит в совершенствовании методов размещения, которые позволяют учитывать специфику архитектуры построения системы проводить процедуру оценки затрат на поддержку модельного мира и оптимизации за счет минимизации занимаемого объема памяти.

3. Результаты моделирования и экспериментального исследования эффективности предложенных алгоритмов. Новизна третьего научного результата состоит в анализе эффективности применения ГА, оценках производительности тренажера и результатов моделирования размещения данных.

4. Методика использования разработанных моделей в тренажёрах, концепция применения современных технологий построения распределенных информационных систем на основе сервисно-ориентированного подхода. Новизна четвертого научного результата состоит в реализации разработанной методики повышения производительности тренажера в составе программных компонент уровня информационного обеспечения распределенного информационного пространства.

Теоретическая ценность работы заключается в построении и исследовании концептуальных моделей размещения объектов данных в информационных системах, конструировании алгоритмов и разработке численных методов с учетом ограничений систем реального времени.

Практическая ценность работы заключается в реализации и использовании разработанных моделей и алгоритмов на этапах проектирования и эксплуатации современных тренажеров, а также для описания, оценки и анализа процессов взаимодействия вычислительных моделей и распределенных данных. Представленные алгоритмы позволяют оценить и улучшить с использованием оптимизации размещения элементов модельного мира производительность системы и как следствие ее эффективность в целом.

Реализация результатов работы. Результаты диссертационной работы использованы при разработке и проектировании аппаратно-программных средств комплексирования «Листва-К», используемых в составе базового комплекса учебно-тренировочных средств «Листва» (г. Обнинск), а также комплексного тренажера «Калина-О» (г. Северодвинск). Использование полученных результатов при проектировании структуры систем позволило сократить затраты на дальнейшее в комплексные тренажеры с распределенным информационным пространством, сократить затраты памяти на хранение модельного мира в среднем на 36 % и получить время доступа к отдельным объектам в пределах 0,3-0,6 мс за счет рационального размещения и использования SOA при построении информационных систем. Задачи оптимизации размещения модельного мира в ТМК, разработанные математические модели и научные результаты работы также внедрены в учебный процесс ЮРГТУ (НПИ).

Апробация работы. Основные положения диссертации и отдельные ее результаты обсуждались и получили положительные отзывы на:

научно-технической конференции студентов и аспирантов ЮРГТУ (НПИ) «Студенческая весна 2007» (г. Новочеркасск);

ежегодных научно-технических конференциях ЮРГТУ (НПИ) «Теория, методы проектирования, программно-техническая платформа корпоративных информационных систем» в период с 2007-2009 гг. (г. Новочеркасск);

VII Международной научно-практической конференции «Моделирование. Теория, методы и средства», 2007 г. (г. Новочеркасск);

Всероссийском смотре-конкурсе научно-технического творчества студентов высших учебных заведений «Эврика-2007» (г. Новочеркасск);

седьмой международной научно-практической конференции «Пилотируемые полеты в космос», 2007 г. (Звездный городок);

межвузовской научно-технической конференции «Перспективы развития средств и комплексов связи. Подготовка специалистов связи», 2009 г. (г. Новочеркасск).

В полном объеме диссертационная работа докладывалась и обсуждалась в ООО «Центр тренажеростроения и подготовки персонала» (г. Москва), ЗАО НИИ «Центрпрограммсистем» (г. Тверь). Получено 2 акта внедрения в комплексные ТМК и акт внедрения научных результатов в учебный процесс ЮРГТУ (НПИ).

Публикации. По результатам выполненных исследований опубликовано 18 научных работ, из них 5 в рекомендованных ВАК изданиях, получено свидетельство о регистрации электронного ресурса.

Структура диссертации. Диссертация содержит 210 страниц основного текста, 88 рисунков, 3 таблицы и состоит из введения, четырех глав, заключения, списка литературы из 143 наименований и трех приложений объемом 40 страниц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертации, научная новизна и практическая значимость, сформулированы цель и задачи работы. Представлены положения, выносимые на защиту, апробация работы и структура диссертации.

В первой главе «Обоснование необходимости решения проблемы оптимизации размещения модельного мира в распределенных информационных системах тренажерно-моделирующих комплексов» приводится обзор существующих архитектур распределенных информационных систем современных ТМК на основе использования локальных вычислительных сетей, комплексирования, элементов технологии SOA. Обобщая различные архитектуры и способы проектирования ТМК, выделяется понятие единого информационного пространства, концептуально являющегося средой объединения входных и выходных информационных потоков всех составных частей ТМК. Рассматривается стандарт IEEE 1516 и высокоуровневая архитектура HLA, а также тренажерная распределенная исполнительная оболочка.

Приведены описанные в литературе подходы к построению математических моделей проектирования и оптимизации структур распределенных информационных систем, в основе которых лежат методики проектирования оптимальных структур распределенных баз данных. Исходя из анализа этапов функционирования и структуры ТМК приводится формализация проблемы размещения данных модельного мира, где в качестве исходных данных задачи служит следующая информация: n – число элементов аппаратной конфигурации PC_i , $i = \overline{1, n}$ с заданными допустимыми объемами памяти на размещение данных V_{0i} ; p – число элементов функциональной конфигурации M_j , $j = \overline{1, p}$; m – число типов данных уровня информационного обеспечения D_k , $k = \overline{1, m}$; матрицы размещения моделей $MC_{i,j}$, $i = \overline{1, n}$, $j = \overline{1, p}$ и потребностей данных $MD_{j,k}$, $j = \overline{1, p}$, $k = \overline{1, m}$; $C_{i,k}$, $i = \overline{1, n}$, $k = \overline{1, m}$ набор матриц размещения объектов модельного мира; $T_{j,k}$, $j = \overline{1, p}$, $k = \overline{1, m}$ – матрица времен функционирования модели j и получения всех необходимых входных параметров от объектов модельного мира по типам $k = \overline{1, m}$ для их обработки.

Логическим продолжением развития архитектурных решений ТМК является использование облачных вычислений (cloud computing) и их модификаций (private clouds), предлагается концептуальное соответствие и предложения по внедрению

данных технологий в распределенные информационные системы ТМК. Архитектура ТМК и информационной системы при использовании концепции private clouds является следующим шагом систем комплексирования, применения SOA и сервисных приложений, а также логическим продолжением и аппаратно-программным решением для построения единого информационного пространства.

Во второй главе «**Модели и алгоритмы оптимизации размещения данных в распределенных информационных системах**» предлагаются модели, основанные на специфике работы системы и параметрах информационных сущностей. Представлены следующие математические модели оптимизации размещения:

1. Модель размещения данных в стандартной схеме информационной системы ТМК, которая предполагает использование методов размещения данных в проектируемых тренажерах, рассматриваются алгоритмы решения, а также результаты их работы на тестовых примерах.

2. Модель размещения данных в представлении облачных вычислений информационной системы ТМК, которая предполагает использование при проектировании тренажера облака данных и концепций центров обработки данных (ЦОД), рассматривается формализация проблемы, математическая постановка, приводится алгоритм решения.

3. Модель размещения данных в условиях формирования кластеров информационной системы ТМК, в основе которой положено выделение взаимосвязей между отдельными моделями и объектами модельного мира и образование кластерных элементов как основных структур информационного пространства, приведен алгоритм решения задачи и его анализ на тестовых примерах.

При постановке задачи оптимизации размещения данных в стандартной схеме информационной системы (рис. 1) дополнительно вводятся следующие условия: детализованная характеристика объема каждого объекта типа данных $VD_{k,j}$, $k = \overline{1, m}$, $l = \overline{1, D_k}$; матрица интенсивности запросов по типам данных от моделей $MDI_{j,k}$, $j = \overline{1, p}$, $k = \overline{1, m}$; временные ограничения моделей $T_{ршт.}$, $T_{лмд}$, $j = \overline{1, p}$, время получения данных t^{net} , t^{mem} объема V_{net} и V_{mem} соответственно.

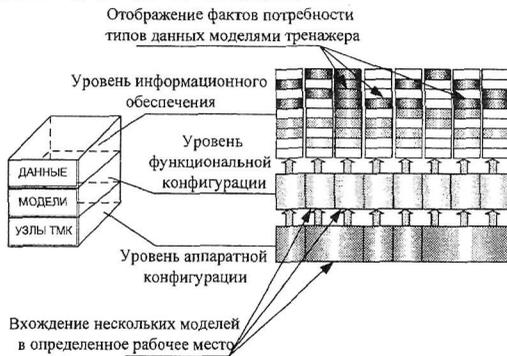


Рис. 1. Иерархия уровней тренажера

Матрица размещения $C_{i,k}$, $i = \overline{1, n}$, $k = \overline{1, z}$, где z – количество объектов модель-

ного мира в соответствии с введенной классификацией по типам: $z = \sum_{k=1}^L D_k$. Для построения модели вводятся следующие обозначения: $M' \subseteq M$ – множество моделей, расположенных на узле $i = \overline{1, n}$; $D' \subseteq D$ – множество типов данных, расположенных на узле $i = \overline{1, n}$; $VD_j^i \subseteq VD_j$, $j = \overline{1, m}$, $i = \overline{1, n}$ – множество объектов модельного мира типа j на узле i . В процессе оптимизации ГА модифицируется матрица $C_{i,k}$ и как следствие множества VD_j^i . Требуется найти такой вариант размещения объектов модельного мира в системе, чтобы обеспечить минимизацию целевой функции:

$$F = \sum_{i=1}^n \left(w_1 * \sum_{k \in D'} \sum_{i \in VD_k^i} VD_{k,j} * C_{i, \sum_{k=1}^m D_k + i} + w_2 * \sum_{j \in M'} \sum_{i=1}^m MDI_{j,i} * MD_{j,i} * \sum_{g \in VD_g^i} VD_{i,g} \right) \rightarrow \min$$

при выполнении ограничений:

$$\sum_{k \in D'} \sum_{i \in VD_k^i} VD_{k,j} \leq V_{0i}, \quad i = \overline{1, n}, \quad k = \overline{1, z}; \quad \sum_{n=1}^m T_{j,n} + T_{j,000} \leq T_{j, \text{критум}}, \quad j = \overline{1, p},$$

$$VD_1^i \cup VD_2^i \cup \dots \cup VD_m^i = VD_i, \quad i = \overline{1, m}$$

Подбор весовых коэффициентов w_1 и w_2 может осуществляться на основе экспертных оценок. Проводится разбиение множества объектов типа данных по моделям, $VD_j^{M_i} \subseteq VD_j$, $j = \overline{1, m}$, $i = \overline{1, p}$ – множество объектов модельного мира типа j , необходимых модели M_i . Таким образом, вводится ограничение на уровне элементов данных определенного типа: $VD_1^{M_1} \cup VD_2^{M_2} \cup \dots \cup VD_m^{M_p} = VD_i$, $i = \overline{1, m}$. Матрица $T_{j,k}$, $j = \overline{1, p}$, $k = \overline{1, m}$ вычисляется следующим образом:

$$T_{j,k} = \begin{cases} t^{mem} * \sum_{\substack{i \in VD_k^{M_i} \\ MC_i, j=1}} \frac{VD_{k,j}}{V_{mem}} + t^{net} * \sum_{\substack{i \in VD_k^{M_i} \\ MC_i, j=0}} \frac{VD_{k,j}}{V_{net}}, & \text{если } MD_{j,k} = 1, \\ 0, & \text{если } MD_{j,k} = 0 \end{cases}$$

Предлагаются стратегии начального размещения данных на основе алгоритма приоритета интенсивности данных (наиболее интенсивно используемые типы данных модельного мира в соответствии с матрицей MDI являются основой межмодельного трафика запросов и основной составляющей критерия эффективности на дополнительные затраты и динамическое обновление данных), минимизации временных затрат (помещение всех необходимых объектов модельного мира на вычислительные узлы исходя из предположения, что основным параметром оптимизации являются временные характеристики), избыточности данных на основе комбинирования типов данных (помещение на каждый вычислительный узел полной копии фрагмента модельного мира по критерию типа потребности моделей).

Постановка задачи оптимизации размещения данных в представлении облачных вычислений информационной системы ТМК (рис. 2 – слева) включает в себя nsr – число ресурсных серверов облака данных $SR = (SR_1, \dots, SR_{nsr})$ с заданными объемами отведенной для хранения объектов модельного мира памяти $VS = (VS_1, \dots, VS_{nsr})$, nsp – число серверов приложений $SP = (SP_1, \dots, SP_{nsp})$ с объемами $VS = (VS_{nsr+1}, \dots, VS_{nsr+nsp})$. Размещение модельного мира в облаке данных допускается

на SR и SP серверах, а также на системе хранения данных (Data Storage), при этом функционирование большинства моделей M_j , $j = \overline{1, p}$ предполагается на SP серверах, фактическое нахождение объекта данных $VD_{k,l}$, $k = \overline{1, m}$, $l = \overline{1, D_k}$ может описываться как интерпретация задачи распределения файлов по уровням памяти вычислительной системы, где при условии распределения объекта $VD_{k,l}$ на нижний уровень (Data Storage) время получения информации составит t^{DS} объема V_{DS} , на средний уровень – SR сервер – время составит t^R для объема V_R , на верхний уровень – SP сервер – t^P и V_P .

На уровне архитектуры облака данных максимальное время получения данных модели составит: $\max(t^{DS}, t^R, t^P)$. nsi – число серверов предоставления информации облака данных $SI = (SI_1, \dots, SI_{nsi})$ с заданными объемами отведенной для хранения объектов модельного мира памяти $VSI = (VSI_1, \dots, VSI_{nsi})$, npc – число элементов корпоративных и удаленных клиентов облака данных $PC = (PC_1, \dots, PC_{npc})$ с объемами памяти $VPC = (VPC_1, \dots, VPC_{npc})$.

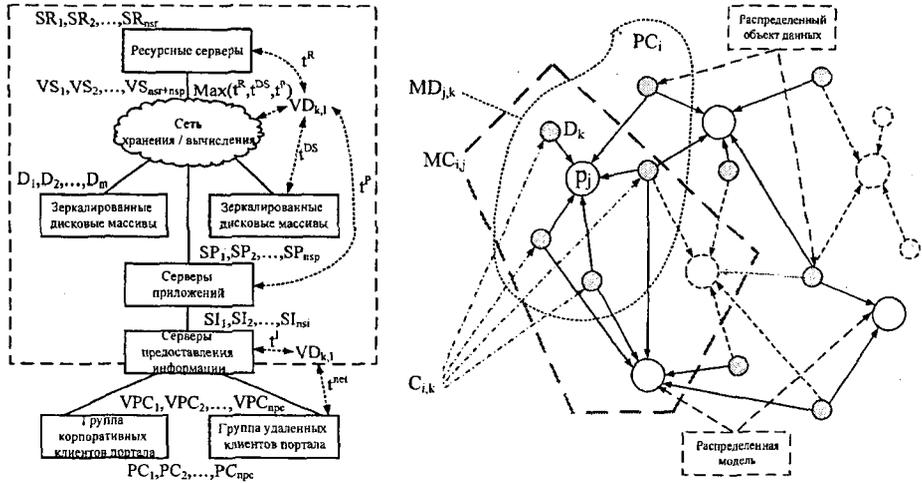


Рис. 2. Задача на основе облака данных и информационного кластера

Для функций предоставления информации размещение объектов модельного мира допускается на SI серверах и удаленных клиентах, время получения информации с SI сервера составит t^I объема V_I , время получения данных клиентом через сетевую инфраструктуру составляет t^{net} . На уровне облака данных и клиентов максимальное время получения данных составит: $\max(t^{DS}, t^R, t^P, t^I) + t^{net}$ при заданном допустимом ограничении t^I . Задается матрица размещения моделей по серверам приложений MS_{ij} , $i = \overline{1, nsp}$, $j = \overline{1, p}$, матрица MD_{jk} , $j = \overline{1, p}$, $k = \overline{1, m}$, MDI_{jk} , $j = \overline{1, p}$, $k = \overline{1, m}$; временные ограничения моделей. Для построения модели вводятся следующие множества: множество моделей, расположенных на серверах приложений

$M' \subseteq M, i = \overline{1, nsp}$; ресурсных серверах $VD_j^R \subseteq VD_j, i = \overline{1, nsr}, j = \overline{1, m}$; серверах приложений $VD_j^A \subseteq VD_j, i = \overline{1, nsp}, j = \overline{1, m}$; серверах предоставления информации $VD_j^I \subseteq VD_j, i = \overline{1, nsi}, j = \overline{1, m}$; клиентах облака данных $VD_j^{PC} \subseteq VD_j, i = \overline{1, npc}, j = \overline{1, m}$.

Матрица $T_{j,k}$ рассчитывается по формуле:

$$T_{j,k} = \begin{cases} t^P * \sum_{k=1}^m \sum_{l \in VD_k^A} \frac{VD_{k,l}}{V_P} + t^R * \sum_{k=1}^m \sum_{l \in VD_k^R} \frac{VD_{k,l}}{V_R} + t^{DS} * \sum_{k=1}^m \sum_{l \in VD_k^I, l \in VD_k^{PC}} \frac{VD_{k,l}}{V_{DS}}, & \text{если } MD_{j,k} = 1, \\ 0, & \text{если } MD_{j,k} = 0 \end{cases}$$

Смысл вычисляемого значения $T_{j,k}$ по отношению к модели j состоит в суммировании объемов элементов модельного мира типа k , расположенных в зависимости от принятой стратегии размещения, тогда в зависимости от размещения на SP , SR – сервере или Data Storage учитывается свое время получения данных. Целевая функция состоит из двух частей: затраты на поддержку серверного комплекса F^S и удаленных клиентов F^{CI} :

$$F^S = \sum_{i=1}^{nsp+nsr} (w_1 * \sum_{k=1}^m \sum_{l \in VD_k^I} VD_{k,l} * C_{i, \sum_{s=1}^{D_s+1}} + w_1 * \sum_{k=1}^m \sum_{l \in VD_k^R} VD_{k,l} * C_{i, \sum_{s=1}^{D_s+1}} + w_2 * \sum_{j=1}^m \sum_{l=1}^m MDI_{j,l} * MD_{j,l} * \sum_{g=1}^{D_j} VD_{l,g}) \rightarrow \min$$

Минимум целевой функции находится при выполнении ограничений на соблюдение отведенного максимального объема для хранения данных на серверах облака данных:

$$\sum_{k=1}^m \sum_{l \in VD_k^A} VD_{k,l} \leq VS_{j,i}, i = \overline{1, nsp}, \sum_{k=1}^m \sum_{l \in VD_k^R} VD_{k,l} \leq VS_{j,i}, i = \overline{1 + nsr, nsp + nsp},$$

предельное время выполнение временных операций на каждом SP сервере:

$$\sum_{h=1}^m T_{j,h} + T_{j,уд} \leq T_{j,спm}, j = \overline{1, p}.$$

Для удаленных клиентов необходимо обеспечить минимизацию функции:

$$F^{CI} = \sum_{i=1}^{npc+nsi} (\sum_{k=1}^m \sum_{l \in VD_k^I} VD_{k,l} * C_{i, \sum_{s=1}^{D_s+1}} + \sum_{k=1}^m \sum_{l \in VD_k^A} VD_{k,l} * C_{i, \sum_{s=1}^{D_s+1}}) \rightarrow \min,$$

Минимум целевой функций находится при выполнении ограничений на максимальный объем для хранения данных на SI серверах и удаленных клиентах:

$$\sum_{k=1}^m \sum_{l \in VD_k^I} VD_{k,l} \leq VSI_{j,i}, i = \overline{1, nsi}, \sum_{k=1}^m \sum_{l \in VD_k^A} VD_{k,l} \leq VPC_{j,i}, i = \overline{1, npc}, t^{nci} + A^{CI} \leq t^{CI},$$

A^{CI} – составляющая на получение модельного мира из облака данных:

$$A^{CI} = \sum_{i=1}^{npc+nsi} \left(t^I * \sum_{k=1}^m \sum_{l \in VD_k^I, l \in VD_k^A} \frac{VD_{k,l}}{V_I} * C_{i, \sum_{s=1}^{D_s+1}} + t^{DS} * \sum_{k=1}^m \sum_{l \in VD_k^I, l \in VD_k^A} \frac{VD_{k,l}}{V_{DS}} * C_{i, \sum_{s=1}^{D_s+1}} \right).$$

Предлагаются стратегии начального размещения данных на основе 100% использования Data Storage (все объекты модельного мира помещены в систему хранения облака данных), алгоритма распределения данных между SR и SP серверами (стратегия распределения объектов данных с высокой степенью интенсивности на SP серверах, а с низкой на SR) и минимизации временных затрат между SP сервера-

ми (помещение всех необходимых объектов данных на SP серверах).

Особенностью задачи оптимизации размещения данных в условиях формирования кластеров информационной системы ТМК (рис. 2 – справа) является способ представления данных и формирования вариантов размещения объектов модельного мира. Вводится понятие распределенной модели и экземпляра данных, кластер строится в виде связей моделей через типы данных, образуя логические цепочки соответствующей функциональности. В данной постановке задачи к алгоритмам начального размещения можно отнести алгоритм конкурирующих моделей данных (выделение набора моделей, которым требуется одинаковый объект, в этом случае объект присоединяется к кластеру, где выполняется условие минимальных затрат размещения по критерию объема), алгоритм разделения данных между кластерами (создание в каждом кластере копий объектов модельного мира), алгоритм агрегации кластеров на основе интенсивности использования данных (интенсивно используемые данные помещаются в рамках каждого кластера, остальные размещаются на основе предыдущих стратегий).

В связи с очевидным увеличением количества переборных вариантов решения при увеличении числа объектов задача размещения модельного мира относится к трудноразрешимым задачам и в качестве основного метода поиска решения предлагается использование ГА. Модификация стандартного ГА заключается во введении нескольких генетических операторов. Это сделано для повышения эффективности применения операторов и включает в себя следующие особенности:

1. Начальная популяция хромосом получается не из случайного набора решений, а вследствие выбранной стратегии начального размещения данных – стратегии включают в себя алгоритмы приближенной оценки занимаемого объема модельным миром в системе и распределения данных, которое оптимизируется на основе ГА.

2. Кроссинговер (скрещивание) двух хромосом на основе обмена наборов объектов модельного мира между вычислительными узлами, а также на основе обмена состава объектов модельного мира на вычислительном узле в различных вариантах решения (рис. 3). Данные операторы отличаются от классических – одноточечного или двухточечного скрещивания тем, что определенным образом выбираются точки разрыва в хромосомах при выбранном методе кодирования, это позволяет получать новые допустимые варианты размещения модельного мира.

3. Мутация хромосомы на уровне объекта модельного мира конфигурации системы означает удаление/добавление в стратегию размещения объекта модельного мира, точка мутации определяется из возможных вариантов, а при мутации на уровне вычислительных узлов производится обмен составом данных между двумя wybranными узлами тренажера.

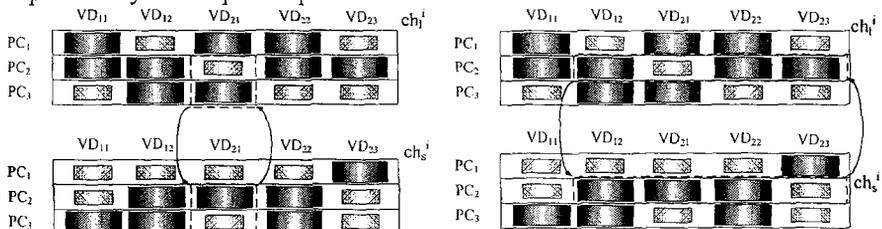


Рис. 3. Различные операторы кроссинговера двух хромосом ГА

В третьей главе «Численные методы и стратегии решения задачи оптимизации размещения модельного мира в информационных системах тренажёрно-моделирующих комплексов» предлагается обобщенный алгоритм оптимизации распределенной информационной системы ТМК на основе анализа моделей и алгоритмов распределения данных. В части аппаратной конфигурации выделяется иерархия уровней размещения данных, где каждая архитектуры построения системы располагает своим набором моделей. Потребности моделей определяются множествами элементов модельного мира различных типов, а алгоритм оптимизация размещения данных основывается на эффективной начальной стратегии размещения, которая предлагает приближенный способ, корректирующийся ГА.

В диссертации производится сравнение производительности различных комбинаций генетических операторов (кроссинговера и мутации), определено время работы ГА и показана зависимость времени решения задачи от размерности исходных данных. Вопросы сходимости ГА рассматриваются на зависимостях коэффициента разнообразия в генах скрещиваемых хромосом и среднего количества хромосом с высокой оценкой фенотипа. В качестве основного критерия отбора в новую популяцию выбран элитизм. Результаты тестирования, усредненные по нескольким запускам, отображены на рис. 4, где в качестве критерия выбиралось отношение величины минимума задачи размещения данных (полученного методом перебора) и найденного значения с помощью ГА. Из рисунка видно, что наилучшие результаты дает применение операторов кроссинговера места размещения и мутации на основе объектов данных.



Рис. 4. Качество генетических операторов

Использование наилучших комбинаций генетических операторов, выявленных в предыдущем эксперименте, позволяет получать в последующих поколениях ГА особей с наилучшим значением целевой функции и приводит общему сокращению времени решения задачи. Общая оценка времени поиска оптимального решения производилась последовательным запуском ГА на различных конфигурациях ТМК, учитывающих количество элементов модельного мира (рис. 5).

Как видно из рис. 5 при увеличении объектов до 400 (количество мест размещения $n = 20$, функциональных моделей $m = 30$) время поиска оптимального решения составляет 43 с (тесты проводились на вычислительных станциях с процессорами Intel Core2Duo). В реальных ТМК, состояние модельного мира которых может

описываться сотнями объектов, а время на конфигурацию и запуск составляет около 7-8 мин (в соответствии с техническими заданиями), данный показатель является приемлемым и позволяет приводить оптимизацию размещения данных модельного мира на начальных этапах работы системы, а также в условиях переустановки и замены функциональных модулей.



Рис. 5. Скорость алгоритма

Для сравнения с ГА выбран прямой перебор и метод ветвей и границ. Применение прямого перебора становится невозможным при увеличении числа объектов модельного мира, а при решении методом ветвей и границ задача размещения сводится к модифицированной минимаксной задаче о назначениях. Основным полученным выводом в результате проведенных экспериментов является то, что разработанный ГА позволяет найти допустимое или оптимальное с точки зрения ограничений задачи решение, являясь при этом разновидностью методов поиска с элементами случайности. Для сложной распределённой системы ТМК требуется найти за время начальной загрузки системы стратегию размещения данных, повышающую производительность и удовлетворяющую предъявляемым ограничениям, а проблема достижения глобального оптимума отходит на второй план.

В четвертой главе «Реализация комплекса программ распределённой информационной системы и применение моделей оптимизации в современных тренажёрах» рассматривается разработка морских ТМК и применение результатов диссертационной работы для проектирования программных компонент уровня информационного обеспечения, приводится аппаратная и программная составляющие общего информационного пространства, проблемы проектирования единых информационных полигонов. Приведена структура программного обеспечения распределённой системы на примере тактического ТМК, где центральную часть занимает специальное программное обеспечение, реализация которого предлагается с использованием сервисно-ориентированного подхода (рис. 6).

На основе применения новейших концепций организации и предоставления ресурсов ЦОД в распределённых информационных системах предлагается ряд архитектурных и технологических решений по построению информационных систем специального назначения. Предлагается структура портала центра тренажерного обучения персонала как перспективной разработки, объединившей в себе преимущества современных ЦОД, технологии private clouds и особенностей проектирования многоуровневой структуры информационных систем.

В данном случае происходит перемещение модели размещения непосредственно в тренажер в составе специального программного обеспечения, где происхо-

дит снятие необходимых входных характеристик из базы данных конфигурирования тренажера. Такое решение позволит в оперативном режиме контролировать производительность тренажера и в случае необходимости производить фоновые операции размещения.

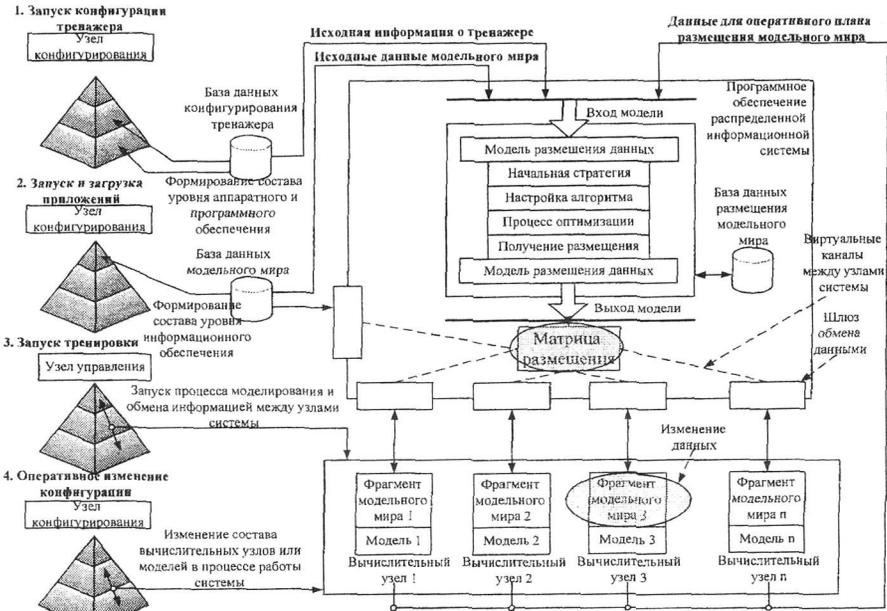


Рис. 6. Взаимодействие программного обеспечения и модели размещения

На примере фрагмента системы рассматривается применение алгоритмов оптимизации распределенного информационного пространства. Объем памяти, затрачиваемый на хранение модельного мира в распределенной системе, удалось снизить в 5 раз по сравнению с первоначальной оценкой и на 36 % по сравнению с наилучшей стратегией начального размещения данных. Время получения сервисами данных составляет 0,3-0,6 мс при использовании программных компонент на основе технологии WCF, что соответствует полученным на практике значениям в ТМК.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ ДИССЕРТАЦИОННОЙ РАБОТЫ

1. Сформулирована задача оптимизации размещения модельного мира в распределенных информационных системах ТМК на основе анализа архитектурных решений проектирования и стандартов распределенного моделирования, отличающихся от известных иерархическим представлением структуры тренажера и взаимосвязей между моделями, данными и вычислительными узлами, что позволяет строить рациональные процедуры и методики проведения распределённых вычислений и удаленного взаимодействия отдельных элементов ТМК.

2. Предложены современные технологические решения проектирования информационных сред на основе ЦОД и облачных вычислений, которые завершаются

концепцией создания универсального портала и перспективных разработок с объединением реальных полигонов проведения учений и ТМК различного назначения, которые отличаются от существующих тем, что позволяют объединять в единое информационное пространство различные средства и методы обучения персонала для повышения качества и эффективности процессов подготовки, а также использовать в данной сфере последние технические и методологические разработки.

3. Сформулирован набор новых математических постановок задач оптимизации размещения данных на различных этапах работы системы и разработаны модели размещения данных на основе стандартной схемы, облачных вычислений и информационных кластеров, отличительной особенностью которых является специфика потоков информации в ТМК и взаимодействия между распределенными данными в рамках моделей, что позволяет в процессе размещения учитывать архитектуру системы, потребности моделей и характеристики объектов модельного мира.

4. Получены алгоритмы оптимизации размещения модельного мира на основе комбинирования подходов начальной стратегии распределения и эволюционных методов ГА, это решение отличает высокая эффективность и скорость получения допустимых вариантов размещения данных, при задании 400 объектов модельного мира около 43 с, что позволяет при процедуре инициализации и загрузки ТМК получать стратегию распределения данных и в дальнейшем при необходимости оперативно изменять отдельные фрагменты информационной системы.

5. Получены результаты анализа производительности и эффективности разработанных алгоритмов, в стандартной схеме лучшие показатели были получены на основе алгоритма минимизации временных затрат в пределах 700 итераций основного ГА, позволяющего распределять объекты на вычислительные узлы в соответствии с их потребностями для уменьшения времени доступа, комбинирование с другими методиками позволило улучшить параметры на 10-20 %. При формировании кластеров наибольшая эффективность получена с конкурирующими моделями данных, которые основаны на присоединении к кластеру, где выполняется условие минимальных затрат размещения по критерию объема для сокращения общего совокупного размера хранения модельного мира в информационной системе.

6. Для сравнения эффективности предложен модифицированный алгоритм задачи о назначениях на основе метода ветвей и границ, отличающийся от существующей стратегией выбора конкурирующих множеств на каждом уровне ветвления, что позволяет тиражировать необходимые объекты модельного мира на различных вычислительных узлах для повышения скорости доступа к данным.

7. Разработаны программные компоненты оптимизации распределенной информационной системы ТМК – Genetic Data Distribution in Simulators (GDDiS), которые отличает разработанная методика оптимизации размещения данных модельного мира и сервисно-ориентированный подход в построении системы на основе технологии WCF, данное решение позволяет внедрять в информационные среды распределенных ТМК логику формирования общетренажерного ресурса данных и повышать производительность системы за счет учета введенных ограничений объема и времени. Проведенные экспериментальные исследования эффективности данного решения показали время получения данных сервисами в пределах 0,3-0,6 мс.

8. Результаты диссертационной работы внедрены в аппаратно-программные средства комплексирования «Листва-К», используемых в составе базового комплек-

са учебно-тренировочных средств «Листва», а также комплексный тренажер «Калина-О», опубликовано 18 печатных работ, в том числе 5 в рекомендованных ВАК изданиях, получено свидетельство о регистрации электронного ресурса.

СПИСОК ОСНОВНЫХ ПУБЛИКАЦИЙ

Публикации в ведущих изданиях, рекомендованных ВАК

1. Янюшкин, В.В. Распределенное информационное пространство и портал современного тренажера / В.В. Янюшкин // Программные продукты и системы. – 2009. – № 3. – С. 67-71.
 2. Янюшкин, В.В. Оптимизация размещения данных модельного мира в распределенной информационной системе тренажерно-моделирующего комплекса / В.В. Янюшкин // Изв. вузов. Северо-Кавк. регион. Техн. науки. – 2008. – № 4. – С. 25-28.
 3. Янюшкин, В.В. Обобщенная математическая постановка задачи оптимизации размещения данных в информационной системе тренажерно-моделирующего комплекса / В.В. Янюшкин // Вопросы современной науки и практики. Университет им. В.И. Вернадского. – 2009. – № 12. – С. 114-121.
 4. Янюшкин, В.В. Кластеры распределенной системы тренажерно-моделирующего комплекса в задаче агрегации фракталов / В.В. Янюшкин // Программные продукты и системы. – 2010. – № 1. – С. 46-54.
 5. Янюшкин, В.В. Модели и алгоритмы оптимизации размещения данных в распределенной информационной системе / В.В. Янюшкин // Изв. вузов. Северо-Кавк. регион. Техн. науки. – 2010. – № 2. – С. 10-16.
- Публикации в сборниках научных статей, трудов и материалов конференций**
6. Янюшкин, В.В. Концептуальный анализ распределенной информационной системы тренажера / А.А. Михайлов, А.И. Евченко, В.В. Янюшкин // Моделирование. Теория, методы и средства: мат. VII междунар. науч.-практ. конф., Новочеркасск, 6 апреля 2007 г. – Новочеркасск: ЮРГТУ(НПИ), 2007. – С. 48-52.
 7. Янюшкин, В.В. Локальная вычислительная сеть комплексного морского тренажера / А.А. Михайлов, А.И. Евченко, В.В. Янюшкин // Моделирование. Теория, методы и средства: мат. VII междунар. науч.-практ. конф., Новочеркасск, 6 апреля 2007 г. – Новочеркасск: ЮРГТУ(НПИ), 2007. – С. 71-74.
 8. Янюшкин, В.В. Распределенный ресурс сеток высот и глубин в структуре морского тренажера / В.В. Янюшкин // Пилотируемые полеты в космос: мат. седьмой междунар. науч.-пр. конф., Звездный городок, 14-15 ноября 2007 г. – Звездный городок: ЦПК, 2007. – С. 118-119.
 9. Янюшкин, В.В. Методы фрактальных кластеров в задаче моделирования информационных систем / В.В. Янюшкин // Теория, методы проектирования, программно-техническая платформа корпоративных информационных систем: мат. VII междунар. науч.-практ. конф., Новочеркасск, 25 мая 2009 г. – Новочеркасск: ЮРГТУ(НПИ), 2009. – С. 173-181.
 10. Янюшкин, В.В. Постановка задач и общие алгоритмы оптимизации размещения данных модельного в тренажерно-моделирующих комплексах / В.В. Янюшкин // Теория, методы проектирования, программно-техническая платформа корпоративных информационных систем: мат. VI междунар. науч.-практ. конф., Новочеркасск, 26 мая 2008 г. – Новочеркасск: ЮРГТУ(НПИ), 2008. – С. 71-80.

11. Яньюшкин, В.В. Задача оптимального размещения элементов данных модельного мира в распределенной информационной системе / В.В. Яньюшкин // Теория, методы проектирования, программно-техническая платформа корпоративных информационных систем»: мат. V междунар. науч.-практ. конф., Новочеркасск, 25 мая 2007 г. – Новочеркасск: ЮРГТУ(НПИ), 2007. – С. 116-123.

12. Яньюшкин, В.В. Проблема размещения фрагментов данных в распределенной информационной системе / В.В. Яньюшкин // Студенческая весна 2007: сборник научных работ студентов и аспирантов ЮРГТУ(НПИ). – Новочеркасск: ЮРГТУ(НПИ), 2007. – С. 80-81.

13. Яньюшкин, В.В. Программные компоненты и архитектурные решения распределенных информационных систем на основе применения технологий cloud computing и WCF / В.В. Яньюшкин // Перспективы развития средств и комплексов связи. Подготовка специалистов связи: мат. межвуз. науч.-техн. конф. – Новочеркасск: НВВКУС, 2009. – С. 239-241.

14. Яньюшкин, В.В. Сериализация объектов модельного мира в распределенной информационной системе морского тренажера / В.В. Яньюшкин // Теория, методы проектирования, программно-техническая платформа корпоративных информационных систем: мат. V науч.-практ. конф., Новочеркасск, 25 мая 2007 г. – Новочеркасск: ЮРГТУ(НПИ), 2007. – С. 123-128.

15. Яньюшкин, В.В. Варианты организации рассылки и получения данных в вычислительной сети тренажера / В.В. Яньюшкин // Теория, методы проектирования, программно-техническая платформа корпоративных информационных систем: мат. V науч.-практ. конф., Новочеркасск, 25 мая 2007 г. – Новочеркасск: ЮРГТУ(НПИ), 2007. – С. 128-133.

16. Яньюшкин, В.В. Распределенный ресурс сеток высот и глубин в структуре морского тренажера / В.В. Яньюшкин, В.В. Торшина // Эврика-2007: сборник научных работ Всероссийского смотря-конкурса науч.-техн. творчества студ. вузов, Новочеркасск, 19-25 ноября 2007 г. – Новочеркасск: ОНИКС+, 2007. – С. 122-124.

17. Яньюшкин, В.В. Специализированные стандарты распределенного моделирования и задачи размещения информационных объектов / В.В. Яньюшкин // Теория, методы проектирования, программно-техническая платформа корпоративных информационных систем: мат. VI науч.-практ. конф., Новочеркасск, 26 мая 2008 г. – Новочеркасск: ЮРГТУ(НПИ), 2008. – С. 52-59.

18. Яньюшкин, В.В. Подход к применению генетических алгоритмов и конструированию структур данных в задаче оптимизации размещения данных / В.В. Яньюшкин // Теория, методы проектирования, программно-техническая платформа корпоративных информационных систем: мат. VI науч.-практ. конф., Новочеркасск, 26 мая 2008 г. – Новочеркасск: ЮРГТУ(НПИ), 2008. – С. 81-85.

19. Свидетельство о регистрации электронного ресурса 15134 «Программные компоненты оптимизации распределенной информационной системы тренажерно-моделирующего комплекса – Genetic Data Distribution in Simulators (GDDiS)» / В.В. Яньюшкин. – Зарегистрировано в объединённом фонде электронных ресурсов «Наука и образование» 24.12.09.

Личный вклад автора в опубликованных в соавторстве работах: в [6] – разработка схем и представления тренажера; [7] – архитектура локальной вычислительной сети; [16] – алгоритм получения данных из распределенного ресурса.

Янюшкин Вадим Вадимович

**МАТЕМАТИЧЕСКИЕ МОДЕЛИ ОПТИМИЗАЦИИ
РАСПРЕДЕЛЁННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ
ТРЕНАЖЁРНО-МОДЕЛИРУЮЩИХ КОМПЛЕКСОВ**

Автореферат

Подписано в печать 31.08.2010

Формат 60×84 1/16 . Бумага офсетная. Ризография.

Усл. печ. л. 1,0. Уч.-изд. л. 1,41. Тираж 100 экз. Заказ 48-730.

Отпечатано в ИД «Политехник»

346428, Новочеркасск, ул. Просвещения, 132

Тел., факс (863-52) 5-56-75