

СЕКЦИЯ 1

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

Среда, 21 апреля 2004 г., читальный зал преподавателей (ауд.232), гл. корпус МГТУ им.Н.Э.Баумана.

Начало в 10.00.

Председатель: профессор, д.т.н. Норенков И.П.

Руководитель экспертной комиссии: к.т.н., доцент Федорук В.Г.

Ученый секретарь: к.т.н., доцент Власов А.И.

Экспертная комиссия:

к.т.н., доцент БОЖКО А.Н., к.т.н., доцент ВОЛОСАТОВА Т. М., к.т.н., доцент ЖУК Д. М., к.т.н., доцент МАНИЧЕВ В. Б., к.т.н., доцент МАРТЫНЮК В. А., к.т.н., доцент ТРУДОНОШИН В. А., ст. преподаватель ГРОШЕВ С. В., ст. преподаватель КНЯЗЕВА С. ст. преподаватель РОДИОНОВ С. В., ст. преподаватель ФЕДОРУК Е. В., ст. преподаватель ШЕСТАКОВ С. А., к.т.н., доцент ПИВОВАРОВА Н. В., к.т.н., доцент КАМЫШНАЯ Э.Н., ассистент КОЛОСКОВ С.В., ассистент ОЯ Д.Р.

1. CRM - СИСТЕМЫ УПРАВЛЕНИЯ ВЗАИМООТНОШЕНИЯМИ С КЛИЕНТАМИ.
ЧТО? КАК? ЗАЧЕМ?

Шаповалов А.Ю., Решетников Р.И.

МГТУ им. Баумана, кафедра ИУ4.

2. МЕТОД СКРЫТИЯ ДАННЫХ В НЕПОДВИЖНЫХ ИЗОБРАЖЕНИЯХ

Фомичева А.М.

МГТУ им. Баумана, кафедра РК6.

3. ИССЛЕДОВАНИЕ ВЕЙВЛЕТ-ВОССТАНОВЛЕНИЯ СИГНАЛОВ В ПАКЕТЕ
WAVELET TOOLBOX СКМ MATHWORKS MATLAB 6.5 R 13.

Учамприн А. В., Кручинин Д. А.

МГТУ им. Баумана, кафедра РК6.

4. ИССЛЕДОВАНИЕ КРИПТОАЛГОРИТМА «RIJNDAEL»

Ильницкий А.О.

МГТУ им. Баумана, кафедра РК6.

5. МЕТОД И ПРОГРАММА ДЛЯ СКРЫТИЯ ТЕКСТОВ

Еранова Е.В.

МГТУ им. Баумана, кафедра РК6.

6. КОДИРОВАНИЕ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

Волокитин Д.В.

МГТУ им. Баумана, кафедра РК6.

7. ЛОГАРИФМИЧЕСКИЕ МЕТОДЫ КОДИРОВАНИЯ ЗВУКА

Соболева Ю.М.

МГТУ им. Баумана, кафедра РК6.

8. ПОСТРОЕНИЕ МОДУЛЯ МОНИТОРИНГА В ПРОГРАММНОЙ СИСТЕМЕ НА
ОСНОВЕ АКТИВНОЙ БАЗЫ ДАННЫХ

Костра В.В.

ИТМ НАН Украины, Днепропетровск, Украина

9. ПАРАЛЛЕЛЬНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМА БПФ

Лукьянец С.В.

МГТУ им. Баумана, кафедра РК6.

7

10. РАЗРАБОТКА АЛГОРИТМА ОБОБЩЕНИЯ ИНФОРМАЦИИ,
ОСНОВАННОГО НА ТЕОРИИ ПРИБЛИЖЕННЫХ МНОЖЕСТВ

Куликов А.В.

МЭИ (ТУ)

11. РАЗРАБОТКА СИСТЕМЫ УПРАВЛЕНИЯ ПРЕДПРИЯТИЕМ НА ОСНОВЕ
СОВРЕМЕННЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Иванцов И. С.

ЗАО «Мидаус»

12. СИСТЕМА МОДЕЛИРОВАНИЯ ПРОЦЕССА ЛИТЬЯ ВЫЖИМАНИЕМ В СРЕДЕ САПР.

Носенко А.Г.

МГТУ им. Баумана, кафедра РК6.

13. ВОПРОСЫ РАЗРАБОТКИ ПРОГРАММНОГО КОМПЛЕКСА МОДЕЛИРОВАНИЯ ДИНАМИКИ ТЕХНИЧЕСКИХ СИСТЕМ НОВОГО ПОКОЛЕНИЯ.

Алиев Е., Баженов С., Пузырев М.

МГТУ им. Баумана, кафедра РК6.

14. РЕАЛИЗАЦИЯ БЫСТРЫХ ПРЕОБРАЗОВАНИЙ ФУРЬЕ НА МНОГОПРОЦЕССОРНЫХ ВЫЧИСЛИТЕЛЬНЫХ УСТАНОВКАХ

Берчун Ю.В.

МГТУ им. Баумана, кафедра РК6.

15. РАЗРАБОТКА КОНЦЕПЦИИ МОДУЛЯ ДИСКРЕТНОГО МАШИНОСТРОИТЕЛЬНОГО ПРОИЗВОДСТВА В РАМКАХ ERP СИСТЕМЫ.

Бондарь Ю.А.

МГТУ им. Баумана, кафедра РК6.

16. СИСТЕМА МОДЕЛИРОВАНИЯ ПРОЦЕССОВ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ

Беломойцев Д.Е.

МГТУ им. Баумана, кафедра РК6.

17. СРАВНЕНИЕ ПОДХОДОВ К ПРОЕКТИРОВАНИЮ ПРОГРАММНЫХ СИСТЕМ

Андронов А.В.

МГТУ им. Баумана, кафедра РК6.

18. ИССЛЕДОВАНИЕ РЕШЕНИЯ ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ МЕТОДОМ РЕГУЛЯРИЗАЦИИ ТИХОНОВА

Воронов А.В.

МГТУ им. Баумана, кафедра РК6.

19. ДВУХУРОВНЕВОЕ МОДЕЛИРОВАНИЕ ОПТИКОЭЛЕКТРОННЫХ СИСТЕМ

Трудоношин И.В.

МГТУ им. Баумана, кафедра РК6.

20. МЕТОД АНАЛИТИЧЕСКОЙ НАСТРОЙКИ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧЕ КОНТЕКСТНОЙ КЛАССИФИКАЦИИ ТЕКСТОВ

Цыганов И.Г.

МГТУ им. Баумана, кафедра Иу4.

21. МЕТРИКИ ТЕКСТОВ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ ОБРАБОТКИ ИНФОРМАЦИИ

Руденко М. И., Цыганов И.Г.

МГТУ им. Баумана, кафедра Иу4.

22. КОМПЬЮТЕРНОЕ АДАПТИВНОЕ ТЕСТИРОВАНИЕ. АЛГОРИТМ И ТЕХНОЛОГИЯ СОЗДАНИЯ ТЕСТОВ.

Мигунов В.О.

МГТУ им. Баумана, кафедра Иу4.

23. ТЕХНОЛОГИИ ЭЛЕКТРОННОЙ ЦИФРОВОЙ ПОДПИСИ

Аксенов А.В.

МГТУ им. Баумана, кафедра РК6.

8

24. ТРЕНАЖЕР ДЛЯ ПОДГОТОВКИ ОПЕРАТОРОВ ПОДВОДНЫХ АППАРАТОВ

Н. Буйда, М. Григорьев, А. С. Медведев, Кошель Р.

МГТУ им.Н.Э.Баумана, СМ11.

25. ВОПРОСЫ ВИЗУАЛЬНОГО МОДЕЛИРОВАНИЯ СИСТЕМ В ТРЕНАЖЕРАХ ПОДВОДНЫХ ТЕЛЕУПРАВЛЯЕМЫХ КОМПЛЕКСОВ

А.Гнеломедов, С. Медведев

МГТУ им.Н.Э.Баумана, СМ11.

26. ВОПРОСЫ МОДЕЛИРОВАНИЯ ОБЪЕКТОВ ВИРТУАЛЬНОГО ПОДВОДНОГО МИРА
В ТРЕНАЖЕРАХ ПОДВОДНЫХ ТЕЛЕУПРАВЛЯЕМЫХ КОМПЛЕКСОВ

Н.Буйда, Кошель Р., М. Григорьев

МГТУ им.Н.Э.Баумана, СМ11.

27. МЕТОД ФОРМИРОВАНИЯ МОДЕЛИ ДИНАМИКИ ПОДВОДНОГО АППАРАТА В
ТРЕНАЖЕРНЫХ СИСТЕМАХ

С.Медведев, А.Гнеломедов, Климов Е.

МГТУ им.Н.Э.Баумана, СМ11.

28. СИСТЕМНОЕ КОГНИТИВНОЕ МОДЕЛИРОВАНИЕ КТР ЭВС

Кутаков С. В.

МГТУ им.Н.Э.Баумана, Иу4. __

CRM - СИСТЕМЫ УПРАВЛЕНИЯ ВЗАИМООТНОШЕНИЯМИ С КЛИЕНТАМИ. ЧТО? КАК? ЗАЧЕМ?

Шаповалов А.Ю., Решетников Р.И.
Научный руководитель: к.т.н., доцент Власов А.И.
МГТУ имени Н.Э. Баумана, Москва, Россия.

CRM - CUSTOMER RELATIONSHIP MANAGEMENT SYSTEMS. WHAT? HOW? WHY?

Shapovalov A.U., Reshetnikov R.I.
Scientific adviser: candidate of technical science, docent Vlasov A. I.
Moscow State Technical University named after N.E. Bauman, Moscow, Russia.
anshap@land.ru, romresh@mail.ru .

Аннотация

В докладе описываются основы концепции CRM (Customer Relationship Management – Управление взаимоотношениями с клиентами). Отражены следующие составляющие концепции: история, структура, необходимость применения, стратегия бизнеса и развития предприятия на основе CRM решений. Показаны цели внедрения CRM системы на предприятии, ее функции и возможности по доработке стандартных модулей. Приводится анализ эффективности внедрения CRM. Особое внимание уделяется отраслевым решениям на базе стандартных модулей на примере радиотехнического предприятия и телекоммуникационной компании. В заключении приводится обзор решений, представленных в России и тенденции развития CRM систем. <http://crm.iu4.bmtu.ru>.

Abstract

The Client Relationship Management (CRM) conception fundamentals are described in this report. The CRM conception constituents, such as history, structure, application necessity, business strategy and CRM based enterprise development are presented here. The aims of the CRM introduction and integration, CRM functionality and standard modules revision possibility are demonstrated. Efficiency analysis is also considered. Special consideration is given to the CRM branch decisions, based on the standard modules, by the example of the radio engineering enterprise and telecommunicational company. In conclusion CRM decisions presented in Russia are reviewed; CRM development tendency is presented.

CRM (Customer Relationship Management – Управление взаимоотношениями с клиентами) - это не программный продукт и не технология. Это даже не набор продуктов.

CRM - это направленная на построение устойчивого бизнеса концепция и бизнес стратегия, ядром которой является "клиенто - ориентированный" подход.

Эта стратегия основана на использовании передовых управленческих и информационных технологий, с помощью которых компания собирает информацию о своих клиентах на всех стадиях его жизненного цикла (привлечение, удержание, лояльность), извлекает из нее знания и использует эти знания в интересах своего бизнеса путем выстраивания взаимовыгодных отношений с ними.

Результатом применения стратегии является повышение конкурентоспособности компании, и увеличение прибыли, так как правильно построенные отношения, основанные на персональном подходе к каждому клиенту, позволяют привлекать новых клиентов и помогают удержать старых.

На уровне технологий CRM – это набор приложений, связанных единой бизнес - логикой и интегрированных в корпоративную информационную среду компании (часто в виде надстройки над ERP) на основе единой базы данных. Специальное программное обеспечение позволяет провести автоматизацию соответствующих бизнес - процессов в маркетинге, продажах и обслуживании. Как результат, компания может обратиться к "нужному" заказчику в "правильный" момент времени, с наиболее эффективным предложением и по наиболее удобному заказчику каналу взаимодействия.

На практике интегрированная система CRM обеспечивает координацию действий различных отделов, обеспечивая их общей платформой для взаимодействия с клиентами. С

этой точки зрения назначение CRM - исправить ситуацию, когда отделы маркетинга, продаж и сервиса действуют независимо друг от друга, причем их видение заказчика часто не совпадает, а действия несогласованны.

С точки зрения управления бизнесом эффект от внедрения CRM проявляется в том, что процесс принятия решения за счет автоматизации переносится на более низкий уровень и унифицируется. За счет этого повышается скорость реакции на запросы, растет скорость оборота средств, снижаются издержки.

Наконец, CRM включает себя идеологию и технологии создания истории взаимоотношений клиента и фирмы, что позволяет более четко планировать бизнес и повышать его устойчивость

Ключевыми функциями CRM-систем являются следующие моменты:

1. Сбор информации. Система позволяет сотруднику организации удобным способом вводить информацию о клиенте в базу данных, либо же самому клиенту вводить эту информацию. В CRM-систему вводится вся доступная информация о клиенте. Естественно, учитывается информация, которая относится к взаимодействию клиент-компания. Также в систему вводится личная информация по клиенту (возраст, семейное положение, ежегодный доход, имущество и пр.). Вся данная информация заносится и обновляется при каждом взаимодействии компании с клиентом, т.е. при любом контакте между двумя сторонами, будь то личное посещение компании клиентом, телефон, почта, факс или интернет.
2. Хранение и обработка. Система позволяет сохранять и ранжировать полученную информацию, в соответствии с заданными критериями. Причем вся информация хранится в стандартной для корпорации форме. Также CRM-система в соответствии с заданными параметрами может анализировать полученную информацию, для последующего экспорта этой информации.
3. Экспорт информации. Предоставление информации CRM-системой является ее главной функцией. Информация, хранящаяся в системе, может быть затребована разными подразделениями и в разном виде. Например, система CRM может определить, какой товар лучше предложить определенному клиенту, на основе экстраполяции исторических данных, или, например, клиент является постоянным покупателем и система напомнит, что клиенту полагается скидка. Также сотруднику компании может просто понадобиться информации об исторических контактах клиента с фирмой, и система позволяет предоставить эту информацию в наглядном виде. Естественно, система позволяет выводить информацию, как по отдельному клиенту, так и по целевой группе в целом. Если для сотрудника отдела продаж наиболее важна информация по определенному клиенту, то для отдела маркетинга, скорее, важна агрегированная информация по определенной группе.

Существует множество проблем, для решения которых компании внедряют эти системы, всех их можно, так или иначе, свести к следующим:

1. Совершенная конкуренция. Современные технологии привели к тому, что покупатель получает доступ к любой части рынка при малых транзакционных издержках, асимметричность информации почти равна нулю и пр. Поэтому основной задачей для компании является удержание имеющихся клиентов.
2. Мультиканальность взаимоотношений. Контакт между клиентом и фирмой может осуществляться разными способами - телефон, факс, web-сайт, почта, личный визит. И клиент ожидает, что вся получаемая по этим каналам информация при следующем взаимоотношении, будет рассматриваться компанией во всей совокупности.
3. Изменение рыночной ориентации компаний. Переход большинства компаний от продукто- и производственно-ориентированных концепций к концепциям маркетинга. Многими компаниями предел качества и минимизации издержек уже достигнуты (в том числе от использования ERP) и клиенты больше обращают внимание на сопровождающие моменты покупки и обслуживание.

Из всего выше сказанного, можно сделать вывод, что CRM-системы - это очень гибкие и разносторонние системы. Каждая компания, при внедрении CRM-системы, может использовать ее для достижения разных целей. Соответственно, в зависимости от целей, которые ставит перед собой компания, внедряя систему CRM, будут изменяться входящие параметры и способ обработки, анализа этих данных. Можно выделить три вида систем по целям, для которых обычно используется стандартные CRM-системы.

- **Оперативное использование.** Система используется сотрудником компании для оперативного доступа к информации по конкретному клиенту в ходе непосредственного взаимоотношения с ним - процессы продажи и обслуживания. В данном случае основным компонентом системы является приложение, которые в наглядном виде позволяет представить сотруднику накопленную информацию по отдельному клиенту. В первую очередь от системы требуется хорошей интеграции между всеми подсистемами, и возможности пополнять базу данных при процессах взаимодействия с клиентом по всем каналам. Данный тип CRM-систем является наиболее распространенным в традиционном бизнесе.
- **Аналитическое использование.** Система используется для анализа различные данные, также ищет статистические закономерности в этих данных для выработки наиболее эффективной стратегии маркетинга, продаж, обслуживания клиентов и т. п. Требуется хорошей интеграции подсистем, большого объема наработанных статистических данных, эффективного аналитического инструментария, интеграции с другими системами, автоматизирующие деятельность предприятия. Данные, генерируемые такими системами, могут быть затребованы, как отделом маркетинга, так и представлены самому клиенту без посредничества сотрудников компании. Данный тип CRM-систем чаще применяется в электронной коммерции, нежели в традиционном бизнесе.
- **Коллаборационное использование.** CRM системы предоставляет клиенту возможность гораздо большего влияния на деятельность фирмы в целом, в том числе на процессы разработки дизайна, производства, доставки и обслуживания продукта. От системы требуется технологий, которые позволяют с минимальными затратами подключить клиента к сотрудничеству в рамках внутренних процессов компании. Опять же, для доступа к данным системам, зачастую, клиент использует сеть интернет, поэтому данные системы наиболее распространены в сфере электронной коммерции.

Таким образом, концепция CRM весьма многогранна. Хотя ее отдельные элементы культивировались и раньше (например, уже давно делается акцент на обеспечении наивысшего качества продукта и сервиса), сами по себе они не достаточны.

Суть же концепции CRM в том, что наиболее желанный и прибыльный клиент имеет право на первоочередное и эксклюзивное обслуживание. Кроме того, концепция CRM ориентирует компанию на длительные взаимоотношения с клиентом.

Главная идея в том, чтобы учиться у своего клиента, иметь обратную связь и работать так, как клиент хочет.

Литература

1. Норенков И.П. Основы автоматизированного проектирования: Учеб. для вузов/ И.П. Норенков. - М.: Изд-во МГТУ им. Н.Э.Баумана, 2002 --ISBN 5-7038-2090-1.
2. Шахнов В.А "Конструкторско-технологическое проектирование электронной аппаратуры". Издательство: МГТУ им. Баумана 2002.
3. spklabs.com.
4. cfin.ru.
5. www.e-commerce.ru.

МЕТОД СКРЫТИЯ ДАННЫХ В НЕПОДВИЖНЫХ ИЗОБРАЖЕНИЯХ

Фомичева А.М.

научный руководитель: к.т.н., доцент Волосатова Т.М.

МГТУ им. Н.Э.Баумана, кафедра РК6, Москва, Россия.

METHOD OF DATA HIDING IN STILL IMAGES

Fomicheva A.M.

The scientific chief: Ph.D., senior lecturer Volosatova T.M.

BMSTU, RK6, Moscow, Russia.

mifom@yandex.ru

Аннотация

В докладе рассмотрен один из методов внедрения цифрового водяного знака в изображение. Этот метод модифицирует синюю цветовую компоненту пикселя таким образом, что изменения не заметны для человека, но легко обнаруживаются численными методами. В соответствии с этим алгоритмом была написана программа, имеющая успешные результаты.

Abstract

In the report one of methods of introduction of a digital watermark in still image is considered. This method modifies blue color component of pixel in such a manner that changes are not appreciable for the person, but are easily found out by numerical methods. According to this algorithm the program having successful results has been written.

Введение

На сегодняшний день существует два основных направления решения задачи защиты информации: криптография и стеганография. Цель криптографии – скрытие содержимого сообщения за счет их шифрования. В отличие от этого, при стеганографии скрывается сам факт существования тайного сообщения.

Слово «стеганография» имеет греческие корни и буквально означает «тайнопись». Исторически это направление появилось первым, но затем во многом было вытеснено криптографией. Тайнопись осуществляется самыми различными способами. Общей чертой является то, что скрываемое сообщение встраивается в некоторый безобидный, не привлекающий внимания объект, который затем открыто транспортируется адресату. При криптографии наличие шифрованного сообщения само по себе привлекает внимание противников, при стеганографии же наличие скрытой связи остается незаметным.

Развитие средств вычислительной техники в последнее десятилетие дало новый толчок развитию компьютерной стеганографии.

Сообщения встраивают теперь в цифровые данные, как правило, имеющие аналоговую природу – речь, аудиозаписи, изображение, видео и даже текстовые файлы и исполняемые файлы программ, преобразованные в цифровую форму.

Скрытие данных в неподвижных изображениях

Можно выделить две причины популярности исследований в области стеганографии в настоящее время: ограничение на использование криптосредств в ряде стран мира и появление проблемы защиты прав собственности на информацию, представленную в цифровом виде. Первая причина повлекла за собой большое количество исследований в духе классической стеганографии (то есть скрытие факта передачи информации), вторая – еще более многочисленные работы в области так называемых водяных знаков. Цифровой водяной знак (ЦВЗ) – специальная метка, незаметно внедряемая в изображение или другой сигнал с целью контролировать его использование.

ЦВЗ могут применяться в основном для защиты от копирования и несанкционированного использования. Невидимые ЦВЗ анализируются специальным декодером, который выносит решение об их корректности. ЦВЗ могут содержать некоторый аутентичный код, информацию о собственнике либо какую-нибудь управляющую информацию. Наиболее подходящими объектами защиты при помощи ЦВЗ являются неподвижные изображения, файлы аудио- и видеоданных.

Большинство исследований посвящено использованию в качестве стегоконтейнеров изображений. Это обусловлено следующими причинами:

- Существованием практически значимой задачей защиты фотографий, картин, видео от незаконного тиражирования и распространения;
- Относительно большим объемом цифрового представления изображений, что позволяет внедрять цифровые водяные знаки (ЦВЗ) большого объема либо повышать робастность внедрения, то есть повышать устойчивость ЦВЗ к различным преобразованиям, производимым над файлом-контейнером;
- Заранее известным размером контейнера, отсутствием ограничений, накладываемых требованиями реального времени;
- Слабой чувствительностью человеческого глаза к незначительным изменениям цветов изображения, его яркости, контрастности, содержанию в нем шума, искажениям вблизи контуров;
- Хорошо разработанными в последнее время методами цифровой обработки изображений.

Надо отметить, что последняя причина вызывает и значительные трудности в обеспечении робастности ЦВЗ: чем более совершенными становятся методы сжатия, тем меньше остается возможностей для встраивания посторонней информации. Развитие теории и практики алгоритмов сжатия изображений привело к изменению представлений о технике внедрения ЦВЗ. Если первоначально предлагалось вкладывать информацию в незначимые биты для уменьшения визуальной заметности, то современный подход заключается во встраивании ЦВЗ в наиболее существенные области изображений, разрушение которых приведет к полной деградации самого изображения. Не случайно поэтому стегоалгоритмы учитывают свойства системы человеческого зрения (СЧЗ) аналогично алгоритмам сжатия изображений. В стегоалгоритмах зачастую используются те же преобразования, что и в современных алгоритмах сжатия (дискретное косинусное преобразование в JPEG, вейвлет-преобразование в JPEG2000). При этом существуют, очевидно, три возможности. Вложение информации может производиться в исходное изображение либо одновременно с осуществлением сжатия изображения-контейнера, либо в уже сжатое алгоритмом JPEG изображение.

Выполнение линейных ортогональных преобразований изображений – вычислительно трудоемкий процесс, несмотря на наличие быстрых алгоритмов. Поэтому в некоторых случаях можно ограничиться встраиванием информации в пространственной области изображения.

Предложенный алгоритм внедряет ЦВЗ в область исходного изображения. Его преимущество в том, что для внедрения ЦВЗ нет необходимости выполнять вычислительно громоздкие линейные преобразования изображений. ЦВЗ внедряется за счет манипуляций яркостью и цветовыми составляющими (r, g, b) .

Алгоритм внедрения ЦВЗ

24-битное изображение имеет RGB -кодировку. Встраивание выполняется в канал синего цвета, так как к синему цвету система человеческого зрения наименее чувствительна. Рассмотрим алгоритм передачи одного бита секретной информации.

Пусть s_i – встраиваемый бит, $I = \{R, G, B\}$ – контейнер, $p = (x, y)$ – позиция, в которой выполняется вложение. Секретный бит встраивается в канал синего цвета путем модификации яркости

$$l(p) = 0.299r(p) + 0.587g(p) + 0.114b(p)$$

$$b_1(p) = \begin{cases} (b(p) - ql(p)) & \text{if } s_i = 1 \\ (b(p) + ql(p)) & \text{if } s_i = 0 \end{cases}$$

где q - константа, определяющая энергию встраиваемого сигнала. Ее величина зависит от предназначения схемы. Чем больше q , тем выше робастность вложения, но тем сильнее его заметность.

Алгоритм извлечения ЦВЗ

Извлечение бита получателем осуществляется без наличия у него исходного изображения, то есть вслепую. Для этого выполняется предсказание значения исходного, немодифицированного пикселя на основании значений его соседей. Для получения оценки пикселя можно использовать значения 2 пикселей, расположенных справа и слева от оцениваемого в той же строке.

Оценка b получается в виде:

$$Avg = \frac{1}{5} \cdot \sum_{i=1}^5 b_i$$

Знак разности и значение

$$Avg - b_3$$

определяют значение встроенного бита.

При использовании исходного файла значение встроенного бита определяется знаком разности значений b для закодированного и закодированного изображения.

Результаты работы программы

Разработанная программа встраивает ЦВЗ в 24-битный bmp-файл, а также извлекает его обоими вышеописанными методами. ЦВЗ представляется в виде вертикальных полос: белые пиксели соответствуют 1, а черные – 0, серые пиксели не содержат информации. В результате раскодирования создаются файлы, отображающие вид встроенного сигнала. В случае использования метода оценки яркости соседних пикселей точность определения ЦВЗ намного ниже, но тем не менее полосы четко различимы на фоне всего изображения. Размер исходного файла при внедрении в него ЦВЗ не изменяется. Программа функционирует в среде Windows 2000, XP.

Литература

1. Грибунин В.Г., Оков И.Н., Туринцев И.В. Цифровая стеганография. – М.: СОЛОН-Пресс, 2002.
2. Bender W., Gruhl D., Morimoto N., Lu A. Techniques for Data Hiding // IBM System Journal. 1996.
3. Мюррей Дж. Энциклопедия форматов графических файлов. 1996.

ИССЛЕДОВАНИЕ ВЕЙВЛЕТ-ВОССТАНОВЛЕНИЯ СИГНАЛОВ В ПАКЕТЕ WAVELET TOOLBOX СКМ MATHWORKS MATLAB 6.5 R 13.

Учамприн А.В., Кручинин Д.А.

Научный руководитель: к.т.н., доцент Волосатова Т.М.
МГТУ им. Н. Э. Баумана, кафедра САПР (РК6), Москва, Россия

RESEARCH OF SIGNAL'S WAVELET-RESTORING IN WAVELET TOOLBOX MATHWORKS MATLAB 6.5 R 13.

Uchamprin A.V., Kruchinin D.A.

The scientific chief: Ph.D., senior lecturer Volosatova Tamara Mihajlovna
MSTU named after Bauman, RK6 (CAD/CAM/CAE) department, Moscow, Russia
blaze@aviel.ru, bounseden@yandex.ru, <http://blaze.aviel.ru>

Аннотация

В данной работе при помощи пакета расширения системы 6.5 Wavelet Toolbox 2/2.5 исследуется преобразование, фильтрация и вейвлет-восстановление различных сигналов:

вейвлет-анализ сигналов, непрерывное одномерное вейвлет-преобразование, дискретное одномерное вейвлет-преобразование, обратное вейвлет-преобразование, прямое восстановление из одномерных вейвлет-коэффициентов, многоуровневое одномерное вейвлет-разложение, многоуровневое одномерное вейвлет-восстановление, восстановление одиночной ветви из одномерных вейвлет-коэффициентов.

Abstract

By means of using add-on 6.5 of system Wavelet Toolbox 2/2.5, in this research investigates transformation, filtration and Wavelet-restoring of some signals: Wavelet-analysis of signals, continuous plain Wavelet-transformation, discrete plain Wavelet-transformation, reverse Wavelet-transformation, direct reconstruction from plain Wavelet coefficients, multi-level plain Wavelet decomposition, multi-level plain Wavelet reconstruction, reconstruction of a single branch from plain Wavelet coefficients.

Введение

В последнее время наметилась тенденция к использованию широкополосных импульсных и цифровых сигналов (видеоимпульсная локация, видеосредства компьютеров и т.д.), которые относятся к классу нестационарных. Ряды Фурье плохо годятся для представления коротких локальных особенностей сигналов и функций, таких как перепады и скачки, поскольку они предельно локализованы в частотной области и очень плохо локализованы во временной. Таким образом назрела необходимость в создании нового математического аппарата приближения функций и сигналов.

Термин «вейвлет», введённый впервые Морле (Morlet), в переводе с английского означает «короткая или маленькая волна». Вейвлеты занимают промежуточное положение между крайними случаями – синусоидой и импульсной функцией. Довольно грубо вейвлеты можно представить как некоторые волновые функции, способные осуществлять преобразование Фурье не по всей временной оси, а локально по месту своего расположения.

С помощью вейвлетов сигнал представляется совокупность волновых пакетов - вейвлетов, образованных на основе некоторой исходной (базовой) функции. Эта совокупность, разная в разных частях временного интервала определения сигнала и корректируемая множителями, имеющими вид сложных временных функций, и представляет сигнал с той или иной степенью детализации. Такой подход и называют вейвлет-анализом сигналов.

Число используемых при разложении сигнала вейвлетов задаёт уровень декомпозиции сигнала. При этом за нулевой уровень декомпозиции часто принимается сам сигнал, а последующие уровни декомпозиции образуют ниспадающее вейвлет-дерево того или иного вида. Точность представления сигнала по мере перехода на более низкие уровни декомпозиции снижается, но зато появляется возможность вейвлет-фильтрации сигналов, удаления из сигналов шумов и эффективной компрессии сигналов, т.е. становится возможной вейвлет-обработка сигналов.

Прямое вейвлет-преобразование означает разложение произвольного входного сигнала на принципиально новый базис в виде совокупности волновых пакетов – вейвлетов, которые:

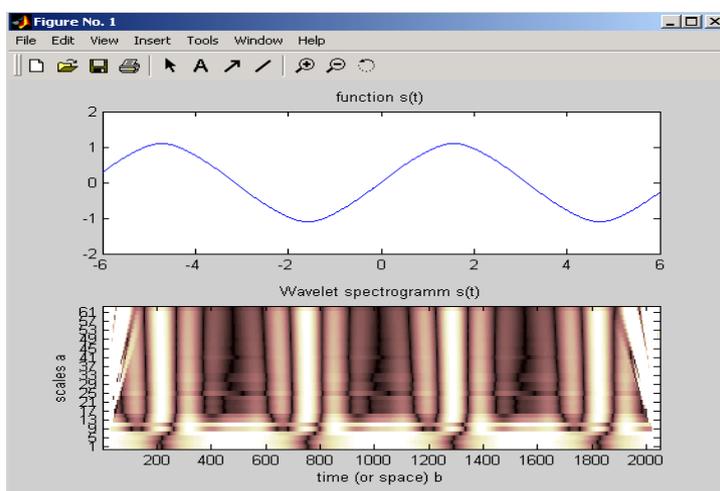
- 1) имеют вид коротких, локализованных во времени (или пространстве), волновых пакетов с нулевым значением интеграла.
- 2) Обладают возможностью сдвига по времени.
- 3) Способны к масштабированию (сжатию/растяжению).
- 4) Имеют ограниченный (или локальный) частотный спектр.

По обилию типов вейвлетов и функций для обработки сигналов пакет Matlab 6.0/6.5 Wavelet Toolbox 2/2.5 является лучшим среди пакетов расширения для систем компьютерной математики (СКМ) в этой области (пакеты расширения по вейвлетам есть и в новых реализациях СКМ Mathcad и Mathematica).

Непрерывное одномерное вейвлет – преобразование

Вейвлет-спектрограмма слегка искажённой функции синуса:

Пусть имеется функция, описывающая сигнал вида $\sin(t)+0.1*\sin(t)^3$. Составим программу построения графика этой функции и спектрограммы на основе первых 64 вейвлетов:



Вейвлет-спектрограмма отчётливо выделяет все особенности функции в точках перегиба. Светлые столбы спектрограммы отчётливо выделяют экстремумы функции, но между ними видны представления каких-то на первый взгляд непонятных локальных особенностей данной функции.

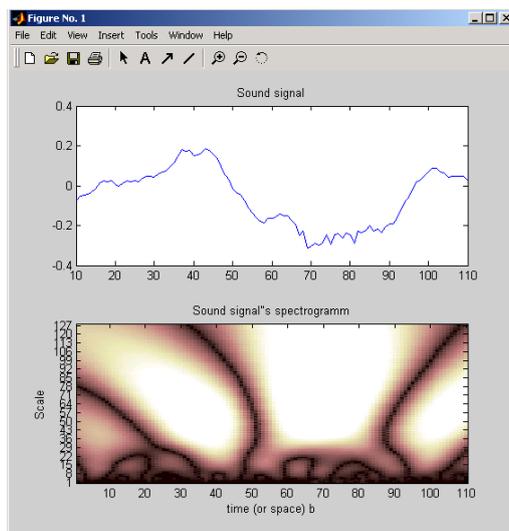
В этом примере хорошо видно, что увеличение числа вейвлет-коэффициентов уже не бесполезно – картина спектра для больших по номеру коэффициентов выглядит более стабильной, чем для малых. Но и здесь увеличение числа вейвлет-коэффициентов выше 40 перестаёт влиять на вид спектра.

Вейвлет-анализ реальных звуковых сигналов

Построим временную диаграмму и вейвлет-спектрограмму реального звукового сигнала, загружаемого из файла mtlb с выборкой в 110 отсчётов:

Изменив число выборок на 1010 можно получить те же графические зависимости. Прекрасно выделяются частотные компоненты звукового сигнала, имеющие разные частоты. Так, в нижней части отчётливо видны частые изменения яркости спектрограммы, указывающие на наличие периодических высокочастотных компонент, тогда как в верхней части заметны изменения яркости менее частые, соответствующие более низкочастотным компонентам.

Спектрограмма утрачивает детальность в большом интервале времени. Это объяснимо – вейвлеты принципиально приспособлены к детализировке локальных изменений сигнала, занимающих небольшие промежутки времени.



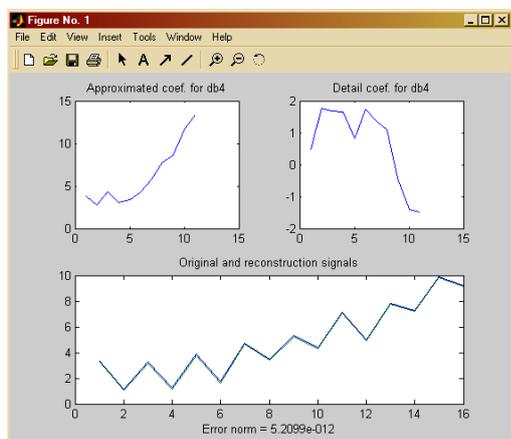
Как показывают приведённые примеры, вейвлет-спектрограммы наиболее пригодны для анализа тонкой структуры сигналов, содержащих резкие скачки, переходы производных через 0 и т.д. К таким сигналам сейчас относятся звуковые сигналы речи и музыки и сигналы изображений.

Дискретное одномерное вейвлет-преобразование

Наряду с непрерывным вейвлет-преобразованием существует дискретное вейвлет-преобразование. Его главным преимуществом является наличие эффективных алгоритмов быстрого вейвлет-преобразования, которое напоминает быстрое преобразование Фурье. В частности, для быстрого вейвлет-преобразования (БВП) может эффективно использоваться пирамидальный алгоритм с прореживанием по частоте. Благодаря этому появляется возможность анализа больших выборок за вполне приемлемое для практических целей время. Но эти возможности реализуются не для всех типов вейвлетов.

Одноуровневое обратное вейвлет-преобразование

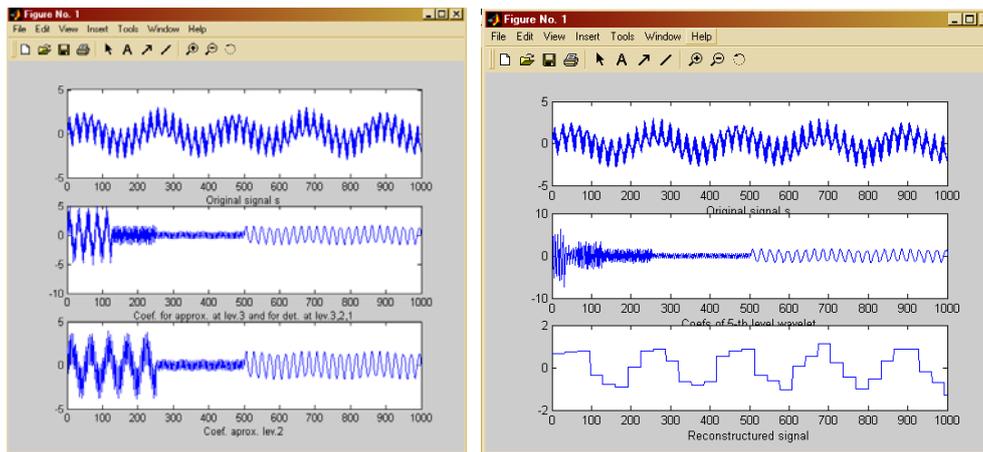
Таким образом, погрешность реконструкции равна $5,2 \cdot 10^{-12}$, что близко к точности машинных расчётов двойной точности.



Известно, что в реальных условиях при ограничении числа гармоник точное восстановление случайного сигнала при использовании прямого и обратного преобразований Фурье невозможно. Вейвлет – преобразование способно обеспечить точную реконструкцию сигнала после его прямого и обратного преобразований.

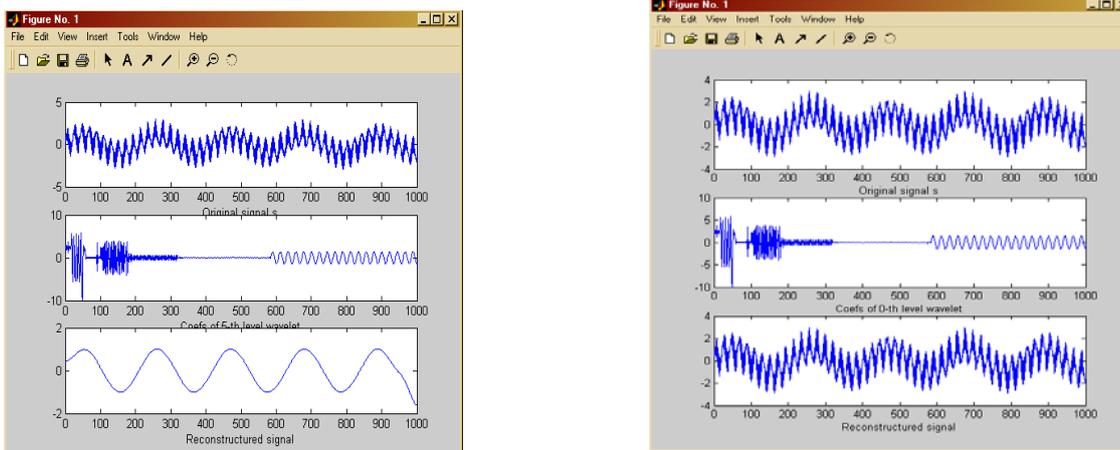
Многоуровневое одномерное вейвлет-разложение.

Применим данную функцию к сигналу `sumsin` – сумма синусов с разной частотой и амплитудой:



Восстановление одиночной ветви из одномерных вейвлет-коэффициентов

Видно, что в данном случае вейвлет Хаара не обеспечивает приемлемой степени реконструкции низкочастотной синусоиды. Выходной сигнал имеет вид ступенчатой кривой, в которой угадывается низкочастотная компонента сигнала. Высокочастотная компонента отсеивается, т.е. реконструкция высокого уровня сопровождается эффектом фильтрации. Применяя в предыдущем случае вейвлет Добеши 'db10' :
 Теперь низкочастотная компонента сигнала выглядит как синусоидальная функция (за исключением небольшого участка в конце). Высокочастотная компонента отсеяна. Снизив уровень восстановления, получим полностью восстановленный сигнал.



Итак, уже можно сделать вывод, что вейвлет-преобразование наиболее пригодно для работы с сигналами на небольших промежутках времени, сигналами, содержащими резкие скачки, переходы производных через 0, а на больших промежутках времени восстановление утрачивает точность и детальность.

Литература

1. И. Добеши «Десять лекций по вейвлетам». R&C Dynamics, 2001
2. В.П. Дьяконов «Вейвлеты – от теории к практике». Солон – Р, 2002
3. <http://www.exponenta.ru>

ИССЛЕДОВАНИЕ КРИПТОАЛГОРИТМА «RIJNDAEL»

Ильницкий А.О.

Научный руководитель: к. т. н., доцент Волосатова Т.М.
МГТУ им. Н.Э. Баумана, кафедра РК6, Москва, Россия.

RESEARCH OF CRYPTOALGORYTHM «RIJNDAEL»

Иlnitskiy A.O.

The scientific chief: Ph.D., senior lecturer Volosatova T.M.
BMSTU, RK6, Moscow, Russia.

Аннотация

В докладе рассматриваются основные особенности и характеристики блочного криптоалгоритма «RIJNDAEL» и предлагается его программная реализация.

Abstract

The report describes main features and characteristics of block cryptoalgorhythm «RIJNDAEL» and its program realization.

Введение

Основной целью криптографической защиты или криптографического закрытия информации является защита от утечки информации, которая обеспечивается путём обратимого однозначного преобразования сообщений или хранящихся данных в форму, непонятную для посторонних или неавторизованных лиц. Все современные шифры базируются на принципе Кирхгофа, согласно которому секретность шифра обеспечивается секретностью ключа, а не секретностью алгоритма шифрования. Стойкость криптосистемы зависит от сложности алгоритмов преобразования, длины ключа, а точнее от объёма ключевого пространства, метода реализации. Качественный шифр невозможно раскрыть способом более эффективным чем полный перебор по всему ключевому пространству, при этом криптограф должен рассчитывать только на то, что у противника не хватит времени и ресурсов, чтобы это сделать.

Описание алгоритма

Криптоалгоритм «RIJNDAEL» является блочным шифром, который был разработан бельгийскими учёными. В 2002 году по результатам конкурса был выбран за основу американского стандарта шифрования AES.

Этот алгоритм обладает следующими преимуществами:

- высокая эффективность на любых платформах
- высокий уровень защищённости
- хорошо подходит для реализации в smart-картах из-за низких требований к памяти
- быстрая процедура формирования ключа
- хорошая поддержка параллелизма на уровне инструкций
- поддержка разных длин ключа с шагом 32 бита

Недостатки:

- уязвим к анализу мощности

«RIJNDAEL» - это итерационный блочный шифр, имеющий архитектуру “Квадрат”. Промежуточные результаты преобразований, выполняемых в рамках криптоалгоритма, называются состояниями (*State*). Состояние можно представить в виде прямоугольного массива байтов (рис. 1).

a0	a4	a8	a12
a1	a5	a9	a13
a2	a6	a10	a14
a3	a7	a11	a15

рис. 1 (*State*)

Ключ шифрования также представлен в виде прямоугольного массива с четырьмя строками. Число столбцов N_k этого массива равно длине ключа, делённой на 32.

Раунд состоит из четырёх различных преобразований:

- Замена байтов *SubBytes()* – побайтовой подстановки в S-блоках с фиксированной таблицей замен.
- Сдвиг строк *ShiftRows()* – побайтового сдвига строк массива состояния на разное количество байт.
- Перемешивание столбцов *MixColumns()* – умножение столбцов состояния, рассматриваемых как многочлены над $GF(2^8)$, на многочлен третьей степени $g(x)$ по модулю $x^4 + 1$.
- Сложение с раундовым ключом *AddRoundKey()* – поразрядного XOR с текущим фрагментом развёрнутого ключа.

Характеристики производительности

В таблице 1 приведены данные по времени выполнения RIJNDAEL на процессорах семейства Pentium Pro и Pentium II. Результаты были получены с использованием компилятора Visual C++ (версия 6).

таблица 1. (характеристика производительности)

Длина ключа / размер блока, бит	Количество циклов, необходимых для разворачивания ключа		Производительность шифрования	
	RIJNDAEL	RIJNDAEL ⁻¹	Скорость, Мбит/сек	Число циклов / блок
(128, 128)	305	1389	70,5	363
(192, 128)	277	1595	59,3	432
(256, 128)	374	1960	51,2	500

На основе данного алгоритма была разработана программа «RIJNDAEL», оперирующая блоками в 128 бит с длиной ключа 128 бит. Программный продукт «RIJNDAEL» позволяет зашифровать и в последствии расшифровать любую последовательность байт, представленных в виде файла любого типа. Программа предоставляет удобный интерфейс для операционных систем семейства Windows. Программа может быть легко перекомпилирована под операционную систему семейства Unix с предварительной разработкой интерфейса под конкретную ОС.

Литература

1. Баричев С.Г., Гончаров В.В., Серов Р.Е. Основы современной криптографии: Учебный курс. – 2-е изд., испр. и доп.- М.: Горячая линия – Телеком, 2002. – 175 с.: ил.
2. Зенин О.С., Иванов М.А. Стандарт криптографической защиты – AES. Конечные поля / Под ред. М.А. Иванова – М.: КУДИЦ-ОБРАЗ, 2002.- 176 с. – (СКБ – специалисту по компьютерной безопасности).

МЕТОД И ПРОГРАММА ДЛЯ СКРЫТИЯ ТЕКСТОВ

Еранова Е.В.

научный руководитель: к.т.н., доцент Волосатова Т.М.

МГТУ им. Н.Э.Баумана, кафедра РК6, Москва, Россия.

METHOD AND PROGRAMM FOR TEXT HIDING

Eranova E.V.

The scientific chief: Ph.D., senior lecturer Volosatova T.M.

BMSTU, RK6, Moscow, Russia.

sherry-kar@mtu-net.ru

Аннотация

В работе представлен метод сокрытия изображений (текстов) и его программная реализация.

Abstract

In the document the method for image (texts) hiding and its program realization is considered.

Введение

Проблема надежной защиты информации от несанкционированного доступа является одной из древнейших и не решенных до настоящего времени задач. Способы и методы сокрытия секретных сообщений известны с давних времен, причем, данная сфера человеческой деятельности получила название стеганография. Это слово происходит от греческих слов *steganos* (секрет, тайна) и *graphy* (запись) и, таким образом, означает буквально “тайнопись”, хотя методы стеганографии появились, вероятно, раньше, чем появилась сама письменность (первоначально использовались условные знаки и обозначения).

В современной компьютерной стеганографии существует два основных типа файлов: сообщение— файл, который предназначен для сокрытия, и контейнер—файл, который может быть использован для сокрытия в нем сообщения. При этом контейнеры бывают двух типов. Контейнер—оригинал (или “Пустой” контейнер) — это контейнер, который не содержит скрытой информации. Контейнер—результат (или “Заполненный” контейнер) — это контейнер, который содержит скрытую информацию. Под ключом понимается секретный элемент, который определяет порядок занесения сообщения в контейнер.

Основными положениями современной компьютерной стеганографии являются следующие:

1. Методы сокрытия должны обеспечивать аутентичность файла. Аутентичность предполагает абсолютное совпадение скрываемого и восстановленного текста.

2. Предполагается, что эксперту полностью известны возможные стеганографические методы.

3. Безопасность методов основывается на сохранении стеганографическим преобразованием основных свойств открыто передаваемого файла при внесении в него секретного сообщения и некоторой неизвестной противнику информации — ключа.

4. Даже если факт сокрытия сообщения стал известен эксперту, извлечение самого секретного сообщения представляет собой задачу, решение которой требует значительных затрат на создание и эксплуатацию программно-аппаратных средств дешифровки.

В связи с возрастанием роли глобальных компьютерных сетей становится все более важным значение стеганографии. Анализ информационных источников компьютерной сети Internet позволяет сделать вывод, что в настоящее время стеганографические системы активно используются для решения следующих основных задач:

1. Защита конфиденциальной информации от несанкционированного доступа.
2. Преодоление систем мониторинга и управления сетевыми ресурсами.
3. Защита авторского права на некоторые виды интеллектуальной собственности.

В настоящее время методы компьютерной стеганографии развиваются по двум основным направлениям:

1. Методы, основанные на использовании специальных свойств компьютерных форматов.
2. Методы, основанные на избыточности аудио и визуальной информации.

Текстовая стеганография относится к первому направлению. Это методы использования известного смещения слов, предложений, абзацев, имитирующих функции (частным случаем этого метода является акrostих), специальных свойств полей форматов, неотображаемых на экране.

Один из этих методов был исследован и реализован в данной работе.

Методы использования известного смещения слов, предложений, абзацев

Методы основаны на изменении положения строк и расстановки слов в предложении, что обеспечивается вставкой дополнительных пробелов между словами.

Этот класс методов является самым старым, он появился еще в "довиртуальную" эпоху и был просто перенесен на компьютеры с учетом их особенностей и возможностей. Данной категории принадлежит очень много алгоритмов. К полностью автоматическим относится, например, форматирование текста количеством пробелов отличным от единицы. Скажем, один пробел соответствует биту "0", а два - "1". Программа получает произвольный текст в качестве контейнера и внедряет в него какое-либо стегосообщение.

Алгоритм встраивания информации в текстовый файл

Программа переводит в двоичный код текст встраиваемого сообщения и изменяет количество пробелов между словами в тексте файла-контейнера. При этом «1»-обозначается двумя пробелами, «0»- одним пробелом.

После того как прочитан файл – сообщение и создан массив нулей и единиц, файл удаляется, далее происходит кодирование и запись файла-результата, либо уже существующего, либо созданного. Из файла-контейнера читается слово и к нему присоединяется один или два пробела, в соответствии с массивом нулей и единиц, и получившаяся строка записывается в файл – результат. Когда массив закончен, записывается знак табуляции, а затем слова записываются в файл – результат через один пробел. После окончания записи изменяется режим доступа к файлу - результату: только для чтения.

При раскодировании проверяется наличие в тексте файла знака табуляции, означающего, что файл, возможно, содержит закодированную информацию. Далее в тексте подсчитывается количество пробелов, после каждого слова, пока не встретится знак табуляции, при этом записывается массив нулей и единиц. После происходит обратное преобразование двоичного кода в символы и запись файла-сообщения.

Литература

4. Грибунин В.Г., Оков И.Н., Туринцев И.В. Цифровая стеганография. – М.: СОЛОН-Пресс, 2002.
5. Bender W., Gruhl D., Morimoto N., Lu A. Techniques for Data Hiding // IBM System Journal. 1996.
6. Moskowitz I.O., Longdon G.E. A new paradigm hidden in Steganography // Proceedings of Workshop "New Security Paradigms". ACM Press. 2000.

КОДИРОВАНИЕ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

Волокитин Д.В.

Научный руководитель: к.т.н., доцент Волосатова Тамара Михайловна
МГТУ им. Н.Э.Баумана кафедра РК6, Москва, Россия

CODING OF DIGITAL IMAGES

Volokitin D.V.

Scientific chief: Ph.D., senior lecturer Volosatova Tamara Mikhaylovna
MSTU n.a. Bauman PK6, Moscow, Russia
VolokitinDV@yandex.ru

Аннотация

В данной работе рассмотрена проблема создания стеганографической системы (стегосистемы) для встраивания текстовой информации в изображение. Стеганография - это метод организации связи, который собственно скрывает само наличие связи.

Abstract

Under consideration is the problem of creating of stegosystem for embedding a text information in digital image. Steganography is a method of relation's obtaining which hide presence of relations.

Слово "стеганография" в переводе с греческого буквально означает "тайнопись" (steganos - секрет, тайна; graphy - запись). К ней относится огромное множество секретных средств связи, таких как невидимые чернила, микрофотоснимки, условное расположение знаков, тайные каналы и средства связи на плавающих частотах и т. д.

Стеганография занимает свою нишу в обеспечении безопасности: она не заменяет, а дополняет криптографию. Соккрытие сообщения методами стеганографии значительно снижает вероятность обнаружения самого факта передачи сообщения. А если это сообщение к тому же зашифровано, то оно имеет еще один, дополнительный, уровень защиты.

В настоящее время в связи с бурным развитием вычислительной техники и новых каналов передачи информации появились новые стеганографические методы, в основе которых лежат особенности представления информации в компьютерных файлах, вычислительных сетях и т. п. Это дает возможность говорить о становлении нового направления - компьютерной стеганографии.

Несмотря на то, что стеганография как способ сокрытия секретных данных известна уже на протяжении тысячелетий, компьютерная стеганография - молодое и развивающееся направление.

Как и любое новое направление, компьютерная стеганография, несмотря на большое количество открытых публикаций и ежегодные конференции, долгое время не имела единой терминологии.

Позднее, на конференции Information Hiding: First Information Workshop в 1996 году было предложено использовать единую терминологию и обговорены основные термины.

Стеганографическая система или стегосистема - совокупность средств и методов, которые используются для формирования скрытого канала передачи данных.

В качестве данных может использоваться любая информация: текст, сообщение, изображение и т. п.

В общем же случае целесообразно использовать слово "сообщение", так как сообщением может быть как текст или изображение, так и, например, аудиоданные. Далее для обозначения скрываемой информации, будем использовать именно термин сообщение.

Контейнер - любая информация, предназначенная для сокрытия тайных сообщений.

Пустой контейнер - контейнер без встроенного сообщения; заполненный контейнер или стего - контейнер, содержащий встроенную информацию.

Встроенное (скрытое) сообщение - сообщение, встраиваемое в контейнер.

Стеганографический канал или просто стегоканал - канал передачи стего.

В данной работе рассмотрен алгоритм[2,3] и разработана программа встраивания информации в изображение. Алгоритм основан на манипулировании яркостью $l(x,y)$ за счет канала синего цвета $b(x,y)$. В зависимости от бита S_i , подаваемой на вход встраиваемой информации, модифицируется значение яркости канала синего цвета по следующему алгоритму:

1. Вычисляется яркость пикселя
 $l(x,y) = 0.299r(x,y) + 0.587g(x,y) + 0.114b(x,y)$, где x, y – координаты пикселя.
2. Каждый пиксель модифицируется в соответствии со схемой

$$b'(x,y) = \begin{cases} b(x,y) + 0.01 * l(x,y), & \text{if } S_i = 1 \\ b(x,y) - 0.01 * l(x,y), & \text{if } S_i = 0, \end{cases}$$

где $b'(x,y)$ – модифицированная составляющая пикселя.

3. Декодирование информации происходит путем сравнения исходного изображения со стегоизображением.

Разработанная программа поддерживает работу с 24-битными изображениями в формате BMP. Основным отличием файлов этого формата (DIB) - Device Independent Bitmap (аппаратно-независимый битовый образ) является то, что в них используется кодировка цветов с одной битовой плоскостью. Файлы битовых образов нового формата начинаются со структуры BITMAPFILEHEADER:

```
struct BITMAPFILEHEADER {
    word  bfType;           //тип файла (для битового образа - BM)
    dword bfSize;          //размер файла в dword
    word  bfReserved1;     //не используется
    word  bfReserved2;     //не используется
    dword bfOffbits;       //смещение данных битового образа от заголовка в байтах
}
```

Непосредственно за ней располагается структура BITMAPINFO, со держащая всю информацию о битовом образе. Она делится на две части: структуру BITMAPINFOHEADER, описывающей размеры и цветовой формат битового образа, и массив структур GBQUAD, определяющей цветовую палитру:

```
struct BITMAPINFO {
    BITMAPINFOHEADER  bmiHeader;
    RGBQUAD           bmiColors[1];
}
struct BITMAPINFOHEADER {
    dword biSize;           //число байт, занимаемых структурой
                           //BITMAPINFOHEADER
    dword biWidth;         //ширина битового образа в пикселях
    dword biHeight;        //высота битового образа в пикселях
    word  biPlanes;         //число битовых плоскостей устройства
    word  biBitCount;       //число битов на пиксель
    dword biCompression;   //тип сжатия
    dword biSizeImage;     //размер картинка в байтах
    dword biXPelsPerMeter; //горизонтальное разрешение устройства, пиксел/м
    dword biYPelPerMeter;  //вертикальное разрешение устройства, пиксел/м
    dword biClrUsed;       //число используемых цветов
    dword biClrImportant; //число "важных" цветов
} BITMAPINFOHEADER;
```

Более подробно:

$biSize$ - обычно используется для облегчения доступа к таблице цветов.

biPlanes - определяет число битовых плоскостей; однако, по скольку цвет кодируется последовательными битами, это число всег да равно 1.

biBitCount - этим полем определяется число цветов, используемых битовым образом. В зависимости от способа кодирования, может принимать значения:

1 - битовый образ монохромный, и таблица цветов должна содержать два элемента. Каждый бит в массиве данных кодирует один пиксел. Если значение бита - 0, то пиксел становится первым цветом таблицы; если - 1, пиксел становится вторым цветом таблицы.

4 - битовый образ имеет максимум 16 цветов, и массив bmiColors (таблица цветов) имеет до 16 элементов. Цвет каждого пиксела определяется по таблице цветов при помощи четырехбитного индекса. Например, если первый байт данных имеет значение 3Ah, то при отображении битового образа цвет первого пиксела определяет четвертый элемент таблицы цветов, а цвет второго - одиннадцатый.

8 - битовый образ имеет максимум 256 цветов, и таблица цветов имеет до 256 элементов. Каждый байт массива данных определяет цвет одного пиксела.

24 - битовый образ имеет максимум 2 в 24-й степени цветов. Таблица цветов пуста, а цвет пикселов определяется пакетами из трех байтов, описывающими цветовые интенсивности красного, зеленого и голубого цветов.

biCompression - тип сжатия. Может принимать значения:

BI_RGB - сжатие отсутствует;

BI_RLE8 - сжатие для формата 8 бит на пиксел;

BI_RLE4 - сжатие для формата 4 бита на пиксел.

biXPelsPerMeter и biYPelsPerMeter - могут использоваться для выбора из списка ресурсов пиктограммы, наиболее подходящей для данного устройства.

biClrUsed - число цветов, используемых данным битовым образом. Если 0, то используются все цвета палитры (указанные в массиве bmiColors).

biClrImportant - используется для управления алгоритмом отображения цветов. Так, если четыре различных приложения отображают на экране по одному битовому образу с 75 цветами каждый, то адаптер, выводящий 256 цветов одновременно, не сможет полностью аутентично отобразить на экране все 4 картинки. В этом случае используется механизм замены цветов - ищется битовый образ с наименьшим приоритетом и его "лишние" цвета заменяются наиболее подходящими.

```
struct RGBQUAD {
    byte  rgbRed;      //интенсивность красного
    byte  rgbGreen;    //интенсивность зеленого
    byte  rgbBlue;     //интенсивность голубого
    byte  rgbRerved;   //не используется
} RGBQUAD;
```

После того, как все параметры битового образа определены, в файле идут сами скан-строки битовой плоскости, причем первой скан-строкой в формате DIB считается нижняя скан-строка (т.е. начало координат находится в левом нижнем углу изображения).

Литература

1. В.Г. Грибунин, Цифровая стеганография, М., 2002 г.
2. Martin Kutter, Frdric Jordan, and Frank Bossen. Digital signature of color images using amplitude modulation. In Ishwar K. Sethi and Ramesh C. Jain, editors, Proceedings of the SPIE Storage and Retrieval for Image and Video Databases V, volume 3022, pages 518-526, San Jose, California, 1997.
3. Martin Kutter, Frdric Jordan, and Frank Bossen. Digital watermarking of color images using amplitude modulation. 3022(2):326-332, April 1997.

ЛОГАРИФМИЧЕСКИЕ МЕТОДЫ КОДИРОВАНИЯ ЗВУКА

Соболева Ю.М.

научный руководитель: к.т.н., доцент Волосатова Т.М.

МГТУ им. Н.Э.Баумана, кафедра РК6, Москва, Россия.

LOGARITHMIC METHODS OF SOUND CODING

Soboleva J.M.

The scientific chief: Ph.D., senior lecturer Volosatova T.M.

BMSTU, RK6, Moscow, Russia.

Аннотация

В работе рассмотрены логарифмические методы кодирования звука, а именно μ -Law и A -Law.

Программная реализация позволяет кодировать звуковой сигнал и сравнивать методы между собой и с линейной ИКМ.

Abstract

In the document methods of sound coding (μ -Law and A -Law) is considered. The program realization allows encoding the audio signal and comparing μ -Law with A -Law and also with PCM.

Введение

Не смотря на то, что общепринято различать «сжатые» и «несжатые» звуковые данные, уместнее говорить просто о различных способах «кодирования» звука. Самый простой способ кодировки – это ИКМ. Даже этот простой формат сложнее, чем может показаться, и существует, по крайней мере, четыре формата ИКМ, широко используемые для записи файлов. С помощью алгоритмов компрессии звука ИКМ преобразуют в более компактный формат.

Линейные форматы записи звука, в некотором смысле, являются примитивными. Слух человека устроен так, что мы практически нечувствительны к небольшим искажениям громких звуков, зато очень хорошо замечаем те же искажения в более тихих звуках. Однако при использовании ИКМ уровень погрешности не зависит от величины амплитуды. Один из способов сохранения самых тихих звуков заключается в переопределении уровней с помощью использования нелинейных систем кодирования вместо стандартной линейной ИКМ. Суть данного метода заключается в более эффективном использовании того же количества битов. В результате получается возможность использования большего количества битов для слабых звуков, при записи которых потеря данных более заметна. Эту технологию можно рассматривать как метод компрессии.

Логарифмическая компрессия

У двух наиболее широко используемых нелинейных форматов общих черт больше, чем различий. В методах компрессии как μ -типа (μ -Law), так и A -типа (A -Law), для преобразования линейных отсчётов ИКМ в 8-битные коды применяются логарифмические формулы.

Успех данного подхода обусловлен несколькими причинами. Во-первых, при модуляции звука большей частью получаются небольшие величины. Использование логарифмической компрессии позволяет повысить точность записи выборок небольшой величины за счёт снижения качества записи больших величин, которые встречаются редко. Во-вторых, чувствительность слуха человека также логарифмическая; изменение интенсивности слабого звука заметно сильнее, чем аналогичное изменение силы мощного звука. По существу,

использование логарифмических методов компрессии обеспечивает большую точность именно там, где она необходима для слухового восприятия.

Разработанная программа реализует распространённые схемы логарифмической компрессии. Первоначально эти схемы были разработаны для использования в телефонных системах, а позднее были включены в стандарт ITU G.711.

Логарифм числа можно рассматривать как количество цифр в этом числе. В связи с этим отправной точкой для логарифмических форматов является таблица, в которой для каждого из 8-битных значений указано, сколько битов потребуется для записи этого значения.

Компрессия μ -Law

Так называемая кодировка мю-типа (иногда пишется μ -Law или mu -Law) наиболее широкое распространение получила в Соединённых Штатах и Японии. Чтобы не усложнять записи формул, будем считать, что все моментальные значения представляют собой дробные числа в диапазоне от -1 до +1.

При работе с 16-битными отсчётами необходимо перед применением формул поделить их на 32768.

Допустим у нас есть выборка s . Тогда соответствующий ей отсчёт мю-типа sm вычисляется по формуле:

$$sm = \text{sign}(s) \cdot \frac{\log(1 + 255 |s|)}{\log(1 + 255)}.$$

Эту формулу проще запомнить, если рассматривать её только для положительных значений. В таком случае она принимает вид

$$sm = \frac{\log(1 + 255s)}{\log(1 + 255)}.$$

Для преобразования выборок мю-типа обратно в линейные отсчёты достаточно воспользоваться обратной формулой:

$$s = \frac{256sm - 1}{255}.$$

Это верно лишь для положительных чисел. При работе с отрицательными величинами придётся пользоваться абсолютными значениями и менять знак у полученного в результате преобразования числа.

Рассмотренные формулы напрямую в программе не применяются, так как вычисление логарифма и возведение числа в степень требуют много машинного времени. Затраты времени удалось избежать использованием заранее просчитанной таблицы преобразования. Однако изначально кодирование мю-типа разрабатывалось для аппаратной реализации, а в таком случае справочные таблицы не эффективны. По этой причине в стандарте ITU G.711 описывается особая аппроксимация, которая может быть легко вычислена аппаратно. Результаты, получаемые с помощью этой аппроксимации, очень близки к данным, получаемым с помощью формул. Хотя различий достаточно, чтобы не смешивать два эти способа.

Для аппроксимации логарифма в стандарте G.711 используются простые битовые манипуляции. Производится начальная коррекция, затем вычисляется количество битов, которые требуются для записи старшего байта. Полученное значение комбинируется с четырьмя старшими битами.

Аналогичным образом работает и декодер. Старший бит используется для записи знака, три следующих – для определения, насколько надо сдвинуть 4 младших бита.

Данный вариант записи очень сильно напоминает формат записи 8-битного числа с плавающей точкой: один бит – для знака, три – для экспоненты и четыре – для записи мантииссы. Подобно большинству форматов с плавающей точкой старший бит мантииссы

отбрасывается, и при декодировании приходится восстанавливать его, добавляя 16 к мантиссе.

При кодировании данным методом производится смещение 16-битного отсчёта ИКМ на 128, а затем на 4 перед остальным преобразованием. При декодировании эти значения аналогичным образом вычитаются после преобразования. Если рассматривать такие преобразования как аппроксимацию логарифма, то число 128 используется для обхода разрывности основной функции. Если посмотреть на данный алгоритм внимательнее, нетрудно заметить, что используются вообще только 13 старших бит. Добавление четырёх округляет значение до ближайшего, кратного восьми, повышая тем самым точность.

Компрессия A -Law

По своей концепции кодирование A -типа аналогично кодированию μ -типа. Как и μ -компрессия, первоначально этот метод разрабатывался для телефонных систем. A -компрессия получила широкое распространение в Европе. Подобно μ -компрессии, этот метод кодирования описывается как непрерывной функцией, так и цифровой аппроксимацией.

На самом деле A -компрессия определяет набор кодировок, зависящий от точного значения числового коэффициента A . Чаще всего A даёт значение 87.6. Пусть s - выборка, лежащая в диапазоне от 0 до 1. Тогда значение A -типа sA вычисляется по формуле:

$$sA = \begin{cases} s \cdot \frac{A}{1 + \ln A}, & s \leq \frac{1}{A} \\ \frac{1 + \ln(A \cdot s)}{1 + \ln A}, & \frac{1}{A} \leq s \leq 1 \end{cases}.$$

Цифровая аппроксимация аналогична аппроксимации, используемой в μ -кодировке. Основное отличие заключается в том, что в выходном коде кодировки A -типа биты инвертируются через один.

Свойства логарифмического кодирования

Чтобы понять, что происходит при логарифмическом кодировании достаточно рассмотреть, как ведёт себя отношение сигнал/шум.

В случае линейной ИКМ берётся некоторый аналоговый сигнал, и каждое идеально точное моментальное значение округляется до ближайшего целого числа. Таким образом, разница между идеальной выборкой и значением ИКМ никогда не превышает $\frac{1}{2}$ величины младшего разряда. В случае логарифмического кодирования ситуация усложняется: для больших величин выборок ошибка увеличится, для небольших величин – уменьшится.

Такой меняющийся уровень ошибки влияет на её общую мощность. Если исходный звук становится громче, то величина отсчётов возрастает и, соответственно, возрастает ошибка. То есть при увеличении мощности сигнала возрастает мощность ошибки. В случае же строго логарифмической кодировки отношение сигнал/шум практически постоянно.

Литература

7. Кинтцель Т. Руководство программиста по работе со звуком. – М.: ДМК - Пресс, 2000.
8. Гордеев О.В. Программирование звука в Windows. – СПб.: БХВ – Санкт-Петербург, 2000.

ПОСТРОЕНИЕ МОДУЛЯ МОНИТОРИНГА В ПРОГРАММНОЙ СИСТЕМЕ НА ОСНОВЕ АКТИВНОЙ БАЗЫ ДАННЫХ

Костра В.В.

Научный руководитель: д.т.н., проф. Алпатов А.П.

ИТМ НАН Украины, Днепрпетровск, Украина

MODULE OF MONITORING CONSTRUCTION IN PROGRAM SYSTEM ON BASIS OF ACTIVE DATABASE

Kostr V.V.

prof. Alpatov A.P.

ITM NAS of Ukraine, Dniepropetrovsk, Ukraine

<http://cyberlab.iatp.org.ua>, vladko@rdc.dp.ua

Аннотация

Описывается информационная технология создания модуля мониторинга на основе активной базы данных, для наблюдения за объектами и процессами заданной предметной области. Использование модуля в составе медицинской системы, обеспечивает новые возможности автоматического анализа данных и выработки оперативной информации врачу о состоянии пациента.

Abstract

The information technology of creation of the module of monitoring on basis of active database is described. It is necessary for observation over objects and processes in given subject domain. Use of the module as part of medical system, provides new feature of the data automatic analysis and output information to the doctor about patient state.

ВВЕДЕНИЕ

В настоящее время технологии создания активной базы данных (АБД), относятся к перспективным направлениям при разработке новых возможностей СУБД [1,2]. В рассматриваемом подходе АБД создается для решения задачи оперативного анализа параметров в структурированных медицинских документах, которые сохраняются в компьютерной медицинской карте пациента (КМК) [3].

На предыдущем этапе системного моделирования компонентов медицинской информационной систем (МИС) [4,5], были реализованы информационные технологии обработки параметров в формализованном медицинском документе [6], в процессе его составления для записи в КМК пациента. При этом сами алгоритмы обработки хранятся отдельно и открыты для доступа [6].

Задачи *модуля мониторинга* (ММ) : контролировать изменение (добавление) параметров в КМК, по возможности проводить их интерпретацию, сообщить результаты интерпретации заинтересованному пользователю (по запросу или в автоматическом режиме).

При практической реализации предложенного подхода в МИС модуль состоит из двух частей:

- интерпретатор - правил, который реализуется средствами серверной части МИС, самодостаточен и функционирует независимо от работоспособности клиентских частей;
- интерпретатор - сообщений, который подключается к клиентской части МИС, гарантируя своевременное оповещение пользователя о результатах интерпретации.

Особенностью работы ММ является использование возможностей созданной АБД, которая позволяет выполнять семантические процедуры обработки параметров в БД, заложенных в саму БД, без явных или косвенных указаний пользователя.

ВОЗМОЖНОСТИ АКТИВНОЙ БАЗЫ ДАННЫХ

Информационная ситуация, отражающая состояние объектов наблюдения (пациентов) в МИС, с семиотической точки зрения представляет собой систему отображающих и выражающих ее параметров, находящихся в различных отношениях. В соответствии с основными положениями семиотики каждый параметр можно рассматривать в трех независимых, но взаимосвязанных аспектах: *имя*, *значение*, *смысл* [7]. Первые два аспекта можно реализовать с помощью информационных технологий на основе современных реляционных СУБД: *значение* - сохраняется в таблице параметров, *имя* – описывается элементами метаинформации [4,5], которые описывают параметр. *Смысл* – представляет собой результат процесса понимания параметра, выражающего заложенные в параметр логико-понятийные и образные сущности, и существуют определенные трудности его представления на уровне БД.

При построении АБД, необходимо обеспечить, чтобы СУБД позволила помещать в БД не только "пассивные" факты, относящиеся к предметной области, но и "активные" правила, которые при возникновении указанных ситуаций принуждают систему совершать специфицированные действия. Фактически, набор правил АБД описывает семантику предметной области.

Структуру "активного" правила представим следующей комбинацией: "*событие* – *условие* – *действие*". *Событие* - фиксируется в результате выполнения некоторой стандартной операции в БД. Примерами события в реляционной БД могут быть: обновление некоторого поля некоторой таблицы, удаление записи из таблицы, занесение в таблицу новой записи и т.д. *Условие* - это произвольное логическое выражение, включающее предикаты, определенные над БД. *Действие* - это произвольный оператор или набор операторов, допускаемых СУБД, а также выполнение специальных функций в программном обеспечении МИС. Выполнение этого действия может явиться событием для некоторого другого правила и т.д. Очевидно, что система, поддерживающая работу АБД, должна обладать некоторой встроенной продукционной системой. Результат выполнения действий, при необходимости может передаваться лицу, принимающему решение, тем самым, реализуя между ним и АБД семантический диалог.

Средствами реализации указанных принципов могут являться: ограничения и утверждения, хранимые процедуры, триггеры, функции определяемые пользователем. Эти, и подобные им технические возможности, имеющиеся в современных реляционных СУБД, позволяют реализовывать различные общие модели АБД.

РАБОТА МОДУЛЯ МОНИТОРИНГА

Общую схему принципа работы модуля в МИС представим следующим примером (диаграмма на языке UML, рис. 1). Пользователь системы или прикладная программа (например, считывания данных с медицинского прибора) выполняют некоторое действие над БД. Например, врач-лаборант добавил в КМК пациента параметры теста "Общий анализ крови". Интерпретатор-правил должен сравнить его значения с известными значениями нормы, которые хранятся в шаблоне теста. При отклонении определенного параметра от пределов нормы этот факт сохраняется в БД, и лечащему врачу, может быть выдано соответствующее сообщение (или возможен вариант добавления в КМК специальной записи с этим сообщением).

Интерпретатор-правил, строится на основе триггеров и сохраненных процедур, которые должны поддерживаться СУБД (ORACLE, SQL Server, InterBase и др.). Хранимые процедуры могут быть активизированы приложениями, присоединенными к БД, что позволяет осуществлять модульную разработку БД, обеспечивает легкость её сопровождения и повторного использования разработанных ранее процедур. В реализованном подходе хранения исходных данных, все записи всех пациентов хранятся в одной таблице – doc [8].

Для этой таблицы пишется триггер, который инициируется, при выполнении транзакции на добавление новой записи в БД.

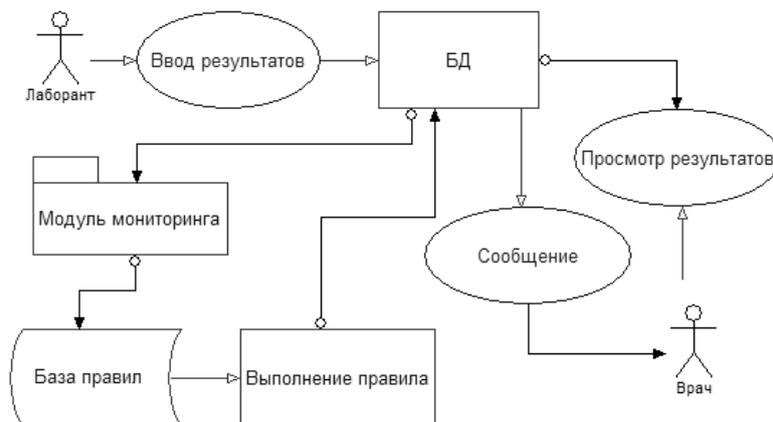


Рис.1 Диаграмма состояний и взаимодействий модуля

Пример записи триггера для СУБД SQL Server:

```

CREATE TRIGGER tr_for_insert ON doc
FOR INSERT
AS
DECLARE @code VARCHAR(4)
DECLARE @params VARCHAR(50)
DECLARE @typz VARCHAR(3)
SELECT @code=code,@params=txtz, @typz=typz FROM INSERTED
EXECUTE obrab_blood @code @params @typz
  
```

При добавлении новой записи в БД создается временная таблица inserted (в автоматическом режиме средствами SQL Server), со структурой аналогичной таблице doc. Из нее считываются данные, добавляемой в таблицу doc записи, которые передаются сохраненной процедуре obrab_blood. Пример сохраненной процедуры:

```

CREATE PROCEDURE obrab_blood
@code INT,
@params VARCHAR(50),
@typz VARCHAR(3) AS
DECLARE @param int
DECLARE @i INT
DECLARE @norm_v INT
SET @i=0;
If @code='0005' AND @typz='Isl'
BEGIN
  Param=Expract_param(3) --процедура извлечения необходимого параметра
  SELECT @norm_v=norm_v FROM Isl0005 WHERE CODE=3 -- верхняя граница нормы
  IF param > norm_v BEGIN
    INSERT INTO warnings VALUES('Количество лейкоцитов больше нормы')
  END
END
END
  
```

В процедуре анализируется добавляемая запись, проверяется значение параметра (количество лейкоцитов), если оно больше нормы, то в таблицу warnings записывается интерпретируемое сообщение. Интерпретатор-сообщений обрабатывает записи в этой таблице, и отображает эту информацию врачу, при просмотре КМК пациента. Для передачи автоматического сообщения врачу (если он в этот момент работает с БД), необходимо генерировать сообщение функцией RAISERROR.

С помощью предложенной технологии можно реализовать различные алгоритмы обработки первичных параметров в медицинских документах (которые вводятся в КМК с

различных рабочих мест МИС) непосредственно после составления документа, или анализ параметров из нескольких документов в динамике.

ЗАКЛЮЧЕНИЕ

Описанная информационная технология реализации и использования модуля мониторинга входит в состав МИС "MedDoc", разработанной автором (под руководством д.т.н. Алпатова А.П. и к.ф.-м.н. Прокопчука Ю.А.), которая используется в ряде лечебных учреждений Днепропетровской области [9]. Предложенный подход позволил реализовать в системе новые возможности автоматического анализа параметров и выработки оперативной информации врачу о состоянии пациента при работе с МИС. Разработанные активные правила хранятся на серверной части МИС и доступны создания и редактирования средствами СУБД (SQL Server). Ведется разработка удобного интерфейса пользователя для создания "активных" правил с помощью модуля мониторинга.

Литература

1. Ершов Н.М., Королев Л.Н., Малютина Э.Э. и др. Применение активных баз данных в прогнозировании. // Вестн.Моск.ун-та. Сер.15, Вычисл.матем. и киберн. 1998, №1 С.32-43
2. С.Д. Кузнецов Основы современных баз данных // www.citforum.ru
3. Костра В.В. Информационная модель компьютерной медицинской карты пациента // Всероссийская научн.-технич. конф. молодых ученых и специалистов "Биомедсистемы-2003":Тезисы докладов.- Рязань,2003.-С.130-131
4. Прокопчук Ю.А., Костра В.В. Средства интеллектуализации открытых информационных систем // Докл. междунар. конф. "Информационные средства и технологии" (Москва, 16-18 октября 2001 г.) в 3-х томах. –М.: Изд-во "Станкин, 2001, Т.2. С.94 – 97
5. Костра В.В., Прокопчук Ю.А. Модульная технология построения информационных систем // III Молод. науч.-техн. конф. "Наукоемкие технологии и интеллектуальные системы 2001". Сборник докладов.-21-22 марта 2001 г., г. Москва, МГТУ им.Н.Э.Баумана, С.146-149
6. Костра В.В., Прокопчук Ю.А. Принципы построения семантического процессора для медицинских информационных систем // IV Молод. научно-технич. конф. "Наукоемкие технологии и интеллектуальные системы 2002". 17-18 апреля 2002 г., г. Москва, МГТУ им.Н.Э.Баумана, С.28-31
7. Перспективы развития вычислительной техники: В 11 кн.: Справ. пособие/ Под.ред. Ю.М.Смирнова. Кн.1: Информационные семантические системы/ Н.М.Соломатин.- М.: Высш.шк., 1989. – 126 с.
8. Ю. А. Прокопчук, В. В. Костра Анализ моделей данных в госпитальных системах // Информационные технологии и кибернетика на службе здравоохранения: Сборник докладов межд. научно-практ. конференции. – Днепропетровск: ИПК ИнКомЦентра УГХТУ, 2003.- С.92-96
9. Прокопчук Ю.А., Харченко О.А., Костра В.В. Госпитальные информационные системы: опыт разработки и эксплуатации // Сборник тезисов IV специализированной выставки и конференции "Информационные технологии в медицине – 2003", Москва, 2003, - С.76-78

ПАРАЛЛЕЛЬНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМА БПФ

Лукьянец С.В.

Научный руководитель: к.т.н., доцент Волосатова Т.М.
МГТУ им. Н.Э. Баумана, кафедра РК6

PARALLEL PROCESSING OF ALGORITHM FFT

Lukynets S.V.

Scientific chief: Ph.D., senior lecturer Volosatova T.M.
BMSTU, PK6

Аннотация

Рассмотрена задача построения параллельного алгоритма быстрого преобразования Фурье. Описаны проблемы и способы их решения для архитектур с общей и распределенной памятью.

Abstract

FFT algorithm development task has been reviewed. Problems and solutions to them for shared and distributed memory architectures have been described.

Алгоритм быстрого преобразования Фурье (БПФ) - это оптимизированный по скорости способ вычисления ДПФ.

Фундаментальное уравнение для получения N-точечного ДПФ выглядит следующим образом:

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) \cdot e^{-j2\pi nk / N} = \frac{1}{N} \sum_{n=0}^{N-1} x(n) [\cos(2\pi nk / N) - j \cdot \sin(2\pi nk / N)]$$

По отношению к этому уравнению следует сделать некоторые терминологические разъяснения. $X(k)$ (прописная буква X) представляет собой частотный выход ДПФ в k-ой точке спектра, где k находится в диапазоне от 0 до N-1. N представляет собой число отсчетов при вычислении ДПФ.

Значение $x(n)$ (строчная буква x) представляет собой n-ый отсчет во временной области, где n также находится в диапазоне от 0 до N-1. В общем уравнении $x(n)$ может быть вещественным или комплексным.

Основная идея БПФ заключается в двух пунктах.

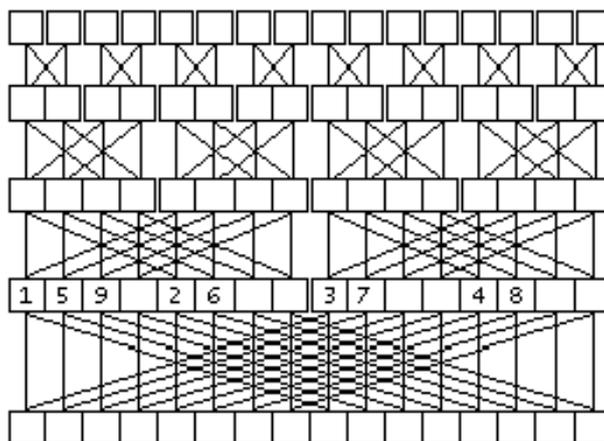
1. Необходимо разделить сумму из N слагаемых на две суммы по N/2 слагаемых, и вычислить их по отдельности. Для вычисления каждой из подсумм, надо их тоже разделить на две и т.д.
2. Необходимо повторно использовать уже вычисленные слагаемые.

Применяют либо "прореживание по времени" (когда в первую сумму попадают слагаемые с четными номерами, а во вторую - с нечетными), либо "прореживание по частоте" (когда в первую сумму попадают первые N/2 слагаемых, а во вторую - остальные). Оба варианта равноценны. В силу специфики алгоритма приходится применять только N, являющиеся степенями 2. Оба варианта являются абсолютно равносильными и дают совершенно одинаковый результат, поэтому будем рассматривать случай прореживания по времени.

Пусть имеется $N=2^T$ элементов последовательности $x_{\{N\}}$ и надо получить последовательность $X_{\{N\}}$. Прежде всего, нам придется разделить $x_{\{N\}}$ на две последовательности: четные и нечетные элементы. Затем точно так же поступить с каждой последовательностью. Этот итерационный процесс закончится, когда останутся последовательности длиной по 2 элемента.

В этом заключается первый этап БПФ, который называется предварительной перестановкой. По сравнению с последующим этапом он занимает мало времени, поэтому способы его параллельной реализации не рассматриваются.

На рисунке проиллюстрирован второй этап вычисления ДПФ. Линиями сверху вниз показано использование элементов для вычисления значений новых элементов. Удобство алгоритма состоит в том, что два элемента на определенных позициях в исходном массиве дают два элемента на тех же местах в преобразованном. Это позволяет обойтись одним массивом данных для исходных данных, результата и хранения промежуточных результатов для всех T итераций.



Именно так, а не иначе: сначала вычисляются элементы с номерами 0 и $N/2$, во всех ДПФ в количестве N_{max}/N штук, потом с номерами 1 и $N/2 + 1$ и так далее. На рис.4 показана последовательность вычислений для $N = 8$. Такая последовательность обеспечивает однократное вычисление W_N^k .

Величина $W_N^k = e^{\frac{-j2\pi k}{N}}$ называется поворачивающим коэффициентом.

Как видно из рисунка выше основная идея распараллеливания БПФ заключается в том, что на каждом процессоре обрабатывается свое подмножество точек.

Наибольшая эффективность в реализации вычислений и обмена достигается в том случае, если число процессоров равно 2^m , это связано со спецификой линейного алгоритма БПФ. Так же не рационально использовать число процессоров, превышающее $N/2$, так как на одном шаге вычислений задействованы сразу 2 элемента массива и для них используется один поворачивающий коэффициент, поэтому 1 процессор за 1 итерацию вычисляет сразу 2 элемента массива.

Для архитектур с общей памятью алгоритм реализован на потоках (threads). В процессе работы каждому потоку отводится своя часть всего массива преобразуемых комплексных чисел. На каждом шаге T ($T=1..log_2N-1$) происходит создание p потоков, каждый из которых в зависимости от своего идентификатора работает с конкретной частью всего массива комплексных чисел. Работа потоков разделена на 2 этапа. На первом этапе происходит независимая работа потоков, так как обрабатываемые части не оказывают влияния на данные в других потоках.

На втором этапе происходит синхронная работа потоков, так как память общая, то каждый поток может обрабатывать $loc_size/2$ данных, где loc_size – это размер блока данных который выделяется определенному потоку. В связи с тем, что каждый поток обрабатывает только половину своих данных и изменяет часть чужих, то при работе на данном этапе необходима синхронизация работы потоков.

Для синхронизации потоков можно использовать несколько вариантов. В работе была реализована синхронизация, в результате которой на каждой итерации вновь создавалась группа потоков для работы с конкретной частью массива, а в конце ожидалось завершение всех потоков группы.

Можно так же организовать синхронизацию с помощью семафоров, но это потребует выделения среди группы потоков 1-го главного, который следил бы за состоянием семафора, а это потребует в свою очередь каким то образом посылать главный процесс в спячку, до конца работы всех потоков, либо задавать потоку пустой цикл с проверкой состояния семафора. Но последний вариант не является хорошим выходом, так как будет затрачиваться процессорное время, которое может быть потрачено гораздо лучше, если на один процессор приходится не один, а несколько потоков, в числе которых и находится главный поток.

Для архитектур с распределенной памятью алгоритм реализован на библиотеке MPI.

Работа алгоритма во многом схожа с многопоточковой реализацией, различия проявляются в способах взаимодействия данных, теперь каждый элемент вычислительной сети работает так же с частью своего массива, предварительно проведя необходимый обмен данными, а в конце вычислений так же передав результат своих вычислений узлу в паре с которым он сейчас работает.

На каждом шаге T после обработки данных осуществляется синхронизация с помощью функции `MPI_Barrier`. А в конце работы производится сбор обработанных данных.

К достоинствам данной реализации можно отнести то, что каждый узел сети не обязан хранить в своей памяти весь массив данных, а только его часть. Но недостатком является то, что такая реализация требует от коммуникативной среды большой пропускной способности, так как в процессе работы происходит постоянный обмен данными между узлами вычислительной среды и никак нельзя допустить больших задержек на передачу данных.

Обратное преобразование практически ничем не отличается, только в процессе вычисления поворачивающего коэффициента необходимо заменить j на $-j$, а в конце поделить все полученные данные на N , распараллеливание чего не составляет никакого труда, так как данные в этом случае абсолютно независимы.

Литература

1. В.В.Воеводин. Информационная структура алгоритмов. М., Изд-во МГУ, 1997 г.
2. В.В.Воеводин, Вл.В.Воеводин. Параллельные вычисления. БХВ-Петербург, 2002 г.
3. MPI – The Complete Reference, The MIT Press, 1996.
4. <http://www.parallel.ru>
5. <http://www.fftw.org>
6. Man pages

РАЗРАБОТКА АЛГОРИТМА ОБОБЩЕНИЯ ИНФОРМАЦИИ, ОСНОВАННОГО НА ТЕОРИИ ПРИБЛИЖЕННЫХ МНОЖЕСТВ

Куликов А.В.

Научный руководитель: д.т.н., проф. Вагин В.Н.

Московский энергетический институт (технический университет)

DEVELOPMENT OF ALGORITHM FOR INFORMATION GENERALIZATION BASED ON THE ROUGH SET THEORY

Kulikov A.V.

Supervisor: doctor of technical sciences, professor Vagin V.N.

Moscow power engineering institute (technical university)

kulikov@apmsun.mpei.ac.ru

Аннотация

Рассматривается проблема обобщения объектов с учетом необходимости обработки неполной и противоречивой информации, хранящейся в реальных массивах данных. Для решения проблемы предлагается использовать средства теории приближенных множеств. Предлагается модификация алгоритмов обобщения, направленная на обработку реальных массивов данных. При этом учитываются такие особенности информации, хранящейся в реальных массивах данных, как большой объем, а также неполнота и противоречивость данных. Работа выполнена при поддержке РФФИ (проект № 02-07-90042).

Abstract

The problem of generalization of objects is considered in view of necessity of processing of the incomplete and inconsistent information stored in real databases. For the decision of this problem it is offered to use means of the rough set theory. Modification of algorithms for the generalization, intended for processing of real databases is offered. Here such features of the information stored in real databases, as great size, and also incompleteness and inconsistency of the data are taken into account. This work is supported by the RFBR (grant No. 02-07-90042).

Необходимость применения методов обобщения информации в интеллектуальных системах обусловлена как построением обобщенных моделей данных, так и обработкой больших массивов экспериментальных данных, которые содержат большое количество скрытых закономерностей. Для выявления таких закономерностей в настоящее время разрабатываются методы обобщения и создаются компьютерные системы, реализующие эти методы. С помощью методов обобщения в системах принятия решений выделяются признаки, характеризующие группу, к которой принадлежит тот или иной объект, и выявляются новые, нетривиальные и полезные знания. Это делается путем анализа уже классифицированных объектов и формирования некоторого набора правил (обобщенной модели). Затем эта обобщенная модель может быть использована для распознавания объектов, не известных системе ранее. Весьма актуальной является задача классификации объектов в условиях избыточной, неполной или противоречивой информации. Цель данного доклада – рассмотреть возможности использования теории приближенных множеств для решения задачи обобщения информации, а также предложить методы, улучшающие работу известных алгоритмов.

Методы теории приближенных множеств предоставляют средства, необходимые для полного исследования множества атрибутов. Алгоритм, реализующий такой подход, строит две системы решающих правил, одна из них дает уверенную классификацию, вторая – возможную, что позволяет обрабатывать неполную исходную информацию, а также работать с неточными и даже противоречивыми исходными данными.

В общем виде работа алгоритма, основанного на теории приближенных множеств, состоит из следующих шагов: поиск классов эквивалентности отношения неразличимости, поиск верхнего и нижнего приближений, поиск среза информационной системы и построение набора решающих правил. Кроме того, для обработки атрибутов с

непрерывной областью значений применяется дискретизация. На этапе дискретизации непрерывная область значений атрибута разбивается на конечное число интервалов. И затем в процессе обобщения значения, лежащие внутри одного интервала, не различаются. Следует отметить, что задачи нахождения оптимального множества таких интервалов (или разграничивающих их делений) и поиска минимального среза NP-сложны, поэтому для их решения необходимы эффективные эвристические алгоритмы. Вычислительная сложность базового и наиболее широко используемого алгоритма дискретизации [1] обладает кубической зависимостью от числа объектов решаемой таблицы, что делает его неприменимым для обработки больших баз данных. Далее мы предлагаем эффективную модификацию, решающую эту проблему.

Разработанный алгоритм направлен на снижение затрат времени и памяти, а также на обработку неполноты и противоречивости данных. Он основан на следующих идеях: а) В качестве существенных признаков рассматриваются признаки, входящие в приблизительные срезы достаточно высокого качества. Идею сокращения множества атрибутов можно обобщить путем введения понятия значимости атрибутов, что позволяет отразить степень важности атрибута в решаемой таблице. б) Предложено одновременное выполнение дискретизации и поиска приблизительного среза. в) Развита идея итерационного вычисления количества пар объектов, различимых делениями непрерывной области значений количественного атрибута. Первоначально данная идея была предложена в [2], однако она применима лишь при наложении ряда ограничений на решаемую таблицу. В предложенной модификации, которая далее будет рассмотрена подробно, эта идея была расширена на случай произвольной решаемой таблицы. г) Для обработки неполноты данных, выраженной в наличии отсутствующих значений атрибутов, реализован метод присвоения всех возможных значений атрибута, относящихся к понятию [3].

Предложенный эвристический алгоритм дискретизации основан на стратегии Джонсона и расширении идеи итерационного вычисления количества пар объектов, различимых делением. Эта идея базируется на наблюдении, что между двумя последовательными делениями существует тесная связь. Так, например, можно заметить, что некоторые пары объектов различимы обоими делениями, а изменения в количестве различимых пар объектов могут быть лишь за счет объектов, значения атрибута которых лежат между двумя этими делениями. В [2] рассматривается ситуация, когда в этот интервал может попасть ровно один объект. Мы обобщаем эту идею на случай произвольного числа таких объектов. Таким образом, наш алгоритм расширяет идею итерационного вычисления количества пар объектов, различимых делением, на общий случай решаемой таблицы. Вычислительная и емкостная сложности предложенного эвристического алгоритма имеют линейную зависимость от количества объектов решаемой таблицы.

Далее в докладе используются терминология и обозначения теории приближенных множеств, введенные в [1]. В частности решаемой таблицей (или решаемой системой) будем называть пару $S = (U, A \cup \{d\})$, где U – непустое конечное множество объектов, называемое обучающим множеством, A – множество условных атрибутов, $d \notin A$ – выделенный атрибут, называемый решением. Кроме того, для изложения предлагаемого алгоритма введем ряд дополнительных обозначений.

Для некоторого деления $p_i^a = (a, c_i^a)$ для атрибута $a \in A$, и некоторого подмножества $X \subseteq U$ будем обозначать:

- $l^X(p_i^a)$ и $r^X(p_i^a)$ - количество объектов из X , имеющих значение атрибута a , меньше (больше) c_i^a ;
- $l_q^X(p_i^a)$ и $r_q^X(p_i^a)$ - количество объектов из X , имеющих значение атрибута a , меньше (больше) c_i^a , и принадлежащих q -му классу решения, где $q = 1, \dots, r(d)$.

Решаемую систему, получаемую в результате дискретизации с использованием множества делений P , обозначим S^P . Изложим две теоремы, которые позволят нам

вычислить значение $W_p(p_{t+1}^a)$, зная $W_p(p_t^a)$, где p_t^a и p_{t+1}^a - два последовательных деления атрибута a , $W_p(p)$ - количество пар объектов, не различимых делениями из P , но различимых делениями из $P \cup \{p\}$.

Теорема 1. Пусть множеству $X \subseteq U$ принадлежит $N^X(p_t^a, p_{t+1}^a)$ объектов, значения атрибута a которых лежат в интервале (c_t^a, c_{t+1}^a) . Тогда:

$$(a) l_q^X(p_{t+1}^a) = l_q^X(p_t^a) + N_q^X(p_t^a, p_{t+1}^a) \quad \forall q = 1, \dots, r(d)$$

$$(b) r_q^X(p_{t+1}^a) = r_q^X(p_t^a) - N_q^X(p_t^a, p_{t+1}^a) \quad \forall q = 1, \dots, r(d)$$

Теорема 2. Пусть имеется K классов эквивалентности отношения неразличимости решающей таблицы $S^P : X_{\alpha_1}, X_{\alpha_2}, \dots, X_{\alpha_K}$, каждому из которых принадлежит $N^{X_{\alpha_i}}(p_t^a, p_{t+1}^a) \geq 1$ объектов, значения атрибута a которых лежат в интервале (c_t^a, c_{t+1}^a) . Тогда

$$W_p(p_{t+1}^a) = W_p(p_t^a) + \sum_{i=1}^K \left[N^{X_{\alpha_i}}(p_t^a, p_{t+1}^a) \cdot (r^{X_{\alpha_i}}(p_t^a) - l^{X_{\alpha_i}}(p_t^a)) - \sum_{q=1}^{r(d)} N_q^{X_{\alpha_i}}(p_t^a, p_{t+1}^a) \cdot (r_q^{X_{\alpha_i}}(p_t^a) - l_q^{X_{\alpha_i}}(p_t^a)) + \sum_{q=1}^{r(d)} \left(N_q^{X_{\alpha_i}}(p_t^a, p_{t+1}^a) \right)^2 - \left(N^{X_{\alpha_i}}(p_t^a, p_{t+1}^a) \right)^2 \right]$$

Теперь представим шаги нашего алгоритма. Назовем его обобщенным итерационным алгоритмом дискретизации (Generalized Iterative algorithm for Discretization, *GID*).

Алгоритм *GID*.

Вход: Непротиворечивая решающая таблица S .

Выход: Субоптимальное множество делений P .

Используемые структуры данных: P – субоптимальное множество делений, $L = [IND(A^P)]$ – множество классов эквивалентности отношения неразличимости таблицы S^P ; A^* - множество возможных делений.

1. $P := \emptyset$; $L := \{U\}$; $A^* :=$ начальное множество делений;
2. Для каждого атрибута $a \in A$ выполнить:

начало

$$W_p(p_0^a) := 0;$$

Для всех $X_i \in L$ выполнить:

$$r^{X_i} := |X_i|; \quad l^{X_i} := 0;$$

$$\text{для } q = 1, \dots, r(d) \text{ присвоить } r_q^{X_i} := |\{x \in X_i : d(x) = v_q^d\}|; \quad l_q^{X_i} := 0;$$

Для каждого деления $p_j^a = (a, c_j^a) \in A^*$ выполнить:

Для всех классов X_{α_i} , объекты которых имеют значение атрибута a из интервала (c_{j-1}^a, c_j^a) , вычислить $N^{X_{\alpha_i}}$ и $N_q^{X_{\alpha_i}}$.

Найти $W_p(p_j^a)$ в соответствии с теоремой 2, т.е.:

$$W_p(p_j^a) = W_p(p_{j-1}^a) + \sum_{i=1}^K \left[N^{X_{\alpha_i}} \cdot (r^{X_{\alpha_i}} - l^{X_{\alpha_i}}) - \sum_{q=1}^{r(d)} N_q^{X_{\alpha_i}} \cdot (r_q^{X_{\alpha_i}} - l_q^{X_{\alpha_i}}) + \sum_{q=1}^{r(d)} \left(N_q^{X_{\alpha_i}} \right)^2 - \left(N^{X_{\alpha_i}} \right)^2 \right].$$

Пересчитать значения $r^{X_{\alpha_i}}$, $l^{X_{\alpha_i}}$ и $r_q^{X_{\alpha_i}}$, $l_q^{X_{\alpha_i}}$ в соответствии с теоремой 1.

конец;

3. Взять в качестве p_{\max} деление с максимальным значением $W_p(p)$ среди всех делений p из A^* .
4. Присвоить $P := P \cup \{p_{\max}\}$; $A^* := A^* \setminus \{p_{\max}\}$;

5. Для всех $X \in L$ выполнить:

Если p_{\max} разделяет множество X на X_1 и X_2 , то удалить X из L и добавить в L два множества X_1 и X_2 .

6. Если все множества в L состоят из объектов, принадлежащих единственному классу решения, то *Конец*, иначе перейти к шагу 2.

7. *Конец*.

Оценим вычислительную сложность предлагаемого алгоритма. Наиболее трудоемкими шагами алгоритма являются второй и пятый.

На шаге 2 в процессе вычисления числа пар объектов, различимых делением $p_j^a = (a, c_j^a)$, изменяются значения $r^{X_{\alpha_i}}, l^{X_{\alpha_i}}, N^{X_{\alpha_i}}$ и $r_q^{X_{\alpha_i}}, l_q^{X_{\alpha_i}}, N_q^{X_{\alpha_i}}$, где $q = 1, \dots, r(d)$. Эти операции выполняются лишь для тех классов эквивалентности X_{α_i} , хотя бы один объект которых удовлетворяет условию принадлежности значения атрибута a интервалу (c_{j-1}^a, c_j^a) . Для одного такого класса эквивалентности будет выполнено $3 \cdot r(d) + 3$ описанных операций. Обозначим это число как α . Оно не зависит от количества объектов n и количества атрибутов k . Число таких классов эквивалентности не может превышать числа n_j объектов, принадлежащих им и значение атрибута a которых лежит в интервале (c_{j-1}^a, c_j^a) . Следовательно, при вычислении $W_p(p_j^a)$ для одного деления p_j^a выполняется не более $\alpha \cdot n_j$ операций. Поэтому в ходе

обработки всех делений одного атрибута будет выполнено $\sum_{j=1}^{n-1} \alpha \cdot n_j \leq \alpha \cdot n$ операций. А для обработки делений всех k атрибутов потребуется $\alpha \cdot kn$ операций. Вторым шагом повторяется $|P|$ раз, а значит, его полная вычислительная сложность составляет $O(|P| \cdot kn)$.

На шаге 5 выполняется разбиение классов эквивалентности. Возьмем худший случай, когда, в конечном счете, в любом классе будет содержаться ровно по одному объекту. Поскольку всего имеется n объектов, то за весь алгоритм будет выполнено $n-1$ операций разбиения. Следовательно, вычислительная сложность пятого шага составляет порядка $O(n)$.

Полная вычислительная сложность предлагаемого алгоритма дискретизации равна $O(|P| \cdot kn) + O(n) = O(|P| \cdot kn)$, что позволяет применять его для обработки больших массивов данных.

Итак, в докладе предлагается модификация алгоритма обобщения, при этом большое внимание уделено процессу поиска наилучшего среза и методике дискретизации непрерывных областей значений атрибутов. Предложенный алгоритм направлен на решение задачи классификации объектов с учетом таких особенностей информации, хранящейся в реальных массивах данных, как большой объем, а также неполнота и противоречивость данных. Проведенные эксперименты показывают, что путем правильного задания уровня аппроксимации в большинстве случаев можно достичь высокого качества классификации новых объектов.

Литература

1. J. Komorowski, Z. Pawlak, L. Polkowski, A. Skowron. Rough Sets: A Tutorial. In Pal, S. and Skowron, A., editors, Rough Fuzzy Hybridization, pp. 3-98. Springer-Verlag, 1999.
2. S.H. Nguyen, H.S. Nguyen. Some efficient algorithms for rough set methods. In: Bouchon-Meunier, Delgado, Verdegay, Vila, and Yager, 1996, pp. 1451-1456.
3. J.W. Grzymala-Busse, M. Hu. A Comparison of Several Approaches to Missing Attribute Values in Data Mining. Rough Sets and Current Trends in Computing 2000, pp. 378-385.

РАЗРАБОТКА СИСТЕМЫ УПРАВЛЕНИЯ ПРЕДПРИЯТИЕМ НА ОСНОВЕ СОВРЕМЕННЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Иванцов И. С.

Научный руководитель: Стучебников В. М.

ЗАО «Мидаус»

ENTERPRISE CONTROL SYSTEM DEVELOPMENT ON THE BASIC OF MODERN INFORMATION SYSTEMS

Ivantsov I. S.

Scientific adviser Stuchebnikov V. M.

Midaus JSC

Аннотация

В работе приводится общая концепция автоматизации современного промышленного предприятия и рассматривается модель АСУП, ее модули и решаемые этими модулями задачи. Описывается реализация системы на примере конкретного предприятия. А также решение проблемы контроля технологического цикла с использованием конечных автоматов.

Abstract

This article contains general concept of modern industrial enterprise automation as well as a basic model of CAM system with description of its modules and their functions. The main part of the work is dedicated to implementation of the mentioned system by the example of existing enterprise. Besides, a solution of a problem of technology circle control with the use of finite state automaton is given in the article.

На нашем предприятии была реализована и успешно функционирует один из модулей системы автоматического управления производством. Модуль предназначен для автоматического сбора данных и расчета параметров выпускаемой продукции. В состав системы входят:

- рабочие станции и специальное программное обеспечение для сбора данных испытаний датчиков и тензомодулей.
- станции обработки и расчета параметров датчиков, ядром которых являются локальные базы данных (передача данных с испытательных участков реализована через файлы).
- центральный сервер СУБД для архивирования данных.

Система сбора данных и расчетов функционирует следующим образом: датчики испытывают на нескольких температурах, полученные результаты сохраняются в выходные файлы. Так как производственный цикл испытаний имеет продолжительность более одного дня, файлов с испытаниями группы датчиков несколько. Далее файлы передаются на компьютер обработки результатов испытаний. Ядро системы расчета - СУБД Paradox. Таким образом, для параллельной обработки оператору требуется переносить данные на разные машины, следить за их непротиворечивостью и целостностью. Наряду с этим в базе данных Paradox имеются проблемы с индексами при больших размерах базы.

Т. е., при том, что данная система позволяет автоматически снимать данные испытаний и рассчитывать параметры датчиков, она имеет ряд ограничений:

- локальные базы данных не обеспечивают одновременную работу нескольких пользователей, что критично при увеличении объема выпускаемой продукции.
- недостаточная надежность локальных баз требует архивации данных.
- локальные базы не позволяют интегрировать производство и базы данных бухгалтерии.

Для устранения вышеупомянутых недостатков отдел автоматизации нашего предприятия решает следующие задачи:

- модернизация модуля автоматического съема параметров выпускаемых датчиков и расчета их характеристик.
- реализация модуля контроля технологического процесса
- интеграция производственной части и бухгалтерии

Наиболее рациональным решением по обеспечению корректного доступа к данным нескольких пользователей является реализация базы данных на ядре MS SQL Server. Созданный сервер значительно упрощает параллельную работу нескольких операторов с более надежными средствами обеспечения целостности данных.

Хранение данных, получаемых с установок системы автоматического съема параметров, решено оставить в файлах, поскольку установки автоматического съема параметров должны работать в независимости от сервера, СУБД или коммуникационного оборудования. Структуру файлов решено изменить по типу COFF формата исполняемых файлов из-за легкости добавления нового блока данных и наличия контрольной суммы.

Для эффективной реализации данной АСУП наш отдел изучил существующие системы управления предприятием и разработал свою концепцию построения этих систем в рамках нашего производства.

На рисунке 1 представлена упрощенная модель промышленного предприятия. Цифрами обозначены зависимости различных систем модели:

- 1 - планирование производства на месяц (квартал, год и т.д.)
- 2 - номенклатура выпускаемых предприятием изделий, предназначенная для использования службой маркетинга в рекламной продукции, каталогах и т.д. Также обновление номенклатуры, в зависимости от потребностей заказчика.
- 3 - номенклатура изделий, используемая в процессе производства
- 4 - сдача на склад готовых изделий, а также периодическая проверка производством изделий со склада.
- 5 - производственный брак, его анализ и, если необходимо, возврат изделия на испытания.
- 6 - брак от потребителя (возвращенное изделие сначала приходит на склад), его анализ и необходимые испытания.
- 7 - реклама, публикации и электронная коммерция.
- 8 - поиск службой маркетинга потенциальных потребителей.
- 9 - планирование закупок комплектующих и оборудования на месяц (квартал, год и т.д.)
- 10 - получение потребителем рекламной продукции, а также заказ изделий через систему электронной коммерции.
- 11 - получение заказанных изделий потребителем со склада и возврат забракованной продукции.



Рис 1

Вообще, АСУП должна выполнять следующие функции:

- Сбор и хранение данных с возможностью их модификации
- Необходимые расчеты
- Возможность публикации необходимой информации
- Функции бухгалтерии и склада
- Грамотное планирование всех процессов на предприятии.

Задача надежного хранения данных с возможностью удобного доступа к ним решена в *базах данных*. На рынке программного обеспечения существует достаточно широкий выбор бухгалтерских пакетов и программ управления складом, которые базируются на распространенных *СУБД*. Большинство современных *internet магазинов* реализованы все на

тех же СУБД. Следовательно, из всех необходимых задач не реализованным остается лишь модуль автоматизации конкретного производства.

В итоге, для реализации АСУП потребуется: приобрести СУБД, удовлетворяющую предприятие по функциональности, надежности и цене, поставить необходимые пакеты бухгалтерии и склада, которые совместимы с выбранной СУБД, и, собственно, разработать клиент - серверное приложение для решения задачи автоматизации производственного процесса и обеспечения передачи необходимых данных остальным модулям системы.

Следующим этапом развития данной системы будет реализация контроля технологического процесса производства и испытаний датчиков (т.е. контроль перехода изделия с одной фазы производственного процесса на другой). Это одна из важнейших функций реализуемой системы.

Очевидно, что весь технологический процесс можно представить в виде графа:

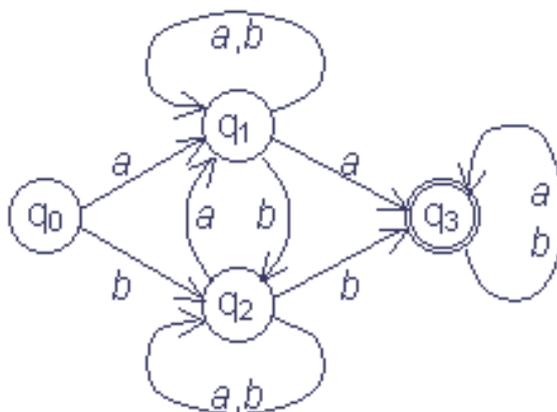


Рис 2

где q_0, \dots, q_n – состояния процесса, а a, b условия переходов. В общем случае это недетерминированный конечный автомат. Любой сложный процесс можно представить в виде графа и построить для него свой конечный автомат.

Формально, *конечным автоматом* называется набор $A = (Q, T, t, q_0, F)$, где

T – *текстовый алфавит*, в нашем случае поле состояний процесса изготовления изделия;

Q – конечное множество *состояний*, $T \cap Q = \emptyset$ (фазы технологического процесса);

$q_0 \in Q$ – *начальное* состояние;

$F \subseteq Q$ – множество *заключительных* состояний;

t – множество *правил перехода*, $t \subseteq Q \times T \times Q$ (условия перехода из одной фазы в другую).

Подобная структура легко реализуется на сервере СУБД (рис. 3) с возможностью контроля переходов на уровне базы данных. Также разумно будет вести историю прохождения технологического процесса изделием. В этом случае можно контролировать каждый этап изготовления продукта, вплоть до электронной подписи ответственного за конкретную операцию лица.

Контроль процесса изготовления – это все же не автоматизация, хотя при наращивании логики приложения на этой же структуре данных реализуется полностью автоматический анализ состояния и переход изделия на следующий этап обработки.

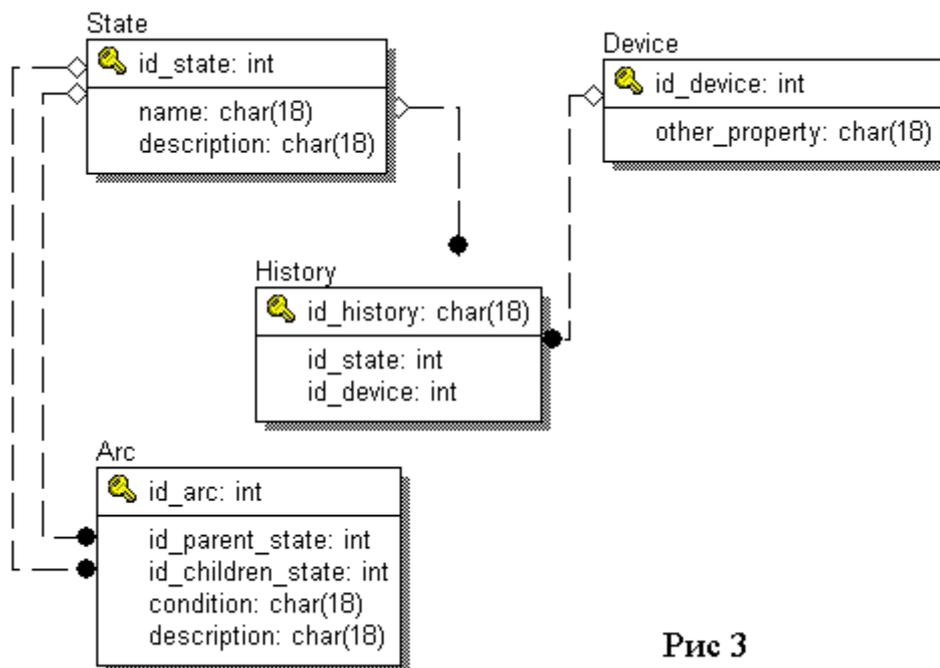


Рис 3

Последним этапом реализации этой системы планируется создание серверного приложения для передачи данных между производственной частью базы и базами склада и бухгалтерии. MS SQL Server предоставляет возможность использования триггеров и распределенных баз данных. Таким образом, решение этой задачи может быть реализовано целиком средствами сервера.

Литература

1. Мамаев, Л. Шкарина, Microsoft SQL Server 2000, “ПИТЕР”, 2001 г.
2. Пушников А.Ю., Введение в системы управления базами данных, <http://www.citforum.ru>
3. С.П. Крицкий, Трансляция языков программирования: синтаксис, семантика, перевод; <http://public.uic.rsu.ru/~skritski/scourses/Transl/>

СИСТЕМА МОДЕЛИРОВАНИЯ ПРОЦЕССА ЛИТЬЯ ВЫЖИМАНИЕМ В СРЕДЕ САПР

аспирант: Носенко А.Г.

Научный руководитель: д.т.н проф. Норенков И.П.

МГТУ им. Н.Э. Баумана кафедра САПР г. Москва, Россия

SYSTEM OF MODELLING OF PROCESS OF MOULDING BY SQUEEZING IN CAD ENVIRONMENT

The post-graduate student: Nosenko A.G.

Scientific assistant: Norenkov I.P.

MSTU name after Bauman.

artem@oct.ru, nosenko@list.ru

Аннотация

Данный доклад дает представление о программном продукте, разработанном для моделирования процесса заполнения жидким расплавом формы, методом литья выжиманием. Приложение написано на Visual C++, под API SolidWorks.

Abstract

This paper proposes a simple about software, developed for modeling process filing liquid alloy, casting by squeezing. This application was developed on Visual C++, with using API SolidWorks.

Работа направлена на технологическое обеспечение высокого качества и ресурса аэрокосмической техники. Пример изделия с такими требованиями: корпусные конструкции летательных аппаратов (обтекатели самолетов) для них требуется высокая надежность и ресурс.

Для решения этой задачи потребовалось выполнить следующие работы:

- 1) Создать геометрическую модель изделия и технологической оснастки.
- 2) Обеспечить поддержку технологической схемы процесса.
- 3) Провести обработку параметров технологического процесса (она проводится на основании имитационных моделей, которые формируют теплогидравлическую картину) основное требование к этим моделям состоит в том, чтобы исключить доработку технологического процесса.

Анализ этих видов работ показал, что следует использовать следующие математические модели:

а) Кинематическая модель – устанавливает перемещение полуформы и увязывает ее с положением расплава, с учетом особенностей конфигурации детали.

б) Тепловая модель – устанавливает аккумуляцию тепла формой и увязывает

теплоотвод с качественными характеристиками получаемого изделия.

Помимо этого необходимо обеспечить оптимизацию тепловых условий формирования изделия и с учетом оптимального режима синтезировать закон управления приводом установки.

В НИАТЕ путем проведения натурных экспериментов были получены необходимые данные для разработки моделей. Были предприняты попытки в конце 80-х годов решить эту задачу с использованием ЭВМ.

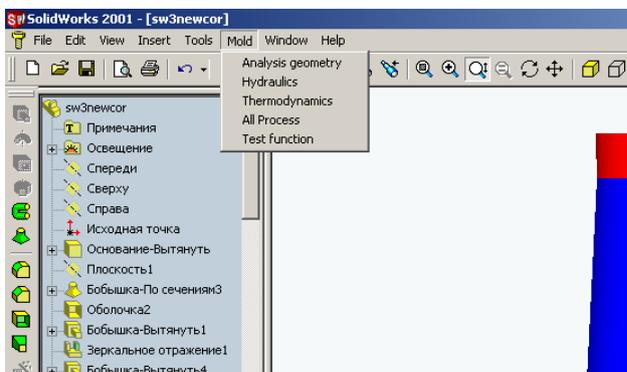


Рис. 1

Отечественные средства имитационного моделирования подобных процессов (Полигон, Ижевский) и зарубежные (MagmaSoft, ProCast, Flow3D, FlowWorks e.t.c) не предусматривают решения подобного рода задач, поскольку их модели не позволяют моделировать заливку расплава в форму с изменяемой геометрией.

В связи с этим возникла необходимость создать программный продукт, для моделирования литья выжиманием.

Модель, отображающая картину теплообмена в движущемся расплаве, реализована описанием ее уравнением Навье-Стокса и дифференциальным уравнением переноса теплоты и условия однозначности.

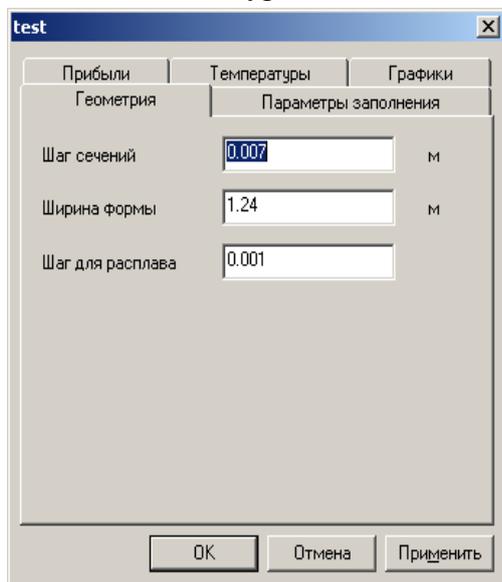


Рис. 2

Данная система моделирования отображает физику процесса полнее чем, это можно реализовать путем снятия экспериментальных данных, поскольку она позволяет получать данные в любой момент времени протекания процесса и на всей длине отливки с любой дискретностью шага.

Для решения этой задачи было принято решение использовать систему SolidWorks, поскольку для нее доступна документация по API и надежное ядро ParaSolid, средством разработки является Visual C++.

В приложении создан интерфейс взаимодействия конечного пользователя со средствами моделирования, он интегрирован с системой SolidWorks. (Рис.1) Исходными данными для выполнения моделирования служат геометрическая модель отливки и параметры, требующиеся для задания начальных условий которые задаются в диалоговом окне (Рис.2). Результатами моделирования являются зависимости, строящиеся по окончании моделирования средствами API SolidWorks (Рис 3), а также текстовые файлы, в которых можно получить необходимые детальные сведения.

Данные, полученные в результате моделирования, использовались при создании литейной машины сотрудниками НИАТ, что позволило сопоставить их с экспериментальными данными и подтвердить их надежность.

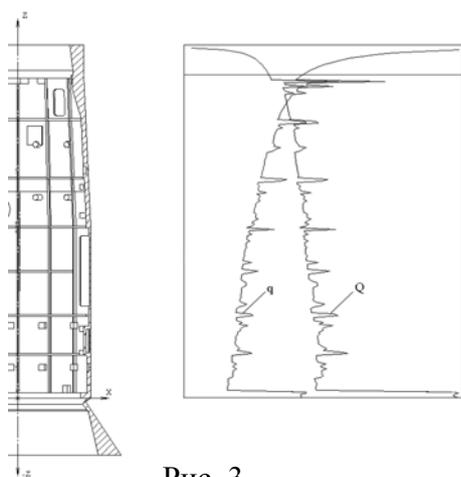


Рис. 3

Литература

1. Рубцов Н.Н., Стебаков Е.С. Новый способ литья тонкостенных крупногабаритных деталей. // Литейное производство, - М., -1956 . - №12. – С.8-10.
2. Баландин Г.Ф. Основы теории формирования отливки. Часть 2, -М.,- Машиностроение, 1979.,с.263.
3. Мусяиченко А.С., Милов В.Н., Глотов Е.Б., Теоретические основы метода литья выжиманием //Приложение к журналу Авиационная промышленность, -М.: Машиностроение. - 1984 .- №5. - С. 10 - 13. (ДСП).
4. Сигалова Е.Е., Ребиндер П.А. Исследование структурно-механических свойств и максотропии в олеоколлоидных системах. //Коллоидный журнал, -М., -1948 г.,10, №3.

ВОПРОСЫ РАЗРАБОТКИ ПРОГРАММНОГО КОМПЛЕКСА МОДЕЛИРОВАНИЯ ДИНАМИКИ ТЕХНИЧЕСКИХ СИСТЕМ НОВОГО ПОКОЛЕНИЯ.

Алиев Е., Баженов С., Пузырев М.

Научный руководитель: к.т.н., доцент Маничев В.Б.
МГТУ им. Н. Э. Баумана, Кафедра РК6, Москва, Россия.

DEVELOPING THE NEW GENERATION OF CAM/CAE SYSTEMS FOR DYNAMIC ANALYSIS.

Aliev E., Bazhenov S., Puziryov M.

The scientific chief: Ph.D., senior lecturer Manichev V.B.
BMSTU, RK6 Dep., Moscow, Russia.

Аннотация:

В докладе рассмотрены основные требования, предъявляемые к программному комплексу моделирования динамики технических систем нового поколения, дается обоснование необходимости разработки программного комплекса приведенного класса; рассмотрен обобщенный метод формирования и решения математических моделей динамики технических объектов в виде систем дифференциально-алгебраических уравнений (ДАУ).

Abstract:

This paper introduces base requirements to the new generation of software for dynamic analysis and explains the reasons for developing this type of software. Also it contains description of common approach to construction and evaluation of DAE systems for different mathematical models.

Введение

В настоящее время непрерывно растут требования со стороны организаций, выпускающих конечную продукцию, по снижению стоимости развертывания производства изделий и сокращению сроков выпуска изделий на рынок при одновременном повышении качества и снижении стоимости продукции. Возрастает роль быстрой и точной оценки принимаемых проектных решений. При этом крайне важна именно требуемая точность оценки, поскольку стоимость ошибок в принятии проектных решений на ранних стадиях проектирования чрезвычайно высока. Для решения задач моделирования и анализа проектных решений применяются системы инженерного анализа проектных решений – Computer Aided Engineering (CAE), которые можно разделить на две группы: системы анализа на распределенном уровне и на сосредоточенном уровне.

Математическими моделями технических объектов при проведении анализа на распределенном уровне являются системы дифференциальных уравнений в частных производных. Моделирование на распределенном уровне характеризуется высокой степенью точности получаемых результатов, но требует значительных затрат времени и вычислительных ресурсов и редко может применяться на ранних стадиях проектирования. Среди программных комплексов, выполняющих моделирование на распределенном уровне, следует упомянуть MSC.Nastran, ANSYS, COSMOS.

Моделирование на сосредоточенном уровне связано с решением систем алгебро-дифференциальных уравнений и позволяет получить точные результаты за приемлемое время без использования больших вычислительных ресурсов. Поэтому моделирование на сосредоточенном уровне находит все большее применение при проектировании изделий промышленности. К программным комплексам моделирования на сосредоточенном уровне относятся MSC.ADAMS, отечественные разработки серии ПА. В настоящее время наметилась тенденция к увеличению рынка подобных систем.

Однако, широкое распространение системы моделирования на сосредоточенном уровне получили только при проектировании изделий электроники и вычислительной техники, где применяются специализированные программные средства, в других же

областях, в частности, при проектировании машиностроительных изделий указанные системы используются лишь эпизодически. Это связано, в том числе, и с отсутствием на рынке универсальных систем, удовлетворяющих требованиям большого числа проектных организаций.

К недостаткам имеющихся на рынке систем следует отнести их узкую специализацию, так лидер на рынке MSC.ADAMS ориентирован преимущественно на проектирование изделий автомобильной промышленности, пополнение библиотеки моделей в программном комплексе MSC.ADAMS крайне затруднено и требует от пользователей навыков программирования. Существующие программные комплексы, за редким исключением, не предоставляют возможностей проведения моделирования систем, состоящих из физически различных подсистем.

В связи с этим могут быть сформулированы требования, которым должны удовлетворять разрабатываемые в настоящее время системы моделирования на сосредоточенном уровне:

- комплексное моделирование технических объектов, состоящих из различных физических подсистем - электрических, механических, гидравлических, тепловых и т.д.
- надежность и высокая точность математических ядер таких программ, возможность легкой замены и расширения математического ядра;
- возможность интеграции с системами геометрического моделирования;
- наличие интерактивных средств создания моделей пользователями, при этом от пользователя не должно требоваться специальных знаний в области программирования и/или численных методов;
- наличие средств параллельной работы разработчиков моделей, возможность выполнения вычислений на специально выделенном сервере или кластере;

Для решения перечисленных задач необходимо создание нового поколения систем моделирования динамических процессов в технических системах. Учитывая приведенные выше требования, а также опыт, накопленный при разработке программ ПА-4, ПА-6, ПА-7, ПА-9, предполагается разработать систему моделирования нового поколения ПА-10.

Математическое ядро

Как уже было сказано выше, программный комплекс ПА10 ориентирован, прежде всего, на задачи, связанные с решением систем дифференциально-алгебраических уравнений (ДАУ). Описание таких систем может производиться многочисленными способами, например, с помощью эквивалентных электрических схем или просто в виде естественной символьной записи.

Язык описания объекта моделирования и самой задачи должен обеспечивать разработчику необходимую гибкость и широту возможностей. Кроме того, такой язык должен быть легок в освоении и в достаточной степени документирован. В предыдущих версиях ПА(ПА7, ПА9) для разработки моделей были применены специализированные языки, описание на которых в дальнейшем конвертировалось в код С и Java соответственно, компилировались и подключались а пакет. Недостатки подобного решения (необходимость изучения как базового языка, так и его расширения, малое количество документации) очевидны.

В ПА10 решено реализовывать несколько иной подход. Разработка моделей ведется непосредственно на одном из популярных языков программирования, доступных в .Net. На

данный момент таких языков насчитывается более 20, включая бесплатные реализации компиляторов для C#, VB и Java.

Немаловажно, что применение объектно-ориентированных языков программирования для описания объектов моделирования позволяет легко перенести ООП в разработку самих моделей. Возможность установления отношений наследования и агрегации между классами, содержащими описание моделей объектов, позволяет нам легко реализовать, например, объединение нескольких моделей в одну или создание частных случаев моделей, не занимаясь разработкой с самого начала, а лишь частично перегрузив функциональность базовой.

Кроме разработки моделей непосредственно на языке программирования, планируется реализовать возможность их создания в визуальной среде, что не требует умения программировать вообще. В этом случае код классов будет генерироваться средой разработки, а затем компилироваться и подключаться в пакет.

Такой подход, на наш взгляд, должен значительно упростить разработку новых моделей и значительно расширить возможности пакета по сравнению с прошлыми версиями.

Рассмотрим далее непосредственно реализацию описанной выше концепции математического ядра ПА10.

Классы-компоненты

Компонент – элементарная единица вычислительного процесса. Каждый компонент описывает часть вычислительной задачи и символизирует собой одно или несколько уравнений в общей системе ДАУ. Можно заметить, что в случае эквивалентных электрических схем ими являются не только какие-то элементы (сопротивления, источники, конденсаторы и так далее), но и узлы. Для того, чтобы какой-то класс стал компонентом, он должен реализовывать интерфейс **IComponent** (см. рис. 1).

Хочется сразу отметить, что, на самом деле, компонент далеко не всегда знает в каких вычислениях ему приходится участвовать. В большом числе случаев он лишь несет описательную информацию, такую как свойства или, например, дифференциальное уравнение в текстовом виде. В компоненте безусловно могут производиться какие-то вычисления, но опять же, не всегда результаты этих вычислений могут быть сразу же, без дополнительной обработки, применены при решении системы уравнений.

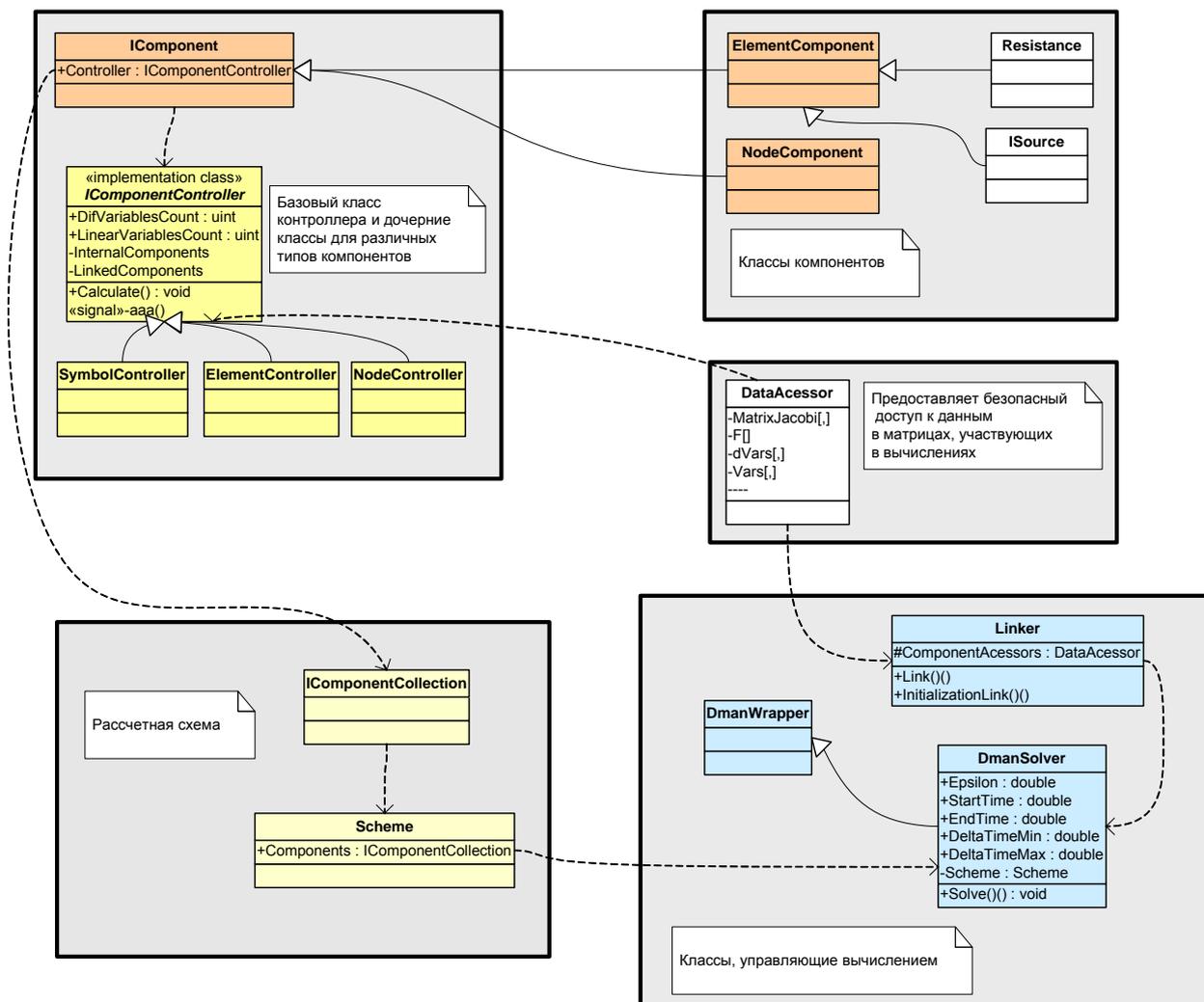


Рис. 1 Математическое ядро

Для перехода от данных и вычислений компонента к непосредственному заполнению матрицы Якоби и вектора значений функций (что собственно и требуется для решения системы ДАУ низкоуровневым математическим ядром) введен класс **ComponentController**. Объекты именно этого класса (или его дочерние) производят необходимые действия для заполнения матриц и другие сопутствующие вычисления.

Удобство такого подхода очевидно, так как значительно облегчает разработку компонентов. Большинство сложных вычислений и операций с матрицами перекладывается на классы-контроллеры, все основные варианты которых можно написать и отладить один раз.

Отметим тут же, что было бы неверным позволять контроллерам компонент иметь полный и неограниченный доступ к матрицам, участвующим в вычислениях. Это не только повышает возможность непреднамеренных ошибок, но и допускает специальную “порчу” данных. Именно поэтому все обращения к матрицам должны быть реализованы через специальный класс **DataAccessor**. Его задача – предоставить возможность доступа только к тем элементам матрицы, в которых компонент действительно нуждается.

Классы для управления процессом вычисления

Теперь следует более подробно рассмотреть классы, которые управляют процессом моделирования.

Во-первых, к таким классам относится сборщик модели, реализованный в классе **Linker**. Как уже упоминалось выше, решаемая система уравнений представлена в виде списка отдельных компонент. Класс **Linker** призван организовывать связь между этим списком и классом **DmanWrapper**, отвечающим за взаимодействие с низкоуровневым математическим ядром.

Его основные задачи:

- Определение размерности задачи для правильной инициализации **DmanWrapper**-а.
- Участие в создании **DataAcessor**-ов компонент для обеспечения безопасного доступа к данным (матрице Якоби, значениям переменных и так далее).
- Управление процессом заполнения матриц в процессе вычисления.

Кроме “сборщика” в процессе составления и решения системы участвует “разборщик”, задача которого управлять получаемыми в процессе моделирования данными. Применение А-устойчивых методов требует периодического взаимодействия с компонентом, предоставляющим уравнение. Разрывы производных и недоопределенность функций могут быть успешно выявлены именно в том случае, когда программная модель компонента информирована о ходе вычисления и промежуточных результатах.

Таким образом, решается задача построения математического ядра, независимого от представления копмонентов и вида уравнений. Соответствующие реализации перечисленных интерфейсов позволяют запрограммировать большинство методов автоматического составления сосредоточенных моделей. Также на их основе возможно построение анализаторов и компоновщиков систем для распределенных моделей.

Работы по построению работающей реализации велись в течение зимнего семестра 2003 года и завершились успешно. Система, разработанная в соответствии с приведенной архитектурой, решает тестовые задачи для небольших контуров, а результаты вычислений точно соответствуют данным, полученным в системе **MathCad**.

Литература

1. «Обобщенный метод моделирования динамики технических систем» Жук Д.М., Маничев В.Б., Папсуев А.Ю., МГТУ имени Н.Э.Баумана.
2. «Анализ требований и проектирование систем. Разработка информационных систем с использованием UML», Лешек А. Мацяшек. Перевод с английского – М.: Издательский дом «Вильямс», 2002. – 432 с.: ил.
3. «Programming C#», Jesse Liberty, O’Reilly, July 2001.

РЕАЛИЗАЦИЯ БЫСТРЫХ ПРЕОБРАЗОВАНИЙ ФУРЬЕ НА МНОГОПРОЦЕССОРНЫХ ВЫЧИСЛИТЕЛЬНЫХ УСТАНОВКАХ

Берчун Ю.В.

Научный руководитель: к.т.н., доцент Волосатова Т.М.
МГТУ им. Н.Э. Баумана, кафедра РК6

REALIZATION OF FAST FOURIER TRANSFORMATIONS ON MULTIPROCESSING COMPUTER PLANTS

Berchun Y.V.

The scientific chief: Ph.D., senior lecturer Volosatova T.M.
BMSTU, PK6

Аннотация

В работе предложен алгоритм разделения задачи нахождения БПФ на подзадачи, допускающие независимую обработку. Исследованы возможности применения этого алгоритма для реализации БПФ на многопроцессорных вычислительных установках. Также исследованы возможности реализации БПФ двумерного сигнала на многопроцессорных вычислительных установках.

Abstract

In the work the separating algorithm of task of FFT-determination to subtasks, allowing independent processing, is proposed. The opportunities of application of this algorithm for realization FFT on multiprocessing computer plants are investigated. The opportunities of realization FFT of two-dimensional on multiprocessing computer plants are investigated too.

Алгоритм разделения задачи нахождения БПФ на подзадачи

Суть предлагаемого алгоритма решения проблемы мультипроцессорной обработки состоит в разделении исходной последовательности отсчётов по алгоритму БПФ с множителями поворота (алгоритму Кули-Тьюки).

Алгоритм множителями поворота состоит в следующем. Пусть $N = N_1 N_2$, где N_1, N_2 — любые положительные целые. В этом случае вычисление N -точечного ДПФ можно свести к вычислению N_1 N_2 -точечных и N_1 N_2 -точечных ДПФ и N умножениям на множители поворота W_N^l . Для этого необходимо выполнить следующую подстановку:

$$k = k_1 + k_2 N_2, \quad n = n_1 + n_2 N_1,$$

где $n_1, k_2 = 0, \dots, N_1 - 1$; $n_2, k_1 = 0, \dots, N_2 - 1$.

Тогда ДПФ можно записать в следующем виде:

$$X(k_1 + k_2 N_2) = \sum_{n_1=0}^{N_1-1} \left[\left(\sum_{n_2=0}^{N_2-1} x((n_1 + n_2 N_1)T) W_{N_2}^{k_1 n_2} \right) W_N^{k_1 n_1} \right] W_{N_1}^{k_2 n_1}$$

Этот алгоритм обычно применяется для выделения последовательностей отсчётов с $N = 2^v$, $v \in \mathbb{N}$, для последующего их вычисления алгоритмами БПФ по основанию 2 (с прореживанием по частоте или по времени).

Однако мы будем использовать его в случае $N = (2^v)^2$, $v \in \mathbb{N}$ для выделения 2^v 2^v -точечных ДПФ, которые также будем рассчитывать при помощи алгоритмов БПФ по основанию 2. После этого последовательность домножается на множитель поворота, подвергается перестановке и вновь выделяются 2^v 2^v -точечных ДПФ.

В данном случае алгоритм не даёт выигрыша в количестве операций, однако позволяет выделить блоки, вычисление которых абсолютно идентично с точки зрения алгоритма, а потому может быть легко распределено между несколькими процессорами.

Приведём общий алгоритм работы программы:

- 1) Перестановка входной последовательности.
- 2) Вычисление 2^v 2^v -точечных ДПФ (1-я степень).

- 3) Домножение на множитель поворота перестановка.
- 4) Вычисление $2^v 2^v$ -точечных ДПФ (2-я ступень).
- 5) Итоговая перестановка и выдача выходной последовательности.

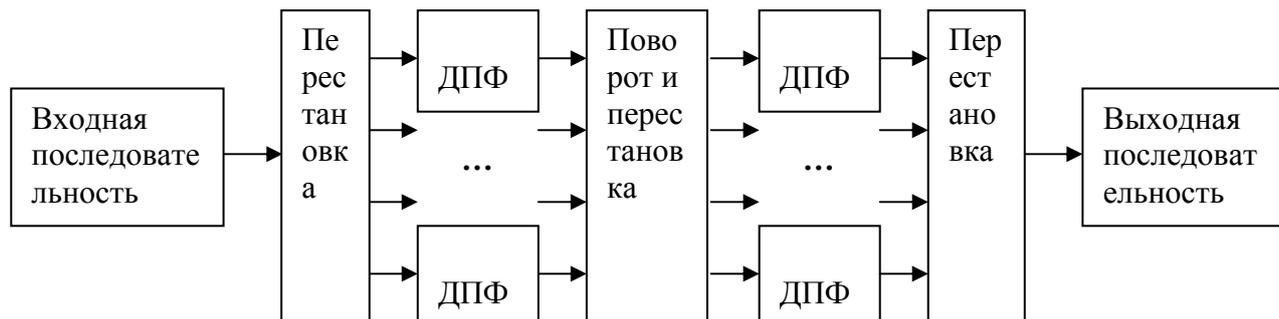


Рис. 1

На рис. 1 условно показан процесс преобразований входной последовательности отсчётов по предлагаемому алгоритму.

В качестве примера приведём направленный граф 16-точечного ДПФ, реализованного по данному алгоритму (рис. 2):

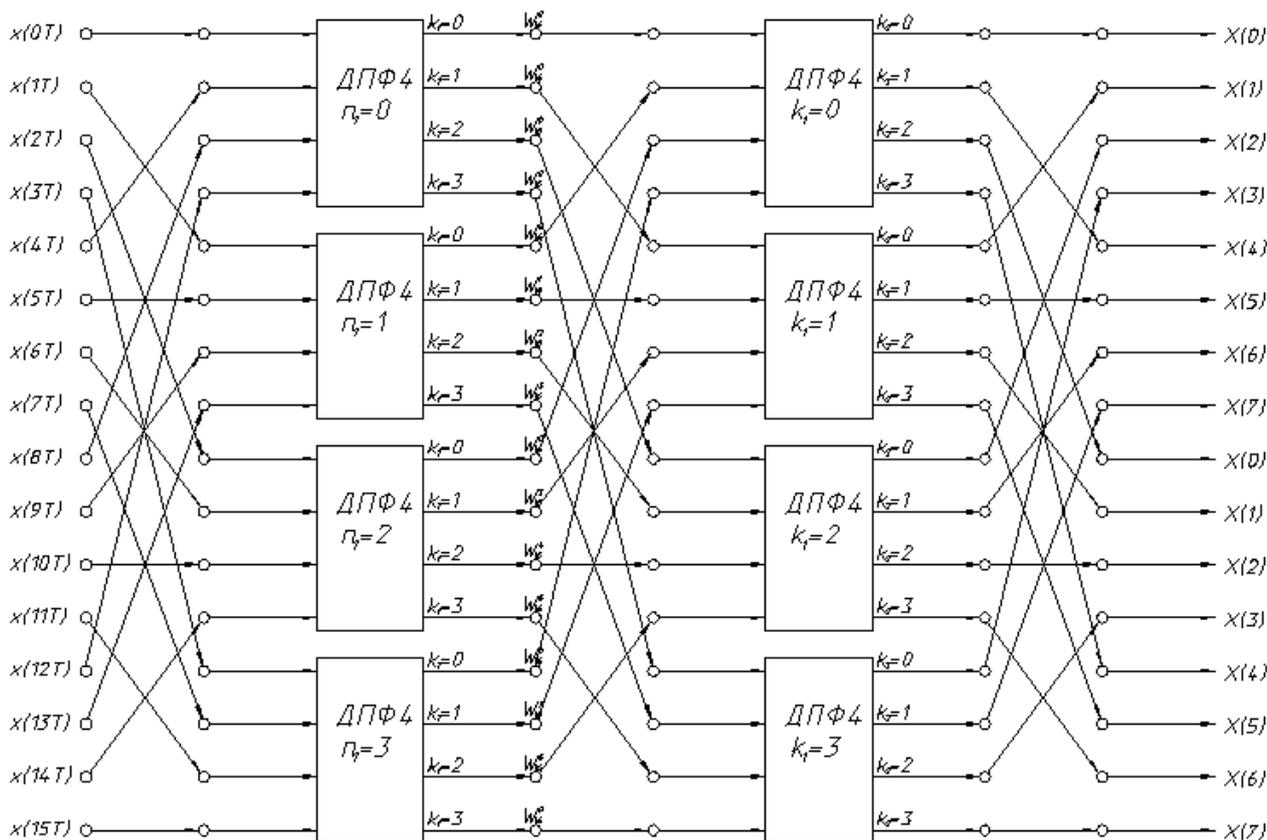


Рис. 2

Возможности применения предложенного алгоритма для реализации БПФ на многопроцессорных вычислительных установках

Проанализировав структуру алгоритма (см. рис. 1), легко видеть, что существует 2 подхода к реализации его на практике. Они опираются, соответственно, на SIMD (Single Instruction Multi Data) и на MISD (Multi Instruction Single Data) архитектуры.

В первом случае все имеющиеся в вычислительной установке процессоры ориентированы на вычисление ДПФ, находящиеся в одной ступени (в первой или второй). Если в системе

работают p процессоров, а число отсчётов равно $N = (2^v)^2$, $v \in \mathbb{N}$, то в каждой ступени требуется вычислить $2^v \cdot 2^v$ -точечных ДПФ. Пусть время вычисления одного ДПФ равно T . Тогда общее время работы программы (не считая затрат машинного времени на перестановки) составляет:

$$T_{\text{общ}} = 2 \cdot \frac{2^v}{p} \cdot T.$$

Во втором случае половина имеющихся процессоров обслуживает первую ступень, другая половина — вторую. Легко видеть, что:

$$T_{\text{общ}} = 2 \cdot \frac{2^v}{p/2} \cdot T,$$

то есть общее время возрастает вдвое. Однако при этом вдвое увеличивается частота появления выходных данных. Это необходимо для систем контроля технологического процесса и т.п., т.е. систем реального времени, в которых идёт постоянное формирование новых входных последовательностей (с перекрытием или без).

При этом стоит отметить, что в первом случае для хранения обрабатываемых данных достаточно одного массива, а во втором случае необходимо 2 массива (для первой и второй ступени). Синхронизация данных в этом случае осуществляется по завершении полной подготовке данных для второй ступени.

Программное управление вычислениями на многопроцессорных установках осуществляется средствами операционной системы. В Unix-подобных системах существуют 3 способа запуска p заданий для исполнения на отдельных процессорах:

- 1) создание новых процессов;
- 2) создание новых задач (threads);
- 3) Message Passing Interface (MPI).

Первые два подхода удобно использовать в системах с общей памятью. MPI применим и в системах с распределённой памятью. При реализации предложенного алгоритма удобнее использовать системы с общей памятью.

Кроме того, предлагаемый алгоритм можно реализовать на аппаратном уровне. Для этого необходимы микросхемы, реализующие 2^v -точечное ДПФ (или DSP-процессоры), а также схемы домножения на множитель поворота.

Возможности реализации БПФ двумерного сигнала на многопроцессорных вычислительных установках

Двумерным ДПФ двумерных (квадратных) последовательностей называется пара следующих преобразований:

$$X(k_1, k_2) = \sum_{n=0}^{N-1} \sum_{n=0}^{N-1} x(n_1 T, n_2 T) W_N^{k_1 n_1} W_N^{k_2 n_2}, \quad k = 0, \dots, N-1$$

$$x(n_1 T, n_2 T) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{k=0}^{N-1} X(k) W_N^{-k_1 n_1} W_N^{-k_2 n_2}, \quad n = 0, \dots, N-1$$

Т.е., достаточно найти ДПФ для всех столбцов в матрице $x(n_1 T, n_2 T)$, после чего транспонировать её и вновь найти ДПФ для всех столбцов (что аналогично нахождению ДПФ для всех строк в не транспонированной матрице). После этого матрица опять транспонируется.

Очевидно, что для вычисления каждого отдельно взятого одномерного ДПФ можно применить известные алгоритмы БПФ. Тогда при выделении каждому из p процессоров N/p заданий по вычислению N -точечных БПФ повышаем общую производительность в p раз. Потери возникают только при транспонировании матрицы.

Способы запуска p заданий одновременно в Unix-подобных операционных системах были рассмотрены выше.

Литература

1. Л.М. Гольденберг, Б.Д. Матюшкин, М.Н. Поляк. Цифровая обработка сигналов. Справочник. М.: "Радио и связь", 1985.
2. Л.М. Гольденберг, Б.Д. Матюшкин, М.Н. Поляк. Цифровая обработка сигналов. Учебное пособие для высших учебных заведений. М.: "Радио и связь", 1990.
3. М.С. Куприянов, Б.Д. Матюшкин. Цифровая обработка сигналов: процессоры, алгоритмы, средства проектирования. СПб.: "Политехника", 2002.
4. А.И. Солонина, Д.А. Улахович, Л.А. Яковлев. Алгоритмы и процессоры цифровой обработки сигналов. Учебное пособие. СПб.: "БХВ-Петербург", 2002.
5. К.Ю. Богачёв. Основы параллельного программирования. М.;БИНОМ. Лаборатория знаний, 2003.

РАЗРАБОТКА КОНЦЕПЦИИ МОДУЛЯ ДИСКРЕТНОГО МАШИНОСТРОИТЕЛЬНОГО ПРОИЗВОДСТВА В РАМКАХ ERP СИСТЕМЫ.

Бондарь Ю.А.
к.т.н. доцент. Жук Д.М.
МГТУ, г.Москва, Россия

MINING OF CONCEPT OF A MODULE DISCRETE MECHANICAL ENGINEERING PRODUCTION IN ERP SYSTEM.

Bondar Y. A.
Ph.D. docent. Ghuk D. M
MSTU, Moscow, Russia
e_mail: sentenel@mail.ru

Аннотация

В настоящее время системы автоматизации предприятия немислимы без автоматизации производства. Проведя анализ нескольких ведущих ERP(MRP2) систем российского производства выяснилось, что ни в одной из этих систем не реализован модуль производства в полном объеме. Например, в системе “ГАЛАКТИКА” сейчас ведется разработка такого модуля. Существуют отдельно разработанные программы, осуществляющие информационную поддержку производства на всех его стадиях, но их интеграция проблематична. Целью данной работы является разработка структуры базы данных и алгоритмов работы модуля дискретного производства в ERP системе. Данная задача является практической, разработанная концепция легла в основу проекта модуля на базе системы “Alfa”.

Abstract

Now systems of automation of company cannot be presented without an automation of manufacturing. By conducting the analysis several carrying on ERP system of the Russian production was found out, that in one of these systems the module of discrete manufacturing in a full volume is not realized. For example, in the system “Galaxy” is now conducted mining of such module. There are separately designed programs executing information support of manufacturing on all its stages, but their integration is problematic. The purpose of the given activity is the mining of frame of the database and algorithms of activity a module of discrete manufacturing in ERP system. The given problem is practical, the designed concept has come in fundamentals of the project a module on the basis system “Alfa”.

Введение

Перед разработкой концепции модуля, я провел ряд встреч с представителями различных фирм, занимающихся разработкой программного обеспечения, в этой отрасли: “БЕСТ”, “Галактика”, “Парус”, “TechnologicS”. После ознакомления с наработками данных фирм, выяснилось, что:

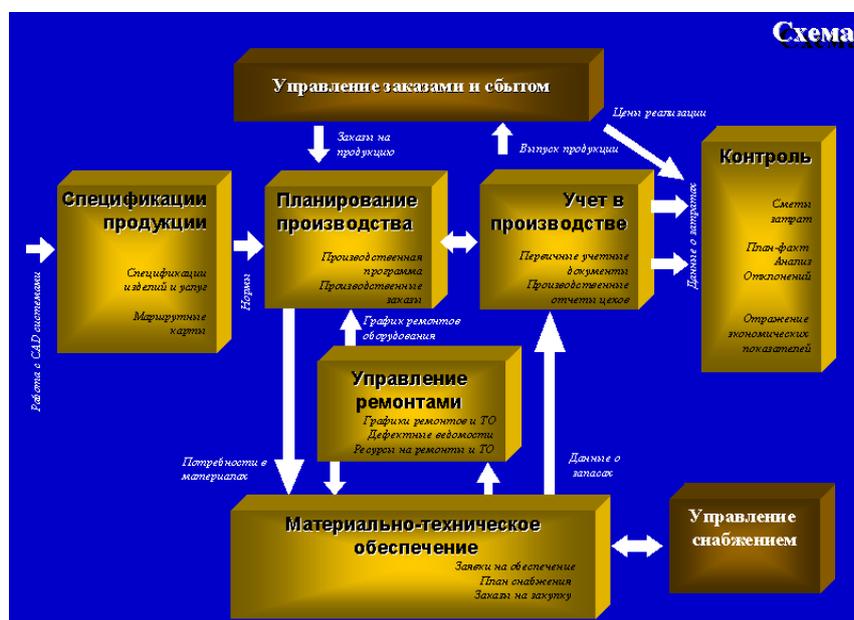
- Дискретное производство плохо поддается формализации из-за сложности бизнес процессов.
- Спрос на подобные разработки появился сравнительно недавно, и рынок программного обеспечения не успел еще среагировать.
- В основном предлагаются учетные системы, без реализации механизма планирования производства.
- Решение, удовлетворяющее большинству из перечисленных требований выпускается отдельно от какой либо КИС. Интеграция зачастую либо невозможна вообще, либо возможна, но в урезанном варианте (TechnologicS).
- Конкуренция с западными аналогами, которые зачастую не способны отразить реалии российского производства (ISO 9000 еще не гарантирует в России работу предприятия в западных стандартах).

Решение этих и других проблем я попытался учесть при разработке алгоритмов работы модуля дискретного производства.

При проектировании нового модуля появилась необходимость разделить весь его функционал на составляющие. Условно все функции модуля можно разделить на пять частей:



Все пять частей, тесно взаимодействуют друг с другом и другими модулями ERP системы. Это достигается использованием единой платформы БД и работой с бизнес объектами (накладными, ордерами и т.д.) единого вида. На рисунке ниже, представлена общая схема работы модуля, и его взаимодействие с уже существующими модулями системы.



Планируется, что информация о изделии будет подгружаться из CAD системы и в дальнейшем использоваться в качестве базовой. Консолидированные данные о потребностях в материалах будут отражаться в модуле управления снабжением. Так же сформированный производственный заказ и информация о выпущенной продукции будет передаваться из модуля управления заказами и сбытом.

Теперь вкратце охарактеризуем каждую из частей модуля. Я расскажу об алгоритмах, положенных в основу их работы и приведу макеты экранных форм, отражающих функционал модуля.

Подсистема “Планирование производства”.

Основные функции, этой подсистемы :

- Формирование производственной программы
- Формирование производственных заказов

Из смежного модуля, отвечающего за формирование заказов, поступает информация о типе и количестве заказанной продукции. Нормы расхода материалов и трудозатрат данная подсистема получает из подсистемы “Спецификации продукции”. Консолидированная информация об общих производственных потребностях передается в модуль системы, отвечающий за складской учет. Данные о плановом техническом оборудовании и внеплановых ремонтных работах передаются в подсистемы в виде графика ремонтов оборудования. После формирования производственной программы данные направляются в подсистему “Учет в производстве”.

Таким образом информация о запланированном производстве находит отражение во всей системе, в тех аналитических срезах, которые требуются для их работы.

Теперь рассмотрим процесс планирования и контроля процесса подготовки производства. При подготовке перед нами встают следующие задачи:

- Планирование работ
- Расчет трудоемкости и расхода материала, и тд
- Контроль комплектности документации
- Заказы на инструментальное производство

На производстве решение этих задач, осуществляется путем формирования производственной документации на различные направления деятельности:

- План работ технического бюро
- Сводные расчеты (трудозатраты, расход материала и тд)
- Техническая документация (CAD,CAM)
- Документы на оснастку производства

В терминах системы эта нормативная документация будет выглядеть следующим образом:

- Создание спецификации
- Нормирование материалов, расцеховка
- Проектирование технологических процессов
- Инструментальное производство

Из всего вышесказанного следует, что входными данными при работе данной подсистемы, в терминах системы, будут являться: заказ отдела сбыта, технологические маршруты изготовления, нормы расходов материала. Решаемые задачи внутри модуля и сервис функции: формирование плана выпуска продукции, анализ плана на выполнимость, формирование производственных заказов, расчет потребностей в ресурсах, формирование заданий по цехам и лимитнозаборных карт, формирование заявок на материально техническое обеспечение (МТО). Все решаемые задачи консолидируются во внешнюю информацию: производственный заказ, заказ на МТО, информация о затратах.

Данная подсистема будет позволять рассчитывать загрузку производственных мощностей, что является необходимым атрибутом современной системы автоматизации производства. Расчет загрузки основан на том, что мы уже знаем производственную программу, а, следовательно, знаем и группы выпускаемой продукции и необходимое оборудование. С другой стороны нам известен производственный календарь и график ремонтов оборудования, следовательно, мы можем составить график работы оборудования. Исходя из производительности оборудования и количества продукции по группам, мы можем рассчитать коэффициенты загрузки мощностей.

Подсистема “Спецификации изделий”.

Основные функции этой подсистемы :

- Создание спецификаций на изделие или услуги
- Генерация маршрутных карт обработки

Из САД систем в данную подсистему передается спецификация и разузловка изделия, чертежи детали и нормы расхода материала на данное изделие. Далее формируются пооперационные трудовые нормы. И консолидированную нормативную информацию передаем в смежную подсистему “Планирование производства”.

В системе изделие можно представить в виде взаимосвязанной модели. Основные атрибуты изделия - это производственная спецификация и маршрутная карта. В свою очередь Производственная спецификация позволяет получить ведомости замен, нормы расхода материала и нормы выхода продукции. Маршрутная карта несет в себе информацию об оборудовании и оснастке и пооперационных трудовых нормах. Эта структура в совокупности с нормами отходов позволяет полностью представить себе технологический процесс изготовления данного изделия.

Рассмотрим процесс проектирования через групповые техпроцессы : Для изготовления изделия нужно произвести расцеховку, то есть указать какие операция будут выполняться в том или ином подразделении. Далее для каждого подразделения составить маршрутный техпроцесс, позволяющий определить очередность операций производимых в каждом подразделении. И наконец операционный техпроцесс, позволяющий определить на каком оборудовании и каким инструментом будет производиться конкретная операция.

Подсистема “Учет в производстве”.

Данная подсистема предназначена для :

- Учета первичных документов
- Создание производственных отчетов цехов

Данная подсистема тесно взаимодействует с “Планированием производства”, для уточнения плана производства производится план-факт анализ. Данные о запасах из модуля складского учета, используются для ведения статистического анализа производственной деятельности. Данные о производственных затратах передаются в бухгалтерский и управленческий учет. Данные, о выпущенной продукции, сформированные на основании плана производства, передаются на уровень учета готовой продукции.

Для учета движения МТО на производстве в подсистеме использована следующая схема: первый уровень – это движение ресурсов в цехе, он состоит из информации об остатках на начало периода, НЗП (незавершённое производство), поступления за период, расход ресурсов, выбытие ресурсов, остатки на конец периода. На втором уровне ведем учет расхода ресурсов на объекты учета затрат: расход материалов и трудозатраты. И третий уровень учета – это учет выпуска продукции и материальных ценностей: объемы выпуска продукции цехом и передача продукции на склады и реализацию.

Формирование производственного отчета цеха за период происходит по следующей схеме: после того как сформирован производственный заказ, формируются лимитно-заборные карты и акты замены, мы получаем все нормативные документы, которые вместе со сменными рапортами и позволяют сформировать требуемый отчет.

Отражение учетных документов можно увидеть на всех параллельных планах счетов: бухгалтерском, налоговом и управленческом.

Подсистема “Управление ремонтами”.

Данная подсистема предназначена для :

- Построение графиков ремонтов и ТО
- Создание дефектных ведомостей
- Планирование ресурсов на ремонт и ТО

При проектировании данной подсистемы, было выявлено ее тесное взаимодействие с материально-техническим обеспечением. В процессе ремонта оборудования, необходимы запасные части, смазочные материалы, инструменты. По графику ремонтов, заранее формируется заказ на оснастку ремонтного производства. После того как все требуемые запасные части и инструменты поступают в ремонтный цех, оборудование уходит на плановое ТО или срочный ремонт.

Сформированный график ремонтов учитывается при создании плана производства на период.

Подсистему можно представить в виде кольца, на котором нанизаны различные операции и исполнителя данных операций в центре.

Подсистема “Контроль”.

Данная подсистема предназначена для :

- Расчета сметы затрат
- Создание план-факт анализа отклонений
- Оперативное отражение экономических показателей

На каждом производстве особняком стоит отдел контроля. Его функции возложены в модуле на данную подсистему. Она взаимодействует с бухгалтерскими модулями ERP системы. Из подсистемы “Учет в производстве”, можно получить данные о затратах в данном периоде. Данные о заказах поступают из модуля оформления заказов.

План-факт анализ – это необходимый инструмент руководителя высшего звена, именно он позволяет принимать решения о увеличении инвестирования в производство предприятия. На основе оперативных экономических показателей производства, можно вносить изменения в текущий план производства. Расчет сметы затрат, и ориентировочной себестоимости позволяет отслеживать затраты на производство.

База данных модуля “Дискретное производство”.

Данная база данных отражает взаимосвязь между основными параметрами и характеристиками производства. Использование справочников в системе позволит связывать между собой разнородные данные, общими аналитическими признаками.

Используя эти признаки в качестве основы для фильтрации информации, можно будет получить отчет в любом аналитическом разрезе.

Все использованные структуры данных, в базе данных являются зарегистрированной торговой маркой фирмы “Информконтакт” и используются с разрешения создателей.

СИСТЕМА МОДЕЛИРОВАНИЯ ПРОЦЕССОВ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ

Беломойцев Д.Е.

Научный руководитель: к.т.н, доцент Волосатова Т.М.
МГТУ им. Баумана, кафедра РК6

PARALLEL PROCESSING MODELING SYSTEM

Belomoytsev D.E.

Scientific chief: Ph.D., senior lecturer Volosatova T.M.
BMSTU, PK6
dimitry.belomoyzev@a4vision.com

Аннотация

Рассмотрены наиболее производительные и наиболее распространенные архитектуры построения вычислительных систем (ВС). Определены решаемые данными ВС области задач. Описаны основные приемы оценки возможности распараллеливания и методы распараллеливания программных реализаций алгоритмов обработки изображений. Составлена схема описания структуры основных узлов ВС для последующей оценки временных затрат на осуществление вычислений. Разработана технология моделирования производительности ВС на основе схемного описания ВС и оценочных времен выполнения элементарных операций.

Abstract

The most powerful and the most wide-spread computer architectures are reviewed. Task entities solved by these computers are determined. General parallelizing ability estimation ways and image processing algorithms' software realizations parallelizing methods are described. Basic computer blocks schematic structure description for further computation performance timing estimation is composed. Computer performance modeling technology based on computer schematic description and elementary operations' execution timing estimations is developed.

На сегодняшний день разработаны и успешно используются для создания компьютеров ряд архитектур, которые отличаются в основном принципом обработки во времени потоков управляющих воздействий и данных. Для каждого архитектурного типа можно выделить род задач, наиболее успешно (в плане производительности) решаемый компьютерами построенными на этой архитектуре. Например, для выполнения задач обработки изображений в силу манипулирования огромными объемами данных и, довольно часто, проведения сложных в плане совокупности действий операций наиболее удобны архитектуры SIMD, MIMD, SPMD [5]. Широкое распространение получили цифровые сигнальные процессоры (DSP) построенные по VLIW-архитектуре [3,5].

Успех программной реализации того или иного алгоритма зависит также и от используемых компилятора (его возможности по оптимизации для выполнения процессором конкретной архитектуры, например, для C/C++ Intel Compiler в сравнении с Microsoft Compiler), операционной системы (организация многозадачного режима).

Также существует ряд «параллельных» языков программирования, например HPF или HPC, предоставляющих возможности для программирования параллельных задач. Однако в данной работе наибольший интерес представляет язык программирования C++, соответствующий стандарту ISO/IEC 14882 [6].

Во многих случаях при изначально стоящей цели разработки программного обеспечения для решения задач обработки изображений и имеющихся аппаратных средствах для решения этих задач становится необходимо найти решение конкретной задачи путем выполнения ряда шагов:

- обзор имеющихся алгоритмов решения подобного класса задач;
- отбор наиболее пригодных с точки зрения производительности, требований к ресурсам и других критериев;
- выбор алгоритма наиболее подходящего архитектуре имеющейся аппаратной части;

- программная реализация (кодирование, оптимизация, тестирование, выполняемые в цикле);
- модификация, если существует острая необходимость, аппаратной части.

Целесообразным при выборе алгоритма для его программной реализации является проведение оценки возможности распараллеливания данного алгоритма. Обладая такой оценкой, можно сделать предположение о времени выполнения алгоритма применив закономерность Амдала

я $R(f) = \frac{1}{f/R_H + (1-f)/R_L}$, где f – доля распараллеливаемого кода в

программной реализации, $f \in [0,1]$; R_H и R_L – границы производительности снизу и сверху, соответственно [1,2].

Получить сведения о степени параллелизма f можно, построив диаграмму зависимостей данных, используя условия Бернштейна [1,2]. Нелишним будет использовать и закономерность Густафсона [2] в случае архитектур с распределенной памятью.

Для достижения более высокой степени распараллеливания возможно использовать приемы развертки циклов, с помощью которых можно повысить как эффективность последующей компиляции, так и эффективность использования кэш-памяти в случае обработки циклом больших объемов данных, распределенных по памяти. Среди таких методов возможно выделить операции с содержимым цикловых действий [1]

- распределение (loop distribution);
- слияние (loop fusion);
- обращение (loop reversal);
- смена порядка прохождения (для циклов по нескольким индексруемым величинам – loop interchanging);
- развертка (loop unrolling);
- свертка (loop collapsing);
- сращивание (loop coalescing);
- распространение (loop spreading);

Таким образом, разрабатывается модуль, в задачи которого входит

- общая оптимизация кода,
- составление диаграммы зависимостей,
- определение степени параллелизма кода,
- предложение вариантов по устранению зависимостей, преобразованию циклов с целью повышения степени параллелизма программной реализации.

Попытку смоделировать возможную производительность конкретной вычислительной системы при выполнении конкретной программной реализации можно провести путем составления схематического описания структуры основных блоков ВС, задания временных затрат на осуществление элементарных операций и последующего разложения на элементарные составляющие действия программы, быстродействие которой оценивается.

С этой целью был выработан перечень элементарных действий, от затрат времени на выполнение которых в наибольшей степени зависит быстродействие системы. В этот перечень вошли:

- время доступа (чтения, записи) к данным в ОЗУ (оперативная память),
- время доступа к данным к кэш-памяти (если в случае с ОЗУ ее объем в высокопроизводительных системах практически всегда достаточно велик, то величина кэш-памяти и сегодня играет важную роль, т.к. «ее много не бывает»),
- время выполнения операций сложения, умножения и т.п.,
- время передачи данных по магистралям (при распределении вычислений м.б. как между процессорами многопроцессорной системы, так и по локальной сети между однопроцессорными станциями).

В перечень оцениваемых операций были включены замеры времени выполнения циклов, различных по своей структуре. Циклы практически всегда используются для обработки больших объемов данных. Поэтому важно определить, происходит ли при выполнении чтение из разных областей памяти или по близким адресам. При этом в одном случае кэш-память будет постоянно обновляться, будет происходить большое количество промахов, однако в другом случае содержимое кэш-памяти будет более или менее статичным, что приведет к меньшему количеству обращений к ОЗУ и, соответственно, к повышению производительности [8, 12].

Оценка времени выполнения внесенных в перечень элементарных операций осуществлялась на базе рабочих станций, построенных на процессорах архитектуры x86 компаний

- Advanced Micro Devices Inc. (K6-2 300MHz (66x4.5, L1=32KB+32KB), K6-2 400MHz (100x4.0, L1=32KB+32KB) [11], K7 750MHz (100x7.5, L1=64KB+64KB, L2=256KB), K7 1666MHz (133x12.5, L1=64KB+64KB, L2=256KB) [10]),
- Intel Corp. Inc. (IP-III 1GHz (133x7.5, L1=16KB+16KB, L2=256KB) [7], IP-IV 2400MHz (400x6.0, L1=8KB+12KB, L2=512KB, L3=2048KB) [9]),

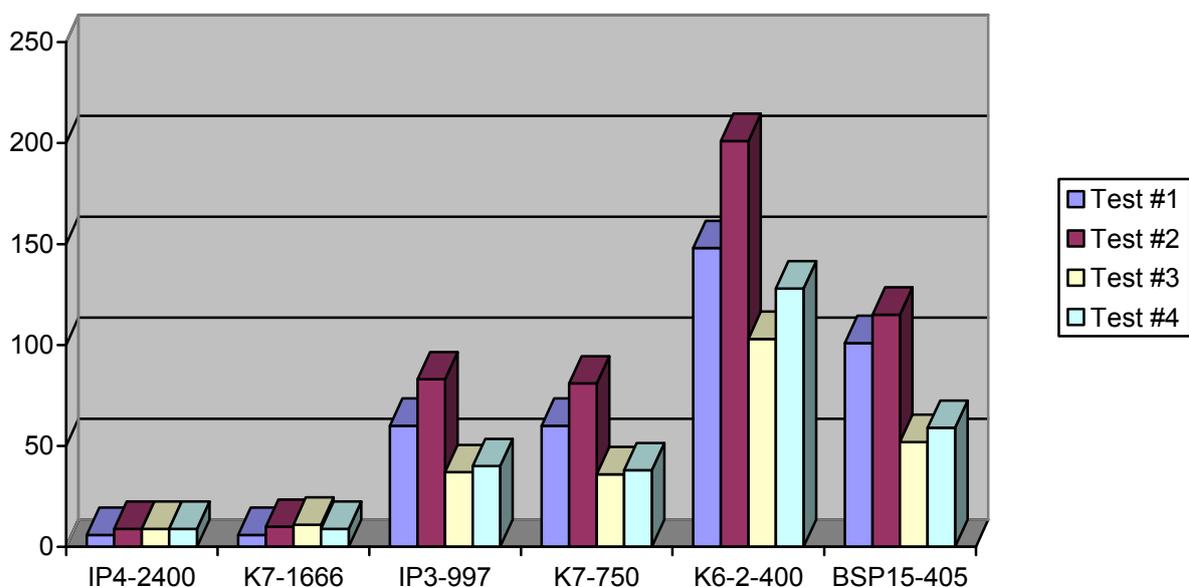
а так же на DSP платформе Tetra компании Equator Technologies Inc. с процессорами MAP-CA BSP-15 (215MHz и 405MHz, L1=32KB+32KB).

Результаты замеров затрат времени на выполнение некоторых тестов в зависимости от типа ОС (MS Windows2000SP2 и RH Linux 9.0) приводятся в табл.1.

Затраты времени на выполнение теста, нс. Таблица 1.

Тест	Платформа													
	IV- Intel Pentium 2400MHz/512MB		AMD Athlon XP 2000+ 1666MHz/512MB		III- Intel Pentium 1GHz/512MB		- AMD Athlon 750MHz/512MB		- AMD K6-2 400MHz/128MB		- AMD K6-2 300MHz/128MB		BSP-15 -405MHz/64MB	BSP-15 -215MHz/64MB
№1	6	6	7	6	64	60	60	60	153	148	189	190	101	193
№2	10	9	10	10	85	83	82	81	209	201	241	237	115	199
№3	9	9	11	11	35	37	39	36	111	103	133	139	52	97
№4	9	9	9	9	39	40	41	38	129	128	151	155	59	106

Результаты замеров в графическом представлении выглядят следующим образом:



Также проводились замеры затрат времени на передачу данных по локальной сети Ethernet 100Mbit/s с использованием TCP и UDP соединений [4].

Разрабатываемый модуль таким образом оценивает затраты на выполнение для конкретной конфигурации моделирует выполнение конкретного программного кода, как это было сделано для перечисленных выше конфигураций.

В настоящий момент проводится расширение базы элементарных составляющих структуры блоков ВС, сбор дополнительных экспериментальных сведений о затратах времени на различных платформах. Также проводится пополнение перечня поддерживаемых методов преобразования программного кода с целью повышения возможности его распараллеливания.

Литература

1. H.F.Jordan, G.Alagband. "Fundamentals Of Parallel Processing". Pearson Education, Inc. Upper Saddle River, New Jersey, 2003.
2. S.H.Roosta. "Parallel Processing And Parallel Algorithms: Theory And Computation". Springer-Verlag. New York, 2003.
3. Tetra CPU Module Hardware Reference Guide, Equator Technologies Inc, 2003.
4. MPI – The Complete Reference, The MIT Press, 1996.
5. Parallel Image Processing Toolkit Programmers` Guide, 2001.
6. International Standard ISO/IEC 14882: Programming Languages – C++. ANSI, 1998.
7. Intel Pentium® III Processor for the PGA370 Socket at 500 MHz to 1.13 GHz Datasheet. Intel Corporation, 2001.
8. IA-32 Intel® Architecture. Software Developer`s Manual. Volume 2: Instruction Set Reference. Intel Corporation, 2003.
9. Intel® Pentium® 4 Processor in the 478-Pin Package at 1.40 GHz, 1.50 GHz, 1.60 GHz, 1.70 GHz, 1.80 GHz, 1.90 GHz, and 2GHz Datasheet. Intel Corporation, 2002.
10. AMD Athlon XP Processor Model 6 Data Sheet. Advanced Micro Devices Inc., 2002.
11. AMD K6-2 Processor Data Sheet. Advanced Micro Devices Inc., 2000.
12. Software Development Guide AMD-K5 Processor. Advanced Micro Devices Inc., 1997.

СРАВНЕНИЕ ПОДХОДОВ К ПРОЕКТИРОВАНИЮ ПРОГРАММНЫХ СИСТЕМ

Андронов А.В.

Научный руководитель: к.т.н., доцент Жук Д.М.

МГТУ им. Баумана, кафедра РК6

THE COMPARISON OF APPROACHES TO DESIGNING PROGRAMS

Andronov A.V.

Scientific chief: Ph. D., senior lecturer Zhuk D.M.

BMSTU, RK6

kutushut@mail.ru

Аннотация

Рассмотрены общие тенденции развития процесса разработки программного обеспечения. Приведены основные этапы разработки. Рассмотрено использование различных подходов: Rational Unified Process, классический последовательный подход, экстремальное программирование. Сделаны выводы по применимости подходов для проектов различной сложности.

Abstract

The general trends of the development of the process of the software development are considered. The main stages of the development are brought. Considered use different approach: Rational Unified Process, classical consequent approach, extreme programming. Findings are made on applicability approach for project of the different difficulty.

Развитие вычислительной техники приводит к усложнению программного обеспечения, что, в свою очередь, требует написания все более сложных программ. Даже понятие «программный продукт» уже ни о чем не говорит, требуя детализации.

Задачи, которые ставятся при проектировании, всегда схожи. Требуется создать систему, максимально соответствующую исходному описанию, с учетом налагаемых на нее ограничений.

Подход к проектированию зависит от степени сложности системы и некоторых других факторов. Среди них – необходимость поддержки пользователей, выпуска обновлений, дополнений, требования к документации.

Системы можно условно разделить на малые, средние и крупные. В настоящее время существует мощная методология создания крупных программных систем – Rational Unified Process (RUP). Она позволяет распределить деятельности в коллективе разработчиков, где каждый отвечает за свою часть проекта, эффективно взаимодействуя. Такой подход полностью оправдан только при наличии серьезной задачи, выполнение которой требует большого коллектива разработчиков и эффективного взаимодействия.

При решении задач средних объемов, как правило, нет необходимости в детально расписанном плане работ с указанием деятельности каждого участника проекта. Упор должен быть сделан на детализацию структуры системы и взаимодействие разработчиков. В данном случае важно направить усилия руководящего звена на обеспечение взаимодействия в рамках поставленных задач, а не на четкое распределение обязанностей.

Малые проекты иногда выполняются вовсе без предварительного проектирования, иногда формируется некоторый набросок системы, который затем уточняется.

Можно заметить, что при переходе от простых систем к сложным происходит переход от проектирования «в уме» к RUP. Соответственно, исключением отдельных элементов из RUP можно получать более простые подходы.

Подходы к проектированию можно разделить на традиционные, основанные на последовательном решении задачи, и итеративные, основанные на выполнении ряда законченных подзадач.

Жизненный цикл программного продукта включает следующие стадии: анализ, проектирование, построение, внедрение. В рамках этих стадий выполняются: анализ, проектирование, реализация и тестирование.

В различных подходах время, отводимое на разные стадии, различается. Хотя время зависит от объема проекта, в рамках схожих систем можно провести сравнение.

Поскольку проектирование во многом творческий процесс, трудно получить достоверные количественные характеристики. Однако можно получить качественные оценки времени, затрачиваемого на отдельные этапы.

Для крупных систем оправдано применение RUP, поскольку в таком случае используются все его преимущества и разработка протекает скорее как производственный процесс, а не как творческий. Другие подходы дают меньшую эффективность из-за неприспособленности к высокой сложности систем.

Для средних систем применение RUP дает выигрыш на стадии внедрения – к этому моменту система уже практически полностью отлажена и не может быть неожиданных задержек при выпуске продукта. Однако на первых стадиях тратится много времени на анализ и распределение обязанностей, поэтому независимо от квалификации разработчиков выпуск займет ощутимое время.

Традиционный подход позволяет сократить время на анализ и проектирование. Однако на стадии внедрения в процессе тестирования могут быть получены неожиданный результаты, отбрасывающие процесс разработки на предыдущие этапы. Такой подход годится для хорошо структурированных систем, однако даже в этом случае существует риск на стадии внедрения. Время на разработку в этом случае будет сокращено по сравнению с RUP.

Экстремальное программирование позволяет взглянуть на систему с точки зрения конечного пользователя. Не вдаваясь в основные принципы этого подхода, рассмотрим его плюсы и минусы для средних систем. Плюсом является высокая скорость разработки. Отдельные подсистемы точно отвечают требованиям ТЗ и проходят жесткое тестирование. В целом система составляется из отдельных компонентов. Минусом такого подхода является принципиальный отказ от «работы на будущее». Любое изменение функциональности влечет за собой затраты времени на доработку. Этот подход позволяет быстро разрабатывать хорошо структурированные системы, тесно работая с заказчиком.

В результате анализа рассмотренных подходов наилучшим по скорости является экстремальное программирование. Однако для плохо структурированных систем стоит применять RUP, чтобы зафиксировать сроки разработки.

Для малых систем RUP не применим, поскольку сроки разработки будут заведомо большими, чем при любом другом подходе.

Экстремальное программирование не очень эффективно, поскольку задачи малого объема трудно разбить на еще более мелкие. К тому же, как правило, ТЗ на такие задачи формулируется в общих чертах, что не позволяет эффективно использовать преимущества данного подхода.

В результате сравнения можно сделать вывод, что универсального подхода к проектированию не существует. Однако для различных типов систем можно выделить предпочтительные подходы. Также хорошим решением будет комбинация различных подходов для каждой конкретной задачи.

Литература

1. Якобсон А., Буч Г., Рамбо Дж. «Унифицированный процесс разработки программного обеспечения» - СПб.: Питер, 2002г.
2. Трофимов С.А.. «CASE-технологии», М.: Бином-Пресс, 2002г.
3. С. Рубцов. «Сравнительный анализ и выбор средств инструментальной поддержки», публикация Интернет (<http://or-rsv.narod.ru>)
4. Тим Ландгрейв. «Экстремальное программирование», публикация Интернет (<http://profi-club.kiev.ua>)
5. Материалы сайта <http://uml.ru>

ИССЛЕДОВАНИЕ РЕШЕНИЯ ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ МЕТОДОМ РЕГУЛЯРИЗАЦИИ ТИХОНОВА

Воронов А.В.

Научный руководитель: к.т.н., доцент Волосатова Т.М.

МГТУ им.Н.Э.Баумана, Москва, Россия.

SOLVING INTEGRAL EQUATIONS USING TIKHONOV'S REGULARIZATION METHOD.

Voronov A.V

The scientific chief: Cand. Tech. Sci., docent Volosatova T.M.

MSTU n.a. Bauman, Moscow, Russia

e-mail: thiverus@yandex.ru

Аннотация

Приведен обзор сфер применения интегральных уравнений в технике. Приведен метод решения конкретной задачи. Изложены общие сведения о некорректных задачах и пути их решения, в частности, методом регуляризации Тихонова. Приведены результаты исследования данного метода.

Abstract

The general review of integral equations and the reasons of using them is given. Specific task is introduced. General information about incorrect tasks and the ways of solving them is given. Results of research of Tikhonov's regularization method are represented.

Область применения интегральных уравнений сейчас достаточно велика. Выбор последних обусловлен такими факторами, как невозможность составления других уравнений, необходимость снижения размерности задачи, возможность компактной формулировки граничных задач, достижение упрощений при вычислениях, возможность простого и естественного перехода к системам конечных уравнений. В каждом конкретном случае во внимание могут быть приняты не все, а иногда и всего лишь один фактор. Можно перечислить следующие области, где интегральные уравнения нашли эффективное применение: физика – задачи диффузии и явления переноса, теория потенциала, теория упругости, теплопроводность и т.д.; механика – многомассовые системы, колебания валов и другие задачи анализа динамики машин и механизмов; аэродинамика – собственные колебания крыльев самолета, флаттер и т.д.; электротехника – анализ и оптимальный синтез электрических цепей; автоматическое регулирование—анализ переходных процессов, расчет оптимальных характеристик, задачи синтеза и идентификации; астрономия – анализ распределения масс в галактиках; радиотехника – анализ поля антенн, синтез антенн и т.д.; акустика – анализ и синтез акустических излучателей и сигналов; теория и практика измерений – восстановление измеряемых сигналов по показаниям приборов; теория надежности – задачи восстановления; теория точности вычислений – получение оценок погрешностей при решении дифференциальных уравнений и т.д. Этот перечень может быть продолжен.

В основу перечисленных задач заложен алгоритм решения уравнений следующего вида:

$$\int_a^b K(x, s)y(s)ds = f(x), \quad c < x < d$$

Данный интеграл – это уравнение Фредгольма первого рода. Здесь $K(x,s)$ – заданная непрерывная по совокупности аргументов функция, называемая ядром интегрального уравнения, $f(x)$ – заданная непрерывная функция, называемая неоднородностью уравнения), $y(x)$ – неизвестная непрерывная функция.

Одним из примеров, где необходимо решать такое уравнение, является задача спектроскопии (изучение спектрального состава светового излучения). Пусть $y(s)$ – искомое распределение плотности энергии по спектру излучения в функции частоты s . Пусть далее

$f(x)$ – экспериментальный спектр, полученный в результате пропускания излучения через измерительную аппаратуру, характеризующуюся аппаратной функцией $K(x,s)$. Тогда задача определения истинного спектра $y(s)$ по измеренному $f(x)$ сведется к решению интегрального уравнения Фредгольма первого рода

$$\int_a^b K(x,s)z(s)ds = f(x), c \leq x \leq d,$$

где a и b – границы истинного спектра, c и d – границы экспериментального спектра.

Данная задача является некорректной. Задача корректная, если:

- решение существует для любой определенной на $[c,d]$ функции $f(x)$
- решение определено однозначно
- решение устойчиво

Данная задача некорректна, в первую очередь, по причине нарушения третьего условия: даже очень малые ошибки правой части $f(x)$ могут приводить к настолько большим ошибкам в решении, что оно совсем не будет похоже с искомой функцией. Если решать такое уравнение методом квадратурных формул, заменяя интеграл конечной суммой по формуле трапеций с шагом $h=const$ и решая получающуюся СЛАУ относительно значений $y(a)$, $y(a+h)$, ... , $y(b)$, то вместо истинного решения зачастую будет получаться так называемая знакопеременная «пила» большой амплитуды, которая при ее подстановке в интеграл, тем не менее, дает заданную $f(x)$.

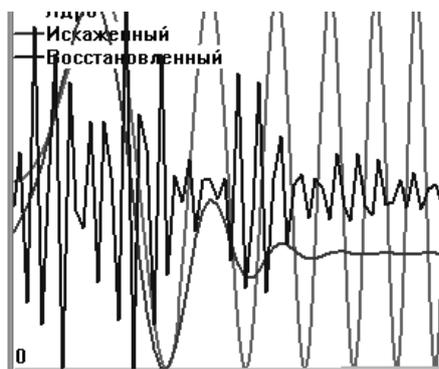


Рис.1. Результат без регуляризации

При этом обнаруживается следующее обстоятельство: чем меньше шаг h и чем, казалось бы, точнее аппроксимируется интегральное уравнение, тем грубее решение (тем больше амплитуда «пилы»).

Хотя задача решения такого уравнения является некорректной, но, если решение искать в компактном множестве, то оно будет устойчивым. Рассмотрим уравнение

$$Ay \equiv \int_a^b K(x,s)y(s)ds = f(x), c \leq x \leq d$$

В методе регуляризации 0-го порядка (слабой регуляризации) Тихонова вводится сглаживающий функционал

$$\Phi_\alpha = \int_c^d [Ay - f(x)]^2 dx + \alpha \int_a^b y^2(s)ds$$

где $\alpha > 0$ -- параметр регуляризации. Этот функционал является регуляризующим, если $\alpha = C(\Delta f)^2$,

где $C > 0$ – некоторая константа, а Δf -- ошибка задания $f(x)$, равная

$$\Delta f = \|\tilde{f}(x) - \bar{f}(x)\| = \sqrt{\int_c^d [\tilde{f}(x) - \bar{f}(x)]^2 dx},$$

где $\bar{f}(x)$ -- точная правая часть, а $\tilde{f}(x)$ -- ее практическое задание.

Из условия минимума функционала Φ_α следует уравнение Эйлера

$$\alpha y(x) + \int_a^b k(x,s)y(s)ds = w(x), a \leq x \leq b, \quad (*)$$

где

$$k(x,s) = \int_c^d K(t,x)K(t,s)dt$$

$$w(x) = \int_c^d K(t,x)f(t)dt$$

В результате вместо исходного уравнения первого рода, надо решать уравнение (*), которое является уравнением второго рода, а его решение при оптимальном выборе значения α является устойчивым.

Наиболее сложным этапом в методах регуляризации является определение параметра регуляризации α . Нужно определить такое значение α , которое, с одной стороны делало бы решение устойчивым, а с другой стороны не сильно искажало бы первоначальное уравнение первого рода.

Для определения такого значения α имеется ряд способов. Например, можно использовать алгоритм выбора квазиоптимального $\alpha_i = \alpha_{opt}$, при котором

$$\|y_{\alpha_{i+1}} - y_{\alpha_i}\| = \min, \text{ где } \|y\| = \max |y(s)|, a \leq s \leq b, \alpha_{i+1} = \delta \cdot \alpha_i, 0 \leq \delta \leq 1, y_{\alpha_i} \equiv y_{\alpha_i}(s).$$

В расчетах использовался метод определения α , основанный на знании исходного (неискаженного) сигнала. Выбирается такой $\alpha_i = \alpha_{opt}$, при котором

$$\|y_{\alpha_i} - y_{исх}\| = \min, \text{ где } \|y\| = \sum |y(s)|, a \leq s \leq b, \alpha_{i+1} = \delta \cdot \alpha_i, 0 \leq \delta \leq 1, y_{\alpha_i} \equiv y_{\alpha_i}(s).$$

Целью численного эксперимента являлось исследование увеличения разрешения по данной спектрограмме, т.е. отношение минимально различимого периода искаженного сигнала к минимально различимому периоду восстановленного сигнала. В качестве исходного сигнала $y(s)$ была взята функция $\sin(x^2)$, а в качестве ядра – гауссоида

$$K(x,s) = \frac{1}{\sqrt{2\pi a}} e^{-\frac{(x-s)^2}{2a^2}}, \text{ где } a - \text{область финитности ядра.}$$

Для получения различных минимально различимых периодов изменялась область финитности ядра – чем она больше, тем больше искажения и минимально различимый период.

Типичная картинка представлена на рисунке.

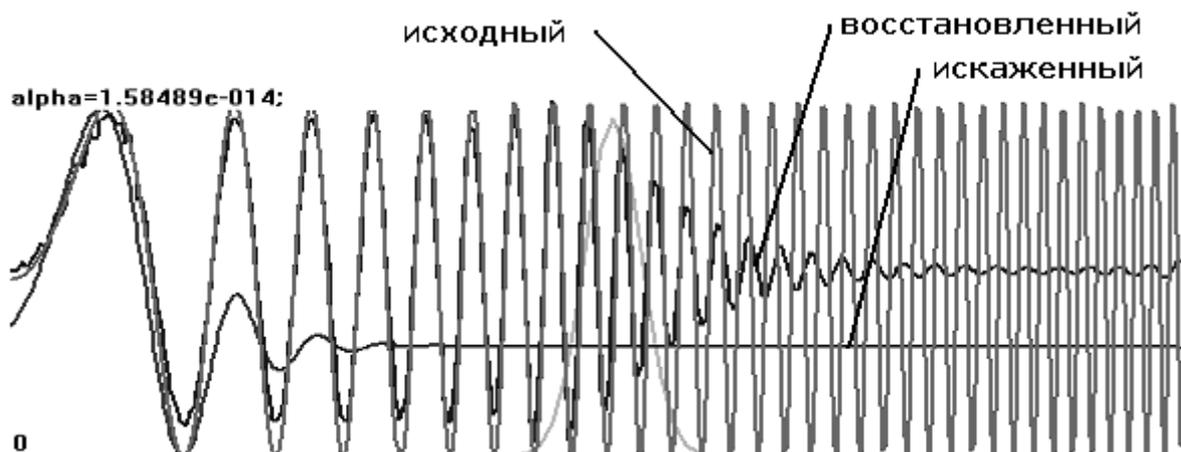


Рис. 3. Результат работы программы.

После анализа результатов для разной величины ширины ядра было получено, что в среднем искомое отношение (сверхразрешение) равно 2,9. То есть мы можем восстановить частоту почти в три раза большую, чем различимую в имеющемся искаженном сигнале.

Также исследовалось время работы метода при уменьшении шага. Было обнаружено, что с ростом числа разбиений количество используемых ресурсов ЭВМ растет экспоненциально. Это проиллюстрировано на рисунке.

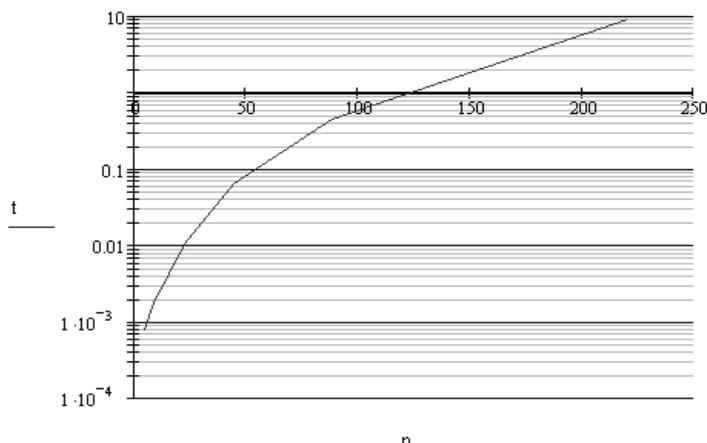


Рис. 4. График зависимости времени выполнения от числа разбиений

Для удобства ввода данных и просмотра результатов к алгоритму была создана программа с графическим (под Windows) интерфейсом.

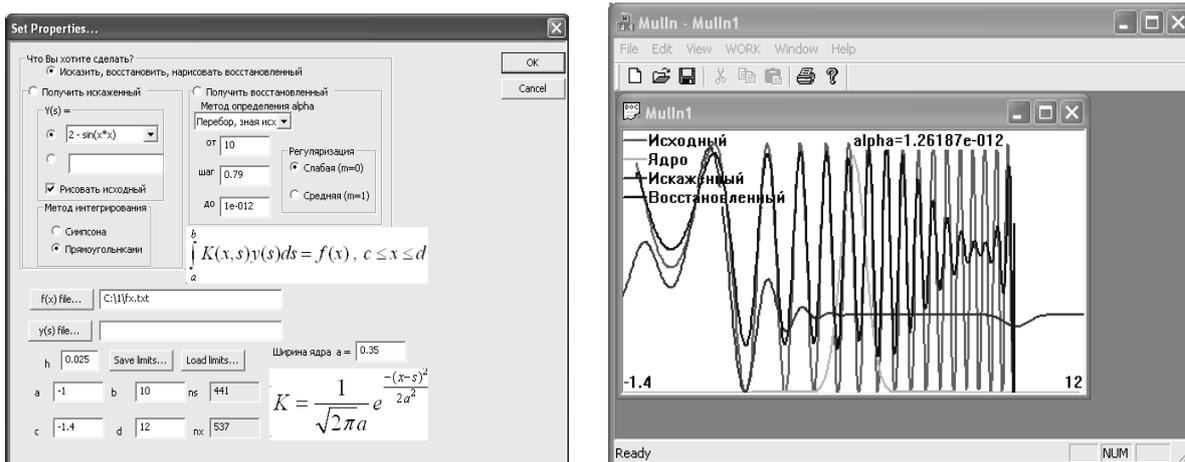


Рис. 5. Интерфейс программы.

Такая программа хоть и требует времени на разработку, но значительно сокращает время обработки результатов, поэтому ее написание вполне оправдано.

В дальнейшем планируется улучшить производительность, отказавшись от метода Гаусса для решения СЛАУ. Также планируется переход к регуляризации в частотной области, используя интеграл свертки и преобразование Фурье, что должно привести к улучшению точности и уменьшению времени вычислений.

ДВУХУРОВНЕВОЕ МОДЕЛИРОВАНИЕ ОПТИКОЭЛЕКТРОННЫХ СИСТЕМ

Трудоношин И.В.

Научные руководитель: к.т.н, доцент кафедры РК-6, Волосатова Т.М.

МГТУ им.Баумана," РК-6", Москва, Россия

TWO-LEVEL MATHEMATICAL SIMULATION OF OPTOELECTRONIC SYSTEMS

Trudonoshin I.V.

The scientific chief: Ph.D., senior lecturer Volosatova T.M.

BMSTU," RK-6", Moscow, Russia

rk6trud@wwcdl.bmstu.ru, <http://wwcdl.bmstu.ru>

Аннотация

Рассматривается подход к математическому моделированию сложных технических систем, содержащих в своем составе электронную и оптическую подсистемы. Поскольку для моделирования таких подсистем используется различный математический аппарат, требуется принятие специальных мер при моделировании в рамках одного расчета. Модели оптических компонентов системы реализованы для комплекса ПА9.

Abstract

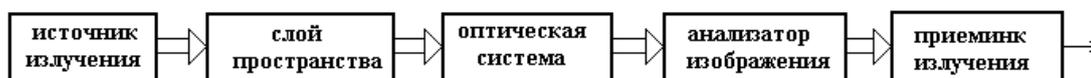
There are many technical systems, which consists of electrical and optical parts. Mathematical models and methods of analyses are different for this parts, and this make troubles to simulate such systems during one computing process. Algorithms and models to make such simulation possible using PA9 was developed.

Оптикоэлектронным прибором (ОЭП) называется прибор, в котором обработка входного сигнала осуществляется как оптическими, так и электронными системами. Переход от обработки оптического сигнала к обработке электрического требует анализа изображения. Входной сигнал ОЭП – всегда многомерный оптический, а выходной – электрический или оптический, в зависимости от типа ОЭП(информационный или измерительный).

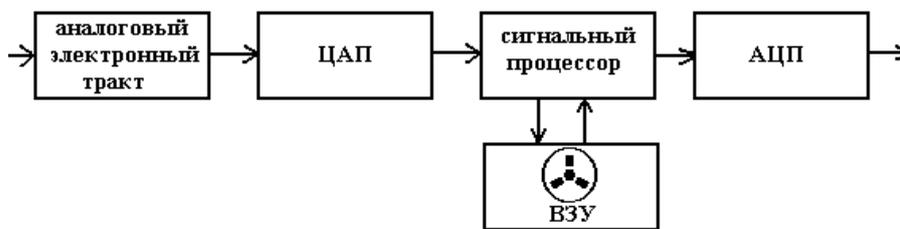
Приемная оптическая система ОЭП преобразует излучение от объекта наблюдения, фонов, помех, которое проходит через слой пространства и поступает на входной зрачок оптической системы. Изображение, построенное оптической системой, модулируется подвижными или неподвижным растром. В результате модуляции на чувствительную площадку приемника излучения поступает переменный во времени поток излучения. Приемник излучения представляет этот поток в виде изменяющегося во времени электрического тока или напряжения.

Образованный приемником излучения электрический сигнал обрабатывается в электронном тракте, который состоит из аналоговой части и, при необходимости, цифровой вычислительной машины. На современном этапе развития техники, выделенный в отдельное устройство аналоговый электронный тракт может отсутствовать. В цифровом электронном тракте обработка информации осуществляется цифровым сигнальным процессором, и поэтому в нем обязательно присутствует аналого-цифровой преобразователь, приводящий сигнал с приемника излучения к виду, пригодному для обработки цифровым процессором. В информационных ОЭП с цифровым процессором, также присутствует цифро-аналоговый преобразователь.

Общий вид оптической части ОЭП как объекта проектирования представлен на рисунке.



На выходе приемника излучения присутствует сигнал $U(t)$, который является входным для электронной части прибора:



В измерительных ОЭП обработка информации заканчивается в цифровом сигнальном процессоре, которая выдается в виде результатов измерений пользователю, или запоминается в запоминающих устройствах для дальнейшей обработки.

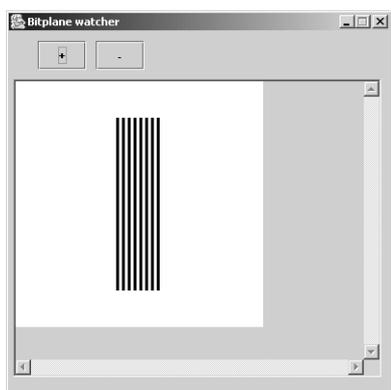
В информационных ОЭП за обработкой сигнала в электрической части ОЭП следует возврат к обработке пространственного сигнала для отображения пользователю прибора на экране.

Работа любого ОЭП невозможна без наличия объекта или совокупности объектов – источников излучения. В модели ОЭП источник излучения рассматривается как источник многомерного оптического сигнала, несущего в себе информацию о состоянии объекта. В соответствии в обобщенной модели оптического тракта ОЭП источник излучения должен описывать яркостное поле $L_\lambda(x, y, z, \lambda, t, \theta, \gamma, \varphi)$. Однако в этой модели применяются некоторые ограничения:

- Источник считается неизменным во времени (исключается t).
- Считается, что фронт волны на входе оптической системы – плоский. (исключается z)
- Считается, что конструктивные особенности источника не приводят к изменению спектрального состава в зависимости от расположения проектируемого прибора (исключаются угловые характеристики θ, γ, φ)

С учетом этих ограничений модель источника излучения выглядит следующим образом: $L_\lambda(x, y) = L_0(\lambda) * L^{OTH}(x, y)$, где L_0 - спектральная плотность яркости источника, учет которой условно переносится в модель приемника излучения, а L^{OTH} - распределение яркости.

Модель источника излучения была реализована для комплекса ПА9 в виде блок-оператора. Она позволяет генерировать параметрически заданные изображения штриховой миры, а так же использовать в качестве входного сигнала произвольное изображение из файла. Пример представлен на рисунке.

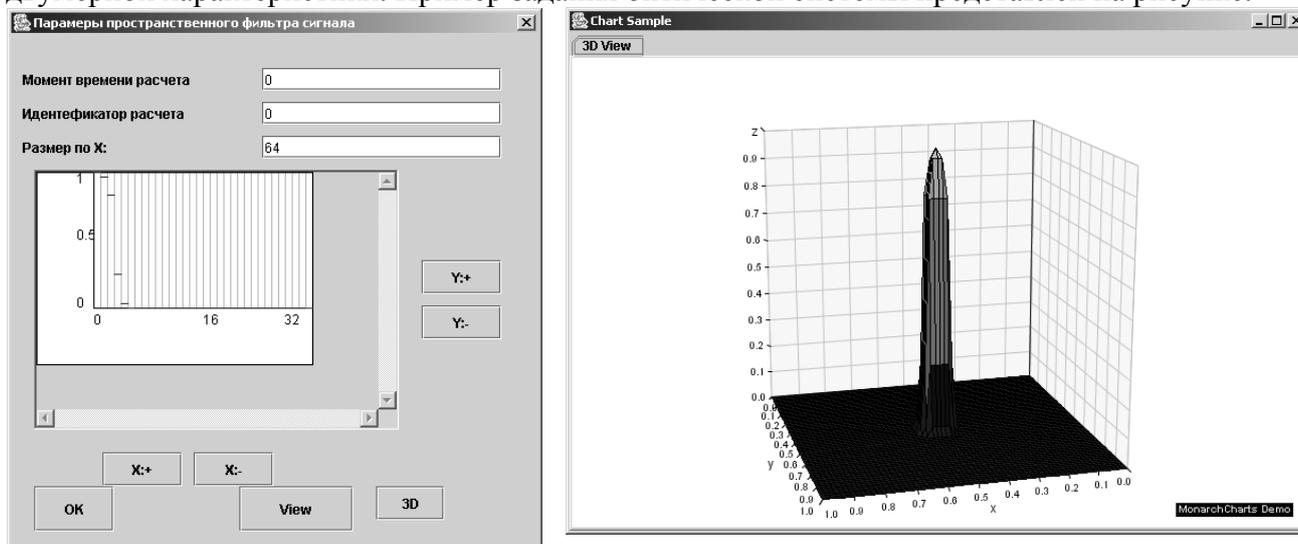


Модель оптической системы на системотехническом уровне проектирования должна отражать общие закономерности преобразования оптического сигнала оптическими системами. Оптическая система, как звено оптико-электронного тракта, преобразует оптический сигнал пространства предметов в распределение поля в пространстве

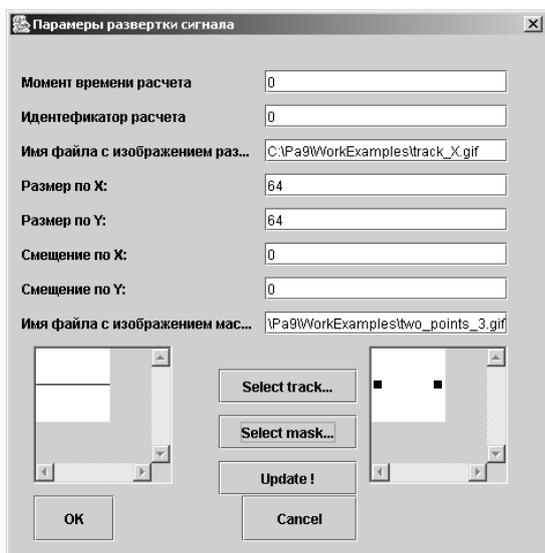
анализа на входе анализатора изображения. Модель оптической системы должна учитывать масштабные преобразования, осуществляемые оптической системой, т. е. передачу размеров и формы поля предметов в изображении; фильтрующее действие оптической системы, т. е. передачу тонкой структуры поля предметов; энергетические преобразования (спектральные и интегральные) оптического сигнала. Преобразующее действие оптической системы зависит от вида входного воздействия: когерентное, некогерентное, частично когерентное. На системотехническом уровне оптическая система определяется совокупностью параметров

и характеристик, выражающих ее действие в оптико-электронном тракте и позволяющих производить ее синтез, анализ и параметрическую оптимизацию.

Модель оптической системы была реализована для комплекса ПА9 в виде блок-оператора. Параметром модели являются импульсный отклик оптической системы, заданный в виде двумерной характеристики. Пример задания оптической системы представлен на рисунке.



Характерной особенностью любого ОЭП как объекта проектирования является наличие подсистемы, с помощью которой выделяется и особым образом кодируется та часть информации об источнике излучения, которая определяет функциональное назначение ОЭП. Такой подсистемой может быть подвижный и неподвижный растры, оправа приемника лучистой энергии, мозаика фоторезисторов и т. п. В вырожденном случае - это неподвижная диафрагма и стоящий непосредственно за ней приемник лучистой энергии. В этой же подсистеме присутствует анализатор изображения - развертывающее устройство, характеризующее некоторым коэффициентом пропускания τ и законом перемещения в поле анализа изображения, а также устройство, осуществляющее преобразование многомерного сигнала в одномерный без искажений во временной координате. Таким устройством может быть, например, безинерционный фотоприемник. В этом случае можно считать, что на вход анализатора изображения поступает сигнал в виде распределения освещенности, создаваемого либо оптической системой, либо слоем пространства.



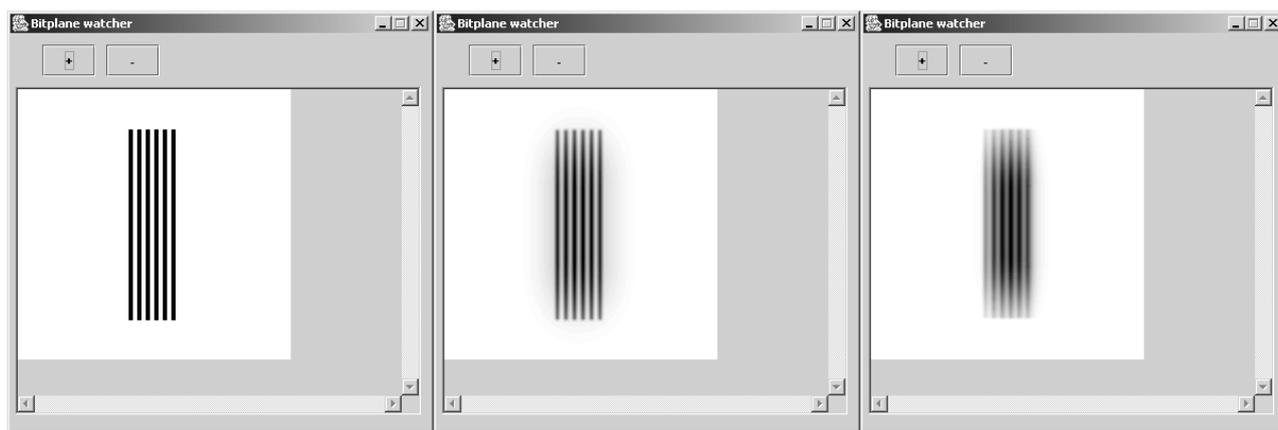
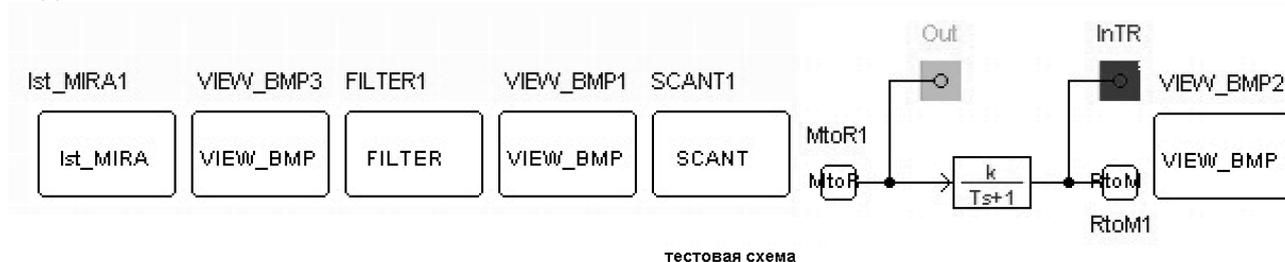
В последнее время все большее внимание уделяется цифровой обработке изображений, причем анализ изображения осуществляется по одной из следующих схем.

Изображение кодируется по закону $x'(t)$, $y'(t)$ в виде аналогового сигнала $\Phi(t)$, а затем с помощью аналогово-цифрового преобразователя формируется цифровой сигнал Φ_i в виде последовательности отсчетов $\Phi(t)$ через равные интервалы времени Δt . Осуществляется дискретизация изображения, затем цифровая обработка двумерного массива отсчетов. Реализацию на ЭВМ модельного представления анализатора изображения можно использовать и для моделирования цифровых и аналоговых

анализаторов изображения.

Модель анализатора изображения была реализована для комплекса ПА9 в виде блок-оператора. Параметрами модели являются траектория (закон) сканирования и форма пятна сканирования. Пример задания параметров модели представлен на рисунке.

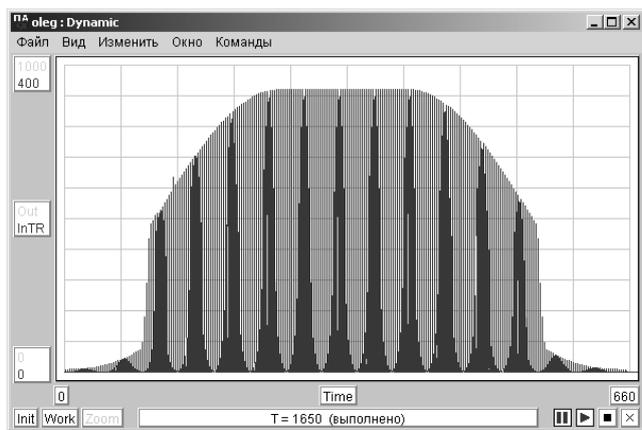
На рисунке ниже приведен пример совместного моделирования оптической и электронной подсистем.



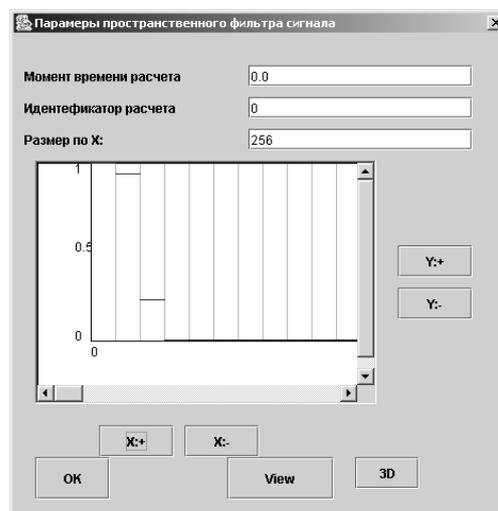
входное изображение

изображение после пространственного фильтра

изображение после электронного тракта



видеосигнал в электронном тракте



параметры пространственного фильтра

Литература

1. «Автоматизация проектирования оптико-электронных приборов», под. ред. Л.П.Лазарева, Машиностроение, 1986
2. «Автоматизация проектирования цифровых устройств», под. ред. С.С. Бадудина, Радио и связь, 1981
3. «Основы теории преобразования сигналов в оптико-электронных системах», Порфирьев Л.Ф., Машиностроение, 1989
4. «Теория преобразования сигналов в оптико-электронных приборах», издательство МВТУ, Машиностроение, 1977

МЕТОД АНАЛИТИЧЕСКОЙ НАСТРОЙКИ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧЕ КОНТЕКСТНОЙ КЛАССИФИКАЦИИ ТЕКСТОВ

Цыганов И.Г.

научный руководитель: к.т.н. Власов А.И.

Московский Государственный Технический Университет им. Н.Э. Баумана,
кафедра “Проектирование и технология производства ЭА”

THE ANALYTICAL TRAINING METHOD OF THE NEURAL NETWORKS IN CONTEXT-SENSITIVE TEXT CATEGORIZATION TASKS

Tsyganov I.G.

the scientific chief: Ph.D. Vlasov A.I.

Bauman Moscow State Technical University,
Department of Design and Equipment of Electronic System.

Аннотация.

Рассмотрен метод построения классификатора текстовых данных на основе логической декомпозиции задачи классификации. Предложен вариант детерминированной аналитической настройки нейронной сети по методу наименьших квадратов с использованием псевдообратной матрицы, получаемой на основе сингулярного разложения матрицы примеров (SVD). Выявлены некоторые недостатки метода. Предложены направления дальнейшего развития метода.

Abstract.

The method of classifier synthesis by the logical task decomposition is considered. The deterministic analytical variant of neural network training by the least-square fitting method is proposed. The method is based on the pseudoinverse matrix obtained by the singular value decomposition (SVD) of term-document matrix. Some disadvantages of the method is detected. The future work direction is proposed.

Введение

Успехи последнего десятилетия в области технологии производства цифровых вычислительных машин и средств телекоммуникаций, приведшие к значительному удешевлению хранения, обработки и передачи информации, сделали реальным и послужили основной причиной для крупномасштабного перехода от традиционных способов представления данных и знаний на бумажных носителях к форме электронных документов. Рост объемов и разнообразия данных, представленных на естественном языке в цифровой форме, заставляет в свою очередь разрабатывать автоматизированные системы организации и управления текстовыми данными как в стационарных задачах библиотечного типа, так и в динамических, при организации интерактивной передачи потоков текстов. Для решения задачи организации текстовых данных широко используются методы категоризации текстов по контекстным признакам.

Существует большое число методов контекстной категоризации текстов, построенных на различных принципах. Например, известны методы на основе деревьев решений [1], правил решений [2], простых Байесовых классификаторов (Naïve Bayes) [3], SVM [4] и т.д. Среди всех существующих методов особое место занимают методы на основе построения нейронных сетей [5,6]. Нейросетевой подход в общем случае оказывается более гибким, поскольку позволяет при изменении числа слоев в сети построить достаточно сложные нелинейные модели. С другой стороны, усилия большинства исследователей сейчас сосредоточены на достаточно простых линейных on-line моделях, таких как Positive Winnow, Balanced Winnow, однослойный перцептрон и т.п. Причинами этого является, во-первых, значительная сложность настройки многослойных сетей, потеря устойчивости алгоритмов при настройке большого числа параметров, во-вторых, лишь незначительный рост

эффективности распознавания текстов при введении ограниченного числа нейронов в дополнительных слоях.

При этом, однако, хорошо известно, что линейные модели в области текстовой обработки являются лишь очень грубой аппроксимацией. Это, например, показано в работе [7], где указано, что векторы текстовых фрагментов в семантическом пространстве признаков, относящиеся к одной и той же категории могут образовывать разрозненные кластеры, что не позволяет применять простую линейную модель в таких задачах. В связи с этим, преимущества нелинейных моделей, построенных с применением нейросетевого базиса становятся совершенно очевидными.

Оптимальный выбор структуры и настройка параметров (весовых коэффициентов) многослойной нейронной сети в общем случае представляет собой достаточно сложную задачу. Ее сложность в основном определяется большим числом факторов, которые должны быть учтены, разряженностью входного сигнала, многоэкстремальностью глобального критерия оптимизации, и многими другими факторами. Поэтому, в данной работе, мы разбиваем общую задачу глобальной минимизации глобального критерия оптимальности на ряд локальных критериев в соответствии с логической декомпозицией структуры задачи на отдельные блоки. В соответствии с указанной декомпозицией общая глобальная задача оптимизации сводится к совокупности однотипных локальных задач оптимизации. Решение каждой локальной задачи может быть получено аналитически, что позволяет сократить время на настройку нейронной сети, а также устраняет проблемы, связанных с устойчивостью алгоритмов. Далее рассматривается глобальный критерий как функция от набора параметров и констант, полученных в результате оптимального выбора с использованием локальных критериев оптимальности. Оптимизация глобального критерия по множеству варьируемых параметров приводит к оптимальному результату для задачи в целом. Указанный подход является аппроксимацией, основанной на логической структуре текстовых данных.

Данный метод основан на принципах, используемых в двух известных методах: k-NN и LLSF. Из метода k-NN заимствован принцип, в соответствии с которым при моделировании свойств категорий используются контекстные свойства отдельных текстовых фрагментов. Высокие результаты, которые неизменно демонстрирует метод k-NN во многих приложениях, наглядно убеждает в существенном значении указанного принципа. В другой стороны, хорошо известны недостатки методов типа k-NN, к которым, во-первых, относится существенная избыточность пространства признаков, во-вторых, отсутствие обоснованных приемов взвешивания отдельных компонент вектора текстового фрагмента. В связи с этим мы применяем метод LLSF, который позволяет значительно сократить число признаков в задаче за счет оптимального выбора коэффициентов.

Формальная постановка задачи категоризации

Рассмотрим формальную постановку задачи категоризации. Пусть имеется набор категорий $C = \{c_1, c_2, \dots, c_{|C|}\}$ и выборка текстовых фрагментов $TV = \{TV_1, TV_2, \dots, TV_{|C|}\}$, где каждая i -ая выборка состоит из текстовых фрагментов $T_{ij} \in TV_i$, предварительно отнесенных к i -ой категории. Задача построения классификатора заключается в том, чтобы на основании имеющегося множества выборок TV и множества соответствующих категорий C построить функцию, осуществляющую отображение произвольного текстового фрагмента на множество категорий: $K : T \rightarrow C$. Такое отображение должно обеспечивать заданную вероятность правильной классификации.

В данном методе производится построение отдельного классификатора для каждой из имеющихся категорий. Такой классификатор, принимая на вход представление текстового фрагмента, осуществляет взвешенную оценку того, насколько входной текстовый фрагмент соответствует данной категории. При этом рассматривается задача, в которой настройка i -го классификатора, соответствующего i -ой категории осуществляется независимо от других

категорий $C_i = \{c_1, \dots, c_{i-1}, c_{i+1}, \dots, c_{|C|}\}$. При построении i -го классификатора в качестве положительных примеров рассматривают выборку текстовых фрагментов: $TV^+ = TV_i$, а в качестве отрицательных примеров выборку вида $TV^- \subset \bigcup_{j \neq i}^j TV_j$. Для выбора отрицательных примеров можно применять метод, описанный в [8]. Этот метод позволяет выбрать из $\bigcup_{j \neq i}^j TV_j$

наиболее близкие текстовые фрагменты к текстам из TV_i . Рассматривая отдельно результаты классификации по всем классификаторам, выбирают ту категорию, что дает наибольшую оценку для данного текстового фрагмента. Мы применяем данный подход на основании того, что в нашей задаче, во-первых, текстовый фрагмент может относиться только к одной категории, во-вторых, все выходы классификаторов нормированы. Таким образом, мы свели общую задачу построения классификатора к группе из $|C|$ отдельных задач. Далее будем рассматривать только задачу построения отдельного бинарного классификатора.

Логическая декомпозиция задачи категоризации

Поскольку задача построения полной модели классификатора является достаточно сложной, т.к. она содержит большое число параметров, оптимизация на поле которых проблематична в связи с наличием многоэкстремальности, то совершенно необходимой становится некоторая аппроксимация. Для этого мы применяем принцип логической декомпозиции, разбивая классификатор на ряд иерархически подчиненных блоков, каждый из которых должен определять относительно самостоятельное и независимое семантическое образование.

В соответствии с логической структурой текстов в данном методе выделяется три уровня в иерархии классификатора (см. рис. 1). На нижнем уровне иерархии (уровень 0) находятся отдельные термины (слова). Каждое слово включено в один или несколько контекстных блоков с некоторым весом, который может быть назначен связи данного слова с данным контекстным блоком. Множество контекстных блоков определяют уровень 1. Каждый контекстный блок определяет контекстную сущность, представленную в виде совокупности входящих в нее слов. Контексты текстовых фрагментов в свою очередь включаются в блок категории (единственный блок уровня 3), а каждому контекстному блоку при этом назначается свой вес.

Один и тот же термин может быть включен в несколько контекстных блоков с

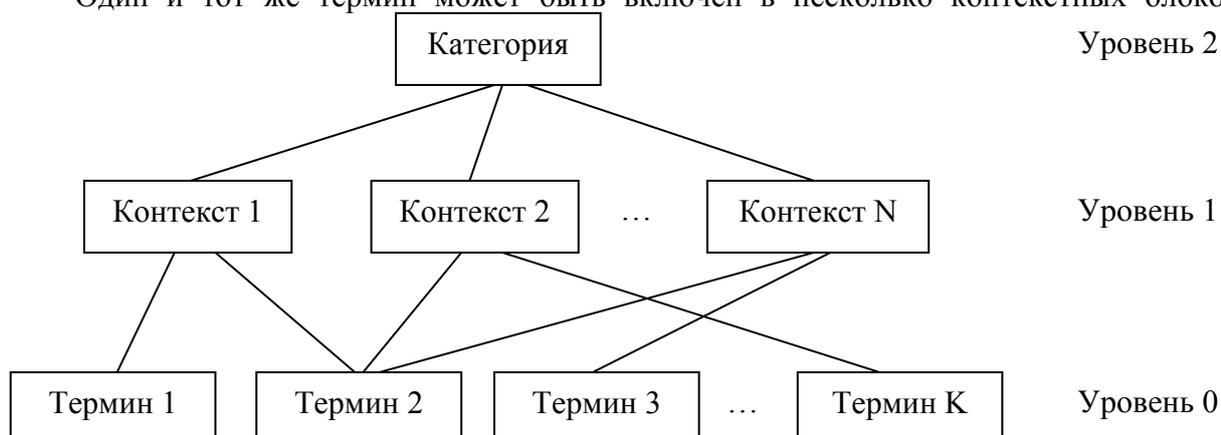


Рис.1. Контекстная иерархия категории

разными весами. В качестве контекстных блоков естественно выбрать некоторое представление имеющихся в TV^+ текстовых фрагментов. Такой подход применяется в частности в методе Expert Network [9]. При этом отдельным связям между термином и контекстным блоком соответствует вес, определяющий значимость данного термина в

данном контекстном блоке. Один и тот же термин может быть включен в несколько контекстных блоков с разными весами. В методе Expert Network, однако, веса отдельных слов назначаются на базе общих для категории правил. В предлагаемом методе веса рассчитываются в соответствии с критерием оптимальности, а число слов сокращается. Выделение контекстного слоя позволяет учесть значимость отдельных слов в различных текстовых фрагментах данной категории. Текстовые фрагменты в данной модели рассматриваются как дифференцированные смысловые единицы, обладающие самостоятельными локальными свойствами. Интуитивно, контекстный блок представляет организованное единство слов данного текстового фрагмента, что позволяет выявить взаимную значимость слов в данном текстовом фрагменте. Контекстный блок при этом представляет собой миниатюрную модель категории в перспективе набора слов данного текстового фрагмента.

Формирование пространства признаков

Ключевым элементами в структуре иерархии классификатора, определяющими качество его работы, являются элементы контекстного слоя. В соответствии с этим выбор информативных признаков в данной задаче определяется локальными свойствами текстовых фрагментов, которые необходимо рассматривать с двух позиций: их соответствие данной категории и их соответствие данному текстовому фрагменту. Поэтому для выбора информационных признаков мы используем следующий алгоритм. Вначале создается словарь V^+ , состоящий из всех слов, встречающихся в текстах выборки TV^+ , при этом из словаря исключаются общеупотребительные термины (такие как союзы, вводные слова и т.д.) по стоп-листу, и каждое слово проходит морфологическую процедуру, в процессе которой из него удаляются приставки, флексии и пр. грамматические суффиксы. Кроме словаря V^+ также создается словарь V^- , строящийся аналогично V^+ , только на основании выборки TV^- .

Далее последовательно рассматриваются отдельные текстовые фрагменты. Каждому слову в каждом текстовом фрагменте назначается вес в соответствии с применением метода $TFxIDF$. Далее все слова данного текстового фрагмента упорядочиваются в соответствии с указанным весом. После чего из полученного множества отбираются W тех, что имеют максимальный вес (W – константа, $W=10$). При этом, вначале отбираются только слова, имеющиеся в V^+ и не имеющиеся в V^- . Если отобранное число слов меньше, чем заданное W , то отбираются прочие слова с максимальным весом.

Таким образом составляются отдельные словари текстовых фрагментов V_i^+ , $i=1, \dots, |TV^+|$, и общий словарь категории $V = \bigcup_i V_i^+$. Отметим, что $|V_i^+| \leq W$. Далее каждый текстовый фрагмент представляется в виде вектора: $x = (x_1, x_2, \dots, x_N)$, где $N=|V|$ – число слов в словаре, x_i – компонента, соответствующая весу, рассчитанному по методу $TFxIDF$ для i -го слова в словаре V . Данный вектор является входным для алгоритма настройки классификатора. Поскольку при подобном “агрессивном” снижении размерности пространства признаков отдельные текстовые фрагменты в словарях V_i будут иметь одинаковое представление, то перед началом работы алгоритма настройки, необходимо удалить все повторяющиеся текстовые фрагменты из выборки TV^+ .

Описание алгоритма настройки

Приведенная выше иерархия блоков естественным образом отображается на трехслойную нейронную сеть. Первому (распределительному) слою соответствуют признаки задачи (слова в словаре V). Второй слой строится из “ассоциативных” нейроподобных элементов (НЭ) контекстов отдельных текстовых фрагментов, последний “командный” НЭ определяет окончательное решение.

Для моделирования каждого контекстного блока используем один НЭ. Количество входов каждого НЭ контекстного слоя определим равным числу слов в словаре V_i , соответствующем данному текстовому фрагменту. Количество НЭ в контекстном слое, таким образом, определяется числом позитивных примеров текстовых фрагментов в обучающей выборке данной категории.

В данном методе производится поэтапный синтез нейронной сети. Вначале определяются параметры НЭ контекстного слоя в соответствии с введенным ниже критерием, после чего определяются параметры единственного НЭ уровня категории. Каждый НЭ контекстного слоя настраивается независимо, что позволяет реализовать параллельную работу алгоритма настройки контекстного слоя. Все НЭ настраиваются в соответствии с одним и тем же критерием.

В качестве функции активации в данной задаче естественно использовать линейную ограниченную функцию активации (или функцию активации R [10]). Каждый вход НЭ имеет ассоциированный с ним вес. Если число входов n , то взвешенная сумма НЭ может быть представлена как

$$S = \sum_{i=1}^n a_i x_i,$$

или, то же самое в матричной записи:

$$S = A \cdot X,$$

где a_i – коэффициент (синапс) i -го входа ($i=1, \dots, n$), x_i – компоненты входного вектора, A – вектор-строка $1 \times n$, X – вектор столбец $n \times 1$. Задача построения нейросетевой модели контекста заключается в оптимальном выборе коэффициентов входов НЭ. Рассматривая каждый контекстный НЭ отдельно, оптимальный выбор коэффициентов определяется минимизацией среднеквадратической ошибки на множестве обучающей выборки на данном НЭ:

$$R = \sum_{j=1}^l e_j^2 = \sum_{j=1}^l \left[\sum_i a_i x_i^j - c^j \right]^2 \rightarrow \min, \quad (1)$$

где e_j^2 – квадрат ошибки на j -ом обучающем примере, l – общее число примеров текстовых фрагментов в обучающей выборке; x_i^j – компонента вектора j -го текстового фрагмента, соответствующая i -му входу данного НЭ, c^j – величина, определяющая желаемое выходное значение для данного нейрона, на j -ом обучающем примере.

В простейшем случае c^j – это бинарная функция, заданная на множестве обучающей выборки:

$$c^j = \begin{cases} 1, & \text{если } T_j \in T_{c_i} \\ 0, & \text{если } T_j \notin T_{c_i} \end{cases}, \quad (2)$$

более общий случай может быть представлен следующим образом:

$$c^j = \begin{cases} k^+, & \text{если } T_j \in T_{c_i} \\ k^-, & \text{если } T_j \notin T_{c_i} \end{cases}, \quad (3)$$

где k^+ – значение, вокруг которого (в соответствии с (1)) будут группироваться выходы НЭ для положительных примеров, k^- – то же для отрицательных примеров.

В наших экспериментах также будем рассматривать следующий вид функции c^j :

$$c^j = \begin{cases} k^+ \frac{n_j}{m_j}, & \text{если } T_j \in T_{c_i} \\ k^- \frac{n_j}{m_j}, & \text{если } T_j \notin T_{c_i} \end{cases}, \quad (4)$$

где n_j – число отличных от нуля компонент вектора x j -го текстового фрагмента, m_j – общее число слов в j -ом текстовом фрагменте. Пропорция в (4) позволяет при минимизации (1)

повысить роль тех примеров, которые наиболее полно “вписываются” в контекст данного НЭ. Например, если в текстовом фрагменте j -го примера нет компонент, отличных от тех, которые соответствуют входам данного НЭ, то пропорция равна k , если это не выполнено, то c^j равна только соответствующей части.

Функция c^j таким образом связана как с текстовым фрагментом, контекст которого представляет данный НЭ, так и с категорией, в модели которой он рассматривается. Чем больше соответствие j -го текстового примера и данного текстового фрагмента в данной категории, тем большие значения данная функция принимает.

Рассмотрим более подробно зависимость (1). В общем случае, когда в обучающей выборке могут быть повторяющиеся текстовые фрагменты или линейно зависимые векторы, решение (1) не может быть получено с помощью прямого применения стандартных приемов дифференцирования по весовым коэффициентам. Поэтому для ее решения удобно использовать хорошо зарекомендовавший себя в области текстовой категоризации прием с получением псевдообратной матрицы на основании сингулярного разложения *SVD* (*singular value decomposition*) [1].

Рассматривая зависимость (1) в матричной форме, запишем:

$$R = \|A \cdot L - C\|^2 \rightarrow \min, \quad (5)$$

где L – матрица ($n \times l$) обучающих примеров:

$$L = (x^1 | x^2 | \dots | x^l) = \begin{pmatrix} x_1^1 & \dots & x_1^l \\ \dots & & \dots \\ x_n^1 & \dots & x_n^l \end{pmatrix},$$

C – матрица-строка ($1 \times l$): $C = (c^1, c^2, \dots, c^l)$, компоненты которой рассчитываются согласно (2), (3), или (4); $\|v\| = \sqrt{\sum_i v_i^2}$ – норма вектора.

Для решения (5) производят следующие действия:

- 1) если $n \geq l$ производят разложение $L = USV^T$,
- 2) если $n < l$, тогда раскладывают $L^T = VSU^T$,

где U – матрица размерности $n \times p$, V – матрица размерности ($l \times p$), содержит левые и правые сингулярные векторы; S – диагональная матрица ($p \times p$), которая содержит p не нулевых сингулярных значений: $s_1 \geq s_2 \geq \dots \geq s_p > 0$, $p \leq \min(l, n)$.

Далее получают матрицу $A = CVS^{-1}U^T$, где

$$S^{-1} = \begin{pmatrix} 1/s_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 1/s_p \end{pmatrix}$$

При реальных экспериментальных исследованиях, однако, выяснилось, что применение данного метода в чистом виде может приводить к тому, что некоторые коэффициенты a_i будут принимать достаточно большие значения и иметь противоположный знак. Анализ ситуаций, когда наблюдалось такое поведение алгоритма, показал, что во многих случаях это происходит когда какие-либо два (или более) слова во всех имеющихся примерах встречаются только вместе и никогда порознь. В этом случае, согласно (5), алгоритм выбирает для одного из этих слов положительное значение, а для другого – отрицательное, чем достигается минимум. Для устранения этого побочного эффекта мы исключали из обучающей выборки все избыточные компоненты. Т.е. последовательно исключалась i -ая компонента $\exists(i, j), j \neq i: \forall x \in TV(x_i > 0 \Leftrightarrow x_j > 0)$. Веса указанных двух компонент всегда совпадают, поскольку они встречаются только совместно, а значит с одинаковой частотой в одних и тех же обучающих примерах. При удалении i -ой компоненты, j -ое значение в векторах x умножается на два. Это можно продемонстрировать на следующем примере:

$$\sum_k a_k x_k - c = \sum_{k;k \neq i; k \neq j} a_k x_k + a_i x_i + a_j x_j - c = \sum_{k;k \neq i; k \neq j} a_k x_k + (a_i + a_j) x_j - c,$$

поскольку обе компоненты не различимы и равнозначны, то не только $x_i = x_j$, но и соответствующие веса также должны совпадать $a_i = a_j$, тогда:

$$\sum_k a_k x_k - c = \sum_{k;k \neq i; k \neq j} a_k x_k + a_j (2x_j) - c,$$

из чего следует, что, для того, что решение (5) не изменилось при удалении i -ой компоненты, достаточно во всех примерах соответствующие j -ые компоненты умножить на 2.

Помимо описанного частного случая, существует и другие варианты комбинаций параметров, которые в совокупности приводят к описанному побочному эффекту. Несмотря на то, что формально в соответствии с (1) решение выбрано оптимальным, результат на текстах, которые не представлены в обучающей выборке с такими значениями параметров может приводить к некорректным значениям.

Заключение

Одним из действенных вариантов сокращения сложности общей задачи построения классификатора является метод ее декомпозиции на ряд более простых задач. В данной работе мы привели один такой метод, ориентированный на учет нелинейностей в контекстной структуре категорий.

В данном методе используется детерминированный выбор коэффициентов нейронной сети в соответствии с принципом наименьших квадратов. Ряд экспериментальных исследований выявил недостатки данного метода, которые связаны с тем, что формальная минимизация без ограничений на изменение параметров может приводить к тому, что расчет некоторых коэффициентов приводит к слишком большим противоположным значениям. Этот эффект, приводя к минимуму на имеющейся совокупности текстовых примеров, однако, на новых текстовых примерах может приводить к неоправданно высоким (или неоправданно низким) оценкам.

Поэтому, очевидным развитием данного метода должна быть минимизация с ограничениями на параметры вида $a_i \geq 0$.

Литература

1. Li, Y. H., Jain, A. K. 1998. Classification of text documents. *Comput. J.* 41, 8, 537–546.
2. Cihem, W. W. AND Singer, Y. Contextsensitive learning methods for text categorization. *ACM Trans. Inform. Syst.* 17, 2, (1999) 141–173.
3. Koller D., Sahami, M. 1997. Hierarchically classifying documents using very few words. In *Proceedings of ICML-97, 14th International Conference on Machine Learning* (Nashville, TN, 1997), 170–178.
4. Joachims, T. Transductive inference for text classification using support vector machines. In *Proceedings of ICML-99, 16th International Conference on Machine Learning* (Bled, Slovenia, 1999), 200–209.
5. Schutze, H., Hull, D. A., Pedersen, J. O. A comparison of classifiers and document representations for the routing problem. In *Proceedings of SIGIR-95, 18th ACM International Conference on Research and Development in Information Retrieval* (Seattle, WA, 1995), 229–237.
6. Dagan, I., Karov, Y., Roth, D. 1997. Mistakedriven learning in text categorization. In *Proceedings of EMNLP-97, 2nd Conference on Empirical Methods in Natural Language Processing* (Providence, RI, 1997), 55–63.
7. Sebastiani F. Machine Learning in Automated Text Categorization. *ACM Computing Surveys*, Vol. 34, No. 1, March 2002, pp. 1–47.
8. Marti Hearst, Jan Pedersen, Peter Pirolli, H. Schutze, G. Grefenstette and D. Hull. Xerox TREC4 site report. In *Proceedings of the Fourth Text Retrieval Conference TREC-4, 1996*.
9. Yang, Y. Noise reduction in a statistical approach to text categorization. In *Proceedings of SIGIR-95, 18th ACM International Conference on Research and Development in Information Retrieval* (Seattle, WA, 1995), 256–263.

МЕТРИКИ ТЕКСТОВ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ ОБРАБОТКИ ИНФОРМАЦИИ

Цыганов И.Г., Руденко М. И.

научный руководитель: к.т.н. Власов А.И.

Московский Государственный Технический Университет им. Н.Э. Баумана,
кафедра “Проектирование и технология производства ЭА”

THE TEXT METRICS IN AUTOMATED INFORMATION PROCESSING SYSTEMS

Tsyganov I.G., Rudenko M.I.

the scientific chief: Ph.D. Vlasov A.I.

Bauman Moscow State Technical University,
Department of Design and Equipment of Electronic System.
e-mail: levent@iu4.bmstu.ru

Аннотация.

В работе проведен подробный анализ основных методов представления текстовых документов в современных системах обработки информации. Рассмотрены различные эвристические методы сокращения размерности пространства признаков. Проанализированы некоторые экспериментальные результаты, полученные при оценке эффективности различных методов.

Abstract.

In the paper the in-depth analysis of text document representation methods in automated information processing systems is proposed. Vast number of heuristic methods of feature space dimension reduction is considered. Some experimental results obtained in evaluation of methods is analyzed.

Введение

Характерной чертой современного этапа развития науки и техники является стремительный рост объемов используемой и обрабатываемой информации, важнейшим видом которой являются данные, представленные на естественном языке. При этом в последнее десятилетие отмечается переход от традиционных способов хранения и передачи текстовых данных в виде бумажных носителей к цифровой форме. Рост и укрупнение массивов доступной текстовой информации поставил множество задач, связанных с управлением имеющимися данными и организацией доступа к ним. В частности, наиболее часто рассматриваются задачи информационного поиска (information retrieval), добычи данных (text mining), автоматизации аннотирования, автоматизации перевода, фильтрации, адаптивного сжатия, автоматического упорядочивания (кластеризации), категоризации текстов по рубрикам и т.д. Во всех этих задачах на начальном этапе решается задача кодирования текстовых данных в вид, понятный для машины, после чего кодированная информация поступает на вход системы обработки.

Способ и используемая метрика кодирования текстовой информации имеет существенное влияние на качество системы текстовой обработки в целом. В процессе кодирования обычно часть информации, имеющейся в структуре текста, удаляется. Это связано, во-первых, с потребностью снижения размерности задачи, во-вторых, с потребностью удаления “шума”. В различных задачах выбор метрики и способа кодирования определяется в зависимости от того, какое именно подмножество элементов структуры естественного языка является в ней существенным и важным. В задаче категоризации текстов, например, часто отказываются от учета последовательности слов, тогда как в задаче автоматизации перевода, последовательность является важной.

В данной работе мы анализируем различные способы кодирования текстовой информации и рассматриваем различные метрики представления текстов в автоматизированных системах обработки текстовой информации.

1. Метрики текстов

Для каждой метрики текстов в общем случае необходимо определить множество объектов (структурных единиц), типы связей между ними, а также операции над этим множеством по его созданию и трансформации [1-20].

1.1. Структурные единицы

В различных работах предлагалось несколько подходов к выбору объектов. Наиболее распространенным является выбор в качестве таких объектов слов (чаще всего их нормальных форм) из текста анализируемых документов.

В некоторых исследованиях за основу брали более крупные семантические части, как: клаузы, фразы (предложения, либо статистически устойчивые сочетания слов). В работах [15-17] было показано, что качество анализа при таком выборе практически не отличается от результатов подхода, единицами которого являются слова. Lewis [17] объясняет это тем, что, несмотря на более высокое «семантическое качество» таких объектов, другие их характеристики, такие как: большое количество терминов, синонимов, меньшая частота встречаемости в коллекции документов - существенно уступают характеристикам слов.

В работе [18] был предложен подход, являющийся, в некотором смысле, комбинацией двух предыдущих. Объектами являлись именные фразы, на которые был наложен статистический фильтр (встречаемость в положительных примерах категории документов коллекции не менее трех раз).

Использование дополнительной информации о типах связей между объектами, например: семантические роли слов в фразах и сверхфразовых единствах, отношения вида «пояснение», «пример», «антитезис» и т.п. - имеет смысл только при использовании клауз, или фраз как структурных единиц.

Операции над полем выбранных объектов можно разделить на два типа: методы кодирования документов и методы получения меры взвешивания слов. Рассмотрим каждый из типов поподробнее в указанном порядке.

1.2. Методы кодирования документов

1.2.1 Простая векторная модель

В простой векторной модели [1], предложенной в конце 80-х годов Салтоном, текст документа представляется в виде вектора, каждый компонент которого соответствует одному слову словаря.

Значение компонента – это функция частоты встречаемости слова в документе и/или функция статистически измеренной важности слова. Ниже будут рассмотрены различные функции подобного рода.

Обозначим: e_i – единичный вектор с i -ой координатой равной единицей и нулями в остальных позициях, n_{jk} – функция частоты встречаемости и значимости слова k в документе j . Используя эти обозначения, вектор j -го документа, записывается следующим образом:

$$n_j = \sum_k n_{jk} e_k$$

1.2.2. Нормирование векторов

Во многих методах часто применяется подход, при котором после получения вектора документа его нормируют. При этом множество векторов текстовой коллекции представляется в виде точек на поверхности сферы заданного радиуса (обычно $r=1$).

Экспериментальное исследование данного подхода, приведенное в работе [20] показало, что качество работы системы автоматической обработки текстов может резко понизиться в этом случае, поскольку значимость отдельных слов начинает зависеть от общего числа слов в документе. В длинных документах количество информативных терминов не увеличивается обычно значительно, тогда как оценка их значимости уменьшается пропорционально общему числу слов в документе.

1.2.3. Скрытое семантическое индексирование

В простой векторной модели предполагается независимость терминов. Отдельные термины служат в качестве ортогонального базиса в векторном пространстве. Поскольку, реально между терминами отдельными терминами существует строгая корреляция, это допущение имеет лишь ограниченную применимость. В методе *LSI* (*Latent Semantic Indexing - LSI*) [2, 3] с использованием сингулярного разложения матриц (*singular value decomposition - SVD*) осуществляется преобразование исходного пространства простой векторной модели в векторное пространство меньшей размерности. В новом векторном пространстве исходные термины и документы описываются как линейная комбинация ортогональных (не коррелированных между собой) индексных переменных. Таким образом осуществляется сокращение избыточности первичного векторного пространства.

Формально метод *LSI* можно интерпретировать следующим способом представления документа:

$$n'_j = \sum_k n_{jk} x'_k,$$

где n_{jk} – гистограмма слов, x'_k – код, который формирует метод *LSI* для k -ого слова при исследовании совместного появления слов в документах.

Преимущество метода в возможности учета семантических связей между словами текста. Основной недостаток – большие вычислительные затраты.

1.2.4. Метод MatchPlus

Метод *MatchPlus* [4] кодирует документы как сумма векторных представлений слов. Представление нового слова рассчитывается на основе наиболее близкого уже существующего представления термина, если два термина появляются рядом в тексте.

1.2.5. Контекстное кодирование

В отличие от простой векторной модели, в контекстном кодировании, каждое конкретное слово документа представляется в форме, содержащей информацию о соседях по документу для данного слова.

Итак, обозначим множество слов, которые появились на расстоянии d от слова i в тексте как $I^{(d)}_i$ (каждое слово может появиться несколько раз). Например, набор слов, которые непосредственно следуют за словом i , обозначится $I^{(1)}_i$. Наша цель – вычислить статистическое описание этого множества $x^{(d)}_i$ и затем объединить описания, полученные для различных расстояний d . Комбинация, полученная объединением описаний, вычисленных для множества расстояний $\{d_1, \dots, d_n\}$, будет

$$x_i = \begin{bmatrix} x_i^{(d_1)} \\ \cdot \\ \cdot \\ x_i^{(d_N)} \end{bmatrix}$$

Вектор x_i – это численное описание контекстов, в которых появляется слово i .

1.2.6. Бионное кодирование

В работе [8] предложен вариант бионного кодирования текстов, которое разработано специально для использования в системах, в которых с большой долей вероятности исходных текстовый материал может содержать опiski, ошибки, различные неологизмы, бытовую лексику, намеренно искаженные слова и прочее.

Каждый документ x_j рассматривается как совокупность определенных некоторой системой правил фрагментов (предложений, абзацев и т.д.), которые называются лексемами. В свою очередь каждая лексема состоит из совокупности слов, а слово представляется в виде непустого неупорядоченного множества информационных кодов, именуемых бионами:

$$\sigma_i = \{ \beta_{i,1}, \dots, \beta_{i,k} \}.$$

1.2.7. Функциональное web-кодирование

Авторами работы[9] были проведены эксперименты, которые показали, что в функциональном отношении каждую страницу в сети Интернет можно классифицировать по таким внешним признакам, как *длина страницы (объем текста), количество рисунков, количество ссылок, процентное соотношение ссылок назад, вперед и за пределы сайта, наличие e-mail и news- ссылок* и т.п.

1.2.8. Стилистическое кодирование

Исходным положением в методах стилистического кодирования является зависимость стиля речи от выполняемой им коммуникативно-общественной функции, от задач общения в соответствующей сфере. Обычно различают пять функциональных стилей речи: научный, художественный, деловой, публицистический, разговорный.

В работе[10] предложена схема стилистического анализатора, включающего набор формальных показателей. При этом производится оценка “поверхностных”, но статистически выраженных показателей, которые легко можно получить с минимальными затратами. Произведено некоторое упрощение задачи, что, однако, значительно повысило эффективность алгоритмов обработки. При этом анализируются такие компоненты языковой структуры как графика (наличие специальных знаков и пр.), морфемика (анализируются специальные приставки и суффиксы), лексика (длина слов, термины общенаучной лексики, названия официальных документов, слова организации логики повествования, и т.д.), морфология (соотношение различных частей речи, существительные среднего рода, краткие прилагательные, формы возвратных глаголов, аббревиатуры, местоимения, частицы и т.д.), синтаксис (средняя длина предложения, экспрессивная пунктуация и т.д.) Подобный стилистический анализ в отличие от стандартной схемы предобработки может потребовать сохранения различных лексических элементов, которые обычно фильтруют с помощью стоп-листов. Подобный анализ может выявить настроение, состояние и пр. характеристики автора текста, что иногда может быть полезно при оценке содержания.

1.3. Методы взвешивания слов и понижения размерности пространства признаков

Взвешивание слов представляет собой процедуру, позволяющую назначить каждому из слов словаря некоторый численный коэффициент, который можно рассматривать в качестве параметра относительной важности слова. Во многих задачах на основании значения этого коэффициента производится сокращение размерности пространства признаков. Это производится за счет выбрасывания тех слов, которые имеют значимость ниже заданного порога. Порог обычно связан с величиной добавочной сложности, возникающей при рассмотрении данного слова, которая рассматривается относительно повышения точности работы системы при его учете.

1.3.1. Взвешивание по категориям

Энтропийное взвешивание слов по категориям применяется в случае когда входные документы могут быть отнесены к одной из предустановленных категорий. Обозначим через $n_g(w)$ частоту встречи слова w в документах, которые отнесены к группе g ($g = 1, \dots, N_g$), нормализованную относительно общего числа слов в группе. Обозначим через $P_g(w)$ вероятность того, что случайно выбранное слово w появляется в группе (категории) g . Энтропия этого слова таким образом определяется по формуле:

$$H(w) = -\sum_g P_g(w) \ln P_g(w) \approx -\sum_g \frac{n_g(w)}{\sum_g n_g(w)} \ln \frac{n_g(w)}{\sum_g n_g(w)},$$

весовой коэффициент для слова $W(w)$:

$$W(w) = H_{max} - H(w), \quad H_{max} = \ln N_g.$$

1.3.2. Взвешивание по документам

Помимо инверсного взвешивания по документам (*IDF*) был предложен метод [37] энтропийного взвешивания слов по документам:

$$W(w) = 1 + \frac{1}{\ln(m)} \sum_{i=1}^m p_i(w) \ln(p_i(w)),$$
$$p_i(w) = \frac{n_i(w)}{\sum_{j=1}^m n_j(w)},$$

где $n_i(w)$ - частота слова w в документе i , а m - общее количество документов. Указанный параметр взвешивания может служить мерой того, насколько данное конкретное слово соответствует особенностям контекста, т.е. значимость слова в массиве текста. Например, слова, часто встречающиеся в разных документах, будут иметь низкую энтропию.

Указанная процедура была получена эмпирически при рассмотрении релевантных слов, которые могли бы служить в качестве индексированных терминов, описывающих документы.

1.3.3. Odds Ratio (Ранг разницы).

Задача - ранжировать документы в соответствии с их релевантностью по отношению к классу C_1 , используя меру появлений слова в качестве признака.

$$OddsRatio(F) = \log \frac{odds(W | C_1)}{odds(W | C_2)} = \log \frac{P(W | C_1)(1 - P(W | C_2))}{(1 - P(W | C_1))P(W | C_2)},$$

где $P(W|C_i)$ - условная вероятность появления слова W при заданном значении класса.

1.3.4. Методы TF-IDF

Каждому слову сопоставляется вес: $v_k^m = \frac{f_k^m \log(N_D / n_k)}{\sum_{j=1}^l f_j^m \log(N_D / n_j)}$, где N_D - количество

документов, n_k - количество документов, в которых появляется термин k , l - количество используемых терминов.

$$f_k^m = \begin{cases} 0, & s = 0 \\ \log(s) + 1, & s > 0 \end{cases}$$

где s - количество появлений термина k в документе m .

Среди методов взвешивания и кодирования можно выделить также класс методов, основная задача которых – сокращение пространства признаков. Рассмотрим их более подробно.

1.3.5. Document frequency (DF) thresholding [12].

Учитывается количество документов, в которых появляется конкретный термин. Для каждого слова словаря назначается параметр DF, задается предел, если конкретное слово имеет частоту DF меньшую этого значения, то данное слово отбрасывается из окончательного словаря. Подобная тактика основывается на том, что редкие слова являются неинформативными в смысле кластеризации. Редкие слова могут рассматриваться в качестве шума.

Преимущества данного метода – высокая скорость работы, недостаток – невысокая гибкость. Например, некоторые слова могут иметь низкую частоту и одновременно быть ключевыми.

1.3.6. Information gain (IG) [12].

IG часто применяется как метод оценки добротности терминов. Он позволяет измерить количество бит информации, полученной для кластеризации из знания того присутствует данный термин или отсутствует в документе.

Обозначим $c_i, i = 1, \dots, m$ - множество категорий. IG для термина t при этом вычисляется как:

$$G(t) = -\sum_{i=1}^m P(c_i) \log P(c_i) + P(t) \sum_{i=1}^m P(c_i | t) \log P(c_i | t) + P(\bar{t}) \sum_{i=1}^m P(c_i | \bar{t}) \log P(c_i | \bar{t})$$

Для каждого термина подсчитывается значение $G(t)$ и из окончательного словаря выбрасываются все термины, значение $G(t)$ которых меньше заданного предела.

1.3.7. Mutual Information (MI) [12].

Этот метод обычно используется при статистическом моделировании ассоциаций слов языка. Для каждого термина рассматриваются параметры: A - количество раз при которых термин t и категория c появляются совместно, B - количество раз, при которых термин t встречается с другими категориями и не встречается с c ; C - количество раз, при которых c встречается без t , N - общее количество документов. При этом MI в такой постановке:

$$I(t, c) = \log \frac{P(t \wedge c)}{P(t) \times P(c)}$$

и его оценка:

$$I(t, c) \approx \log \frac{A \times N}{(A + C) \times (A + B)}$$

$I(t, c) = 0$, если t и c независимы.

Недостатком MI является то, что на меру взвешивания сильно влияет побочная вероятность появления терминов, что можно показать с помощью следующей эквивалентной формулы:

$$I(t, c) = \log P(t | c) - \log P(t),$$

таким образом для терминов с равной условной вероятностью $P(t|c)$, редко встречающиеся термины будут доминировать.

1.3.8. Статистика χ^2 [12].

Метод измеряет оценку независимости величин t и c . A - количество совместных появлений t и c ; B - количество раз, при которых термин t встречается с другими категориями и не встречается с c ; C - количество раз, при которых c встречается без t , D - количество раз, когда не появляется ни t ни c ; N - общее количество документов. Оценка добротности термина производится по следующей формуле:

$$\chi^2(t, c) = \frac{N \times (AD - CB)^2}{(A + C) \times (B + D) \times (A + B) \times (C + D)}$$

Как и для MI $\chi^2(t, c) = 0$, если t и c независимы. Для оценки добротности термина при выборе пространства признаков можно опираться на следующую формулу:

$$\chi_{avg}^2(t) = \sum_{i=1}^m P(c_i) \chi^2(t, c_i)$$

Основным отличием MI и χ^2 является то, что χ^2 - нормализованное значение, и следовательно позволяет сравнивать полученные значения внутри каждой отдельной категории.

1.3.9. Коэффициент корреляции C .

Коэффициент корреляции определяется по следующей формуле:

$$C(t, c) = \frac{\sqrt{N} \times (AD - CB)}{\sqrt{(A + C) \times (B + D) \times (A + B) \times (C + D)}}$$

(см. обозначения для статистики χ^2). В [19] указывается, что коэффициент корреляции является односторонним вариантом χ^2 . Коэффициент корреляции используется для составления локальных словарей категорий. Она позволяет выделить только те слова, которые информативны в смысле их наличия в тексте данной категории, тогда как χ^2 выбирает также и те слова, которые являются информативными в случае их отсутствия в тексте для данной категории.

1.3.10. Term strength (TS) [12].

Метод оценивает важность термина, основываясь на том как термин появляется в близко-классифицируемых документах. Метод использует обучающую выборку для получения пар документов, чья мера близости выше заданного предела. Параметр TS вычисляется как условная вероятность того, что термин появится во второй половине пары рассматриваемых документов при условии того, что он появился в первой. Обозначим, x и y - произвольная пара разных, но близких документов, t - термин, тогда:

$$s(t) = P(t \in y | t \in x).$$

1.3.11. Метод случайного отображения

Если все приведенные выше методы могут использовать как для решения задачи взвешивания отдельных слов, так и для понижения размерности пространства признаков, то данный метод применяется только в задаче выбора информативных признаков.

В работе [5] представлено обоснование возможности использования простого метода случайного отображения (*random mapping*).

В указанной работе единичный вектор e_i , использующийся для векторного представления конкретного слова словаря заменяется на случайный вектор r_i , меньшей размерности. Каждый компонент такого вектора – случайная величина, а сам вектор нормализуется, чтобы иметь единичную длину.

В предлагаемом варианте, j -ый документ коллекции представляют в виде вектора:

$$x_j = \sum_k n_{jk} r_k.$$

Однажды выбранные случайные значения для векторов r_k далее фиксируются и используются как постоянные значения. Чаще всего используется нормальное распределение.

В работе [6] представлено обоснование возможности применения метода случайного отображения.

В работе [7] был показан интересный результат, при котором использовалась бинарный вариант рассматриваемого метода, каждый случайный вектор при этом формировался случайным распределением пяти единиц, остальные элементы оставались равными 0. В этом варианте, точность работы алгоритма оставалась в пределах одного процента по сравнению с алгоритмом простой векторной модели.

Экспериментальные исследования мер оценки значимости отдельных слов

В работе [12] был проведено экспериментальное исследование наиболее известных мер взвешивания слов. Эксперимент проводился на нескольких массивах текстов: Reuters-22173, OHSUMED, для оценки эффективности мер применялись методы категоризации kNN [13] и LLSF [14].

В результате были сделаны следующие выводы:

- наиболее эффективными были признаны методы IG, CHI, DF; TS показал среднюю производительность, MI – низкую;
- показана сильная корреляция между весами терминов в IG, CHI, DF;
- на практике проиллюстрирована важность присутствия общеупотребляемых слов в пространстве признаков;

Литература

1. Salton, G., McGill, M.J., Introduction to Modern Information Retrieval, McGraw-Hill, New York, 1983.
2. M. Berry, S. Dumais, G.W.O'Brien, Using linear algebra for intelligent information retrieval, *SIAM Rev.* 37 (1995), 573-595
3. S. Deerwester, S. Dumais, G. Furnas, K. Landauer, Indexing by latent semantic analysis, *J. Amer Soc Inform. Sci.* 41 (1990) 391-407
4. S. Gallant, W. Caid, J. Carleton, R. Hecht-Nielsen, K. Pu Qing, D. Sudbeck, HNC's MatchPlus system, *ACM SIGIR Forum* 26 (2) (1992) 34-38
5. S. Kaski, T. Honkela, K. Lagus, T. Kohonen, WEBSOM – self-organizing maps of document collection, *Neurocomputing*, vol.21, no. 1-3, p. 107-117 (1998)
6. S. Kaski, Dimensionality reduction by random mapping: fast similarity computation for clustering, *Proc IJCNN'98*, vol. 1, pp. 413-418
7. T. Kohonen, "Self-organization of very large document collections: State of the art", in *Proc. ICANN98*, the 8th Int. Conf. on Artificial Neural Networks, L. Niklasson, M. Boden, and T. Ziemke, Eds., vol. 1, pp. 65-74, Springer, London, 1998.
8. Корж В.В. Диссертация к.т.н. МИФИ. Методы кодирования текстовой информации для построения нейросетевых классификаторов документов, М., 2000 г.
9. Natalija Vlajic, Howard C. Card, An Adaptive Neural Network Approach to Hypertext Clustering.
10. Браславский П.И. Использование стилистических параметров документа при поиске информации в Internet
11. K.E. Lochbaum, and L.A. Streeter, Combining and comparing the effectiveness of latent semantic indexing and the ordinary vector space model for information retrieval, *Information Processing and Management*, 25(6), pp. 665-676, 1989.
12. Yang, Y., Pedersen, J.O., A Comparative Study on Feature Selection in Text Categorization, *Proc. of the 14th International Conference on Machine Learning ICML97*, pp. 412-420, 1997.
13. Y. Yang Expert Network: Effective and efficient learning from human decisions in text categorization and retrieval. In *17th Ann Int ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'94)*, 1994, pages 13-22,
14. Y. Yang and C.G. Chute An example-based method for text categorization and retrieval. *ACM transaction on Information Systems (TOIS)*, pages 253-277, 1994
15. Dumais, S. T. and Chen, H. Hierarchical classification of Web content. In *Proceedings of SIGIR-00, 23rd ACM International Conference on Research and Development in Information Retrieval (Athens, Greece, 2000)*, 2000, 256–263.
16. Apte, C., Damerau, F. J., and Weiss, S. M. Automated learning of decision rules for text categorization. *ACM Trans. on Inform. Syst.* 12, 3, 1994, 233–251.
17. Lewis, D. D. An evaluation of phrasal and clustered representations on a text categorization task. In *Proceedings of SIGIR-92, 15th ACM International Conference on Research and Development in Information Retrieval (Copenhagen, Denmark, 1992)*, 1992, 37–50.
18. Tzeras, K. and Hartmann, S. Automatic indexing based on Bayesian inference networks. In *Proceedings of SIGIR-93, 16th ACM International Conference on Research and Development in Information Retrieval (Pittsburgh, PA, 1993)*, 1993, 22–34.
19. NG, H. T., GOH, W. B., AND LOW, K. L. Feature selection, perceptron learning, and a usability case study for text categorization. In *Proceedings of SIGIR-97, 20th ACM International Conference on Research and Development in experimental inquiry. J. Assoc. Comput. Mach.* 8, 3, *Information Retrieval (Philadelphia, PA, 1997)*, 67–73.
20. DAGAN, I., KAROV, Y., AND ROTH, D. Mistake-driven learning in text categorization. In *Proceedings of EMNLP-97, 2nd Conference on Empirical Methods in Natural Language Processing (Providence, RI, 1997)*, 55–63.

КОМПЬЮТЕРНОЕ АДАПТИВНОЕ ТЕСТИРОВАНИЕ: АЛГОРИТМ И ТЕХНОЛОГИЯ СОЗДАНИЯ ТЕСТОВ.

Мигунов В.О.

Научный руководитель: д.т.н. проф. Шахнов В.А.

МГТУ им. Н.Э. Баумана, Москва, Россия

COMPUTER ADAPTIVE TESTING. ALGORITHM AND TEST MAKING TECHNOLOGY.

Migunov V.O.

DPh. professor Shakhnov V.A.

BMSTU, Moscow, Russia

mr_mb@mail.ru

Аннотация

В данной статье описывается алгоритм проведения адаптивного тестирования, преимущества и достоинства этого метода перед традиционным тестированием. Также описывается технология создания адаптивных тестов.

Abstract

This article describe computer adaptive testing (CAT), advantages and merits of this testing method over the traditional ones. Also there is described computer adaptive test creation technology.

Введение

В последнее время в образовательной среде уделяется большое внимание тестовым технологиям. Это связано не только с предстоящим введением единого государственного экзамена, но и со стандартизацией процедур оценки качества знаний. Проблема точной оценки качества знаний школьников, абитуриентов, студентов, аспирантов и других групп обучающихся становится острой при ориентации на современные технологии тестирования, так как именно тестирование позволяет с достаточной точностью определить уровень знаний испытуемого, более беспристрастно и объективно. Создание хороших тестов, измеряющих и отражающих реальный уровень знаний тестируемых, является целой наукой [3, 4, 6]. Здесь необходимы не только знания в предметной области, но и четкое знание тестологии (представление параметрах заданий, процессе проведении тестирования в целом и т.д.).

Для создания адаптивного теста необходимо иметь хороший банк тестовых заданий. Алгоритм получения качественного банка тестовых заданий представлен на рис 1.



Рис 1. Алгоритм создания качественного банка тестовых заданий для адаптивного теста.

Кратко опишем данный алгоритм. Сначала необходимо подготовить обычный тест с заданиями различной трудности по заданной теме, причем желательно, чтобы трудность заданий распределялась по нормальному закону распределения. Т.е. заданий с очень малой и

очень большой трудности должно быть мало, а средней трудности – много. Проверить тест на валидность (т.е. на соответствие заданной теме) и на точность формулировок (например, отсутствие двусмысленных толкований). Далее следует произвести обычное (бланковое или компьютерное без адаптивного алгоритма) тестирование на достаточной совокупности испытуемых, т.н. апробацию теста. Получить ответы (матрицу ответов A размерностью $n \times m$, n – количество испытуемых, m – число заданий). Используя полученную матрицу ответов, с помощью программы *BILOG* или *LOGIST* получаем латентные (скрытые) характеристики заданий: трудность - β , дифференцирующий параметр - α и фактор угадывания - c [1, 3, 5]. Причем в зависимости от выбранной модели функции успеха этих параметров может быть от одного - β (тогда параметры полагают: $\alpha = 1.7, c = 0$), до трех - α, β, c . Исходя из параметров заданий (модели тестирования) функция успеха будет выглядеть так:

$$p(\theta_i, \alpha_j, \beta_j, c_j) = c_j + (1 - c_j) \left(\frac{1}{1 + e^{-\alpha_j(\theta_i - \beta_j)}} \right),$$

где

θ_i - уровень подготовки испытуемого i , где $i \in 1 \dots n$. Измеряется в логитах $\in (-3; 3)$

α_j - дифференцирующий фактор j -го задания, где $j \in 1 \dots m$.

β_j - трудность задания j -го задания, где $j \in 1 \dots m$. Измеряется в логитах $\in (-3; 3)$

c_j - фактор угадывания j -го задания, где $j \in 1 \dots m$.

$p(\theta_i, \alpha_j, \beta_j, c_j)$ - вероятность правильного выполнения i -м испытуемым j -го задания.

Функция успеха описывает вероятность правильного выполнения i -м испытуемым j -го задания. Параметры j -го задания найдены из матрицы ответов A .

Далее происходит отсев «плохих» заданий и их модификация. После этого процесс апробации теста можно повторить. При отсутствии нареканий к банку заданий можно переходить к использованию адаптивного теста. Подробно о создании банков заданий описано в [4, 6].

Алгоритм адаптивного тестирования

Алгоритм проведения адаптивного тестирования представлен на рис. 2. Адаптивный алгоритм строится на следующих предпосылках, что имеется банк откалиброванных заданий под соответствующую модель тестирования и заданы критерии останова. Описание работы адаптивного алгоритма представлено в работе [7]. Рассмотрим данный алгоритм более подробно.

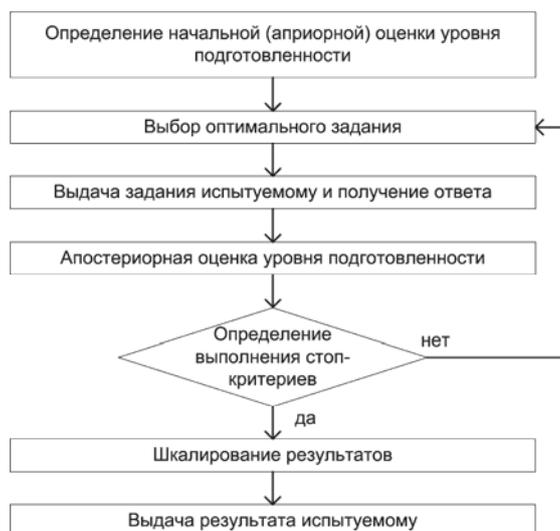


Рис. 2. Алгоритм адаптивного тестирования.

Для начала работы адаптивного алгоритма необходимо правильно выбрать оцениваемый уровень подготовки, поскольку от этого зависит быстрота сходимости

алгоритма. Обычно его выбирают, задав несколько вопросов 5-10 различной трудности и усреднив значения ответов, или полагают равным среднему значению 0 (50%) с добавлением некоторой случайной составляющей (чтобы сделать уникальным порядок следования вопросов). Далее вступает в силу адаптивный алгоритм. Исходя из текущего уровня знаний

θ_s и используя информационную функцию задания $I_j(\theta_s) = \left[\frac{p_j^2(\theta_s, \Delta_j)}{p_j(\theta_s, \Delta_j)(1-p_j(\theta_s, \Delta_j))} \right]$, где

$\Delta_j = (\alpha_j, \beta_j, c_j)$, находим j , для которого информационная функция составляет максимум.

Это и будет следующий задаваемый вопрос. Затем, получив ответ u_s на заданный вопрос, вычисляем апостериорную оценку уровня подготовленности.

$$\theta_{s+1} = \theta_s + \frac{\sum_{k=1}^s S_k(\theta_s)}{\sum_{k=1}^s I_k(\theta_s)}, \quad \text{где } S_k(\theta_s) = (u_k - p_k) \frac{p_k'}{p_k(1-p_k)}, \quad \text{где } p_k = p_k(\theta_s, \alpha_k, \beta_k, c_k)$$

$$p_k' = \frac{dp_k}{d\theta}$$

$$SE(\theta_s) = \frac{1}{\sqrt{\sum I_i(\theta_s)}} \leq \varepsilon, \quad \text{где } SE(\theta_s) - \text{ошибка вычисления } \theta_s$$

Далее определяем критерии остановки. Он может быть один или несколько одновременно. Таким критерием может быть: время, отведенное на тест; минимальное и/или максимальное число заданных вопросов; задано определенное число вопросов; ошибка вычисления уровня подготовленности меньше заданной, остановка теста пользователем. Если критерии выполняются, то тест останавливается, вычисляется итоговый балл и его точность и выводится результат. Иначе, вычисленная апостериорная оценка уровня подготовленности используется в выборе следующего вопроса (по максимуму информационной функции задания) и тест продолжается.

Выводы

Адаптивное тестирование имеет неоспоримое преимущество над обычным компьютерным тестированием, но требует больших затрат при подготовке [2]. Оно целесообразнее, когда необходимо уменьшить время тестирования или протестировать большое количество испытуемых, например абитуриентов.

Литература

1. Карданова Е.Ю., Нейман Ю.М. Основные модели современной теории тестирования // Вопросы тестирования в образовании – М. 2003, №7 – с.12-37.
2. Мигунов В.О. Проблемы развития адаптивного тестирования // Вопросы тестирования в образовании – М. 2003, №8 – с. 102-104.
3. Нейман Ю.М., Хлебников В.А. Введение в теорию моделирования и параметризации педагогических тестов – М.: Прометей, 2000 – 168 с.
4. Рудаков О.В. Типологизации педагогических тестов и заданий в тестовой форме // Вопросы тестирования в образовании – М. 2003, №7 – с. 80-90.
5. Самыловский А.И. Тест как объективный измерительный инструмент в образовании // Вопросы тестирования в образовании – М. 2001, №1 – с. 10-39.
6. Сысоева Л.А. Методика построения логико-семантической модели структуры содержания дисциплины // Вопросы тестирования в образовании – М. 2003, №8 – с. 15-20.
7. Чельшкова М.Б. Адаптивное тестирование в образовании (теория, методология, технология) – М.: Исследовательский центр проблем качества подготовки специалистов, 2001 – 165 с.

ТЕХНОЛОГИИ ЭЛЕКТРОННОЙ ЦИФРОВОЙ ПОДПИСИ

Аксенов А.В.

*Научный руководитель: к.т.н, доцент Волосатова Т.М.
МГТУ им. Баумана, кафедра РК6*

ELECTRONIC DIGITAL SIGNATURE TECHNOLOGIES

Aksenov A.V.

*Scientific chief: Ph.D., senior lecturer Volosatova T.M.
BMSTU, PK6
alexey_aksenov@mail.ru*

Аннотация

В данной работе рассматриваются особенности применения технологии электронной цифровой подписи, дается юридическое обоснование применения электронной цифровой подписи на территории Российской Федерации, дается обзор отечественных программных продуктов, сертифицированных для применения на территории РФ.

Abstract

In the given work features of application of technology of the electronic digital signature are considered, the legal substantiation of application of the electronic digital signature in territory of the Russian Federation is given, the review of the domestic software products certificated for application in territory of the Russian Federation is given.

Введение

Электронно-цифровая подпись (ЭЦП) используется физическими и юридическими лицами в качестве аналога собственноручной подписи для придания электронному документу юридической силы, равной юридической силе документа на бумажном носителе, подписанного собственноручной подписью правомочного лица и скрепленного печатью.

Электронный документ - это любой документ, созданный и хранящийся на компьютере, будь то письмо, контракт или финансовый документ, схема, чертеж, рисунок или фотография.

Провайдер ЭЦП - это программно-криптографическое средство, которое обеспечивает:

1. проверку целостности документов;
2. конфиденциальность документов;
3. установление лица, отправившего документ.

Юридическое обоснование применения электронных документов в рамках действующего законодательства РФ.

1. "Гражданский Кодекс Российской Федерации",

Часть первая, Статья 160, п. 2, в котором говорится:

"Использование при совершении сделок ...электронно-цифровой подписи ...допускается в случаях и порядке, предусмотренных законом, иными правовыми актами или соглашением сторон";

Часть первая, Статья 434, п.1

"Если стороны договорились заключить договор в определенной форме, он считается заключенным после придания ему условленной формы, хотя бы законом для договоров данного вида такая форма не требовалась".

Часть первая, Статья 434, п.2

"Договор в письменной форме может быть заключен ...путем обмена документами посредством почтовой, телеграфной, телетайпной, телефонной, электронной или иной связи, позволяющей достоверно установить, что документ исходит от стороны по договору";

2. Федеральный Закон "Об информации, информатизации и защите информации",

Глава 2, Статья 5, п.3

"Юридическая сила документа, хранимого, обрабатываемого и передаваемого с помощью автоматизированных информационных и телекоммуникационных систем, может подтверждаться электронной цифровой подписью.

Юридическая сила электронной цифровой подписи признается при наличии в автоматизированной информационной системе программно-технических средств, обеспечивающих идентификацию подписи, и соблюдении установленного режима их использования";

Процедура использования ЭЦП.

1. Каждому пользователю, участвующему в обмене электронными документами, генерируются уникальные секретный и открытый криптографические ключи.

Секретный ключ является элементом, с помощью которого производится шифрование документов и формируется электронно-цифровая подпись. Секретный ключ является собственностью пользователя, выдается ему на отдельном носителе (дискета, смарт-карта, touch memo) и держится в секрете от других пользователей.

Открытый ключ используется для проверки ЭЦП получаемых документов-файлов. Владелец должен обеспечить наличие своего открытого ключа у всех, с кем он собирается обмениваться подписанными документами. Кроме того, дубликат открытого ключа направляется в Удостоверяющий Центр, где создана библиотека открытых ключей ЭЦП. В библиотеке Центра обеспечивается регистрация и надежное хранение открытых ключей во избежание попыток подделки или внесения искажений.

2. Пользователь устанавливает под электронным документом свою электронную цифровую подпись. При этом на основе секретного ключа ЭЦП и содержимого документа путем криптографического преобразования вырабатывается некоторое большое число, которое и является электронно-цифровой подписью данного пользователя под данным конкретным документом. Это число добавляется в конец электронного документа или сохраняется в отдельном файле.

В подпись записывается следующая информация:

1. дата формирования подписи;
2. информация о лице, сформировавшем подпись;
3. имя файла открытого ключа подписи.

3. Пользователь, получивший подписанный документ и имеющий открытый ключ ЭЦП отправителя на основании текста документа и открытого ключа отправителя выполняет обратное криптографическое преобразование, обеспечивающее проверку электронной цифровой подписи отправителя. Если ЭЦП под документом верна, то это значит, что документ действительно подписан отправителем и в текст документа не внесено никаких изменений.

Удостоверяющий центр

Для обеспечения безопасности и исключения подмены открытых ключей в соответствии с Федеральным законом "Об электронной цифровой подписи" создается

доверительная организация - Удостоверяющий центр, куда направляются открытые ключи всех пользователей.

Удостоверяющий центр производит сертификацию открытых ключей ЭЦП путем подписания открытого ключа пользователя своим секретным ключом.

Для подтверждения соответствия средства ЭЦП установленным требованиям владельцу ключа подписи выдается сертификат ключа подписи, который содержит следующие сведения:

1. имя владельца, другие идентифицирующие данные;
2. сроки действия ключа;
3. уникальный номер сертификата ключа подписи;
4. наименование средства ЭЦП, с которым используется данный открытый ключ;
5. наименование и местонахождение Центра, выдавшего сертификат ключа подписи;
6. правоотношения, в которых ЭЦП имеет юридическое значение;
7. дата выдачи сертификата;
8. сведения о действии сертификата;

Сертификат ключа подписи в электронном виде, подписанный секретным ключом Удостоверяющего центра, направляется пользователю ключа подписи и вносится в реестр сертификатов ключей подписей Удостоверяющего центра.

Изготовление сертификатов ключей подписей на бумажных носителях осуществляется на основании заявления участника информационной системы. Заявление подписывается собственноручно владельцем сертификата подписи. Содержащиеся в заявлении сведения подтверждаются предъявлением соответствующих документов.

Сертификаты на бумажных носителях оформляются в двух экземплярах, которые заверяются собственноручными подписями обладателя ЭЦП и уполномоченного лица Центра и печатью Центра. Один экземпляр выдается обладателю электронно-цифровой подписи, второй остается в Удостоверяющем центре.

Российские средства электронной цифровой подписи.

1. Крипто-Про

Основные направления деятельности компании:

- Разработка и внедрение криптографических средств, реализующих российские криптографические алгоритмы в соответствии со стандартом Microsoft Cryptographic Service Provider. Интеграция стандартизированных национальных криптографических алгоритмов с продуктами Microsoft.
- Разработка и использование криптографических средств, поддерживающих Microsoft Public Key Infrastructure(PKI). Внедрение PKI.

Крупные корпоративные клиенты: Аппарат Правительства Российской Федерации, Государственный таможенный комитет РФ, Банк Внешней Торговли, ОАО "Альфа-Банк", ОАО "Лукойл", ЗАО "Управляющая компания НИКойл", ЗАО "Группа МДМ", ООО Платежная Интернет-система "Рапида", ОАО "Мосжилрегистрация".

Криптопровайдер КриптоПро CSP разработан ООО "Крипто-Про" и [ГУП НТИЦ "Атлас"](#) по техническому заданию согласованному с ФАПСИ в соответствии с криптографическим интерфейсом фирмы Microsoft - Cryptographic Service Provider (CSP).

Криптопровайдер КриптоПро CSP имеет сертификат ФАПСИ и может использоваться для формирования ключей шифрования и ключей электронной цифровой подписи, шифрования и имитозащиты данных, обеспечения целостности и подлинности информации, не содержащей сведений, составляющих государственную тайну.

2. Анкад

Крупные корпоративные клиенты: Федеральное агентство правительственной связи и информации (ФАПСИ), МВД, Минобороны, Центральный Банк РФ, Сбербанк России.

Основой построения систем криптографической защиты информации служит шифратор серии КРИПТОН - аппаратное устройство КРИПТОН или программный драйвер-шифратор Cryption Emulator.

С шифратором работает прикладное программное обеспечение, рассчитанное на конечного пользователя, и/или средства разработки - библиотеки, предназначенные для встраивания в продукты независимых разработчиков функций шифрования и/или электронной цифровой подписи (ЭЦП).

Ключи шифрования и ЭЦП могут храниться на дискетах, смарт-картах с открытой и защищенной памятью, микропроцессорных смарт-картах, идентификаторах Touch Memoгу, брелоках eToken или ruToken для шины USB.

3. Инфотекс.

Крупные корпоративные клиенты: Сбербанк РФ, ООО "Международная Транспортная Биржа".

Компания представляет линию программных продуктов VipNet на базе Средства криптографической защиты информации (СКЗИ) "Домен-К" предназначены для формирования ключей шифрования и ключей электронной цифровой подписи, шифрования и имитозащиты данных, обеспечения целостности и подлинности информации.

Средствами СКЗИ "Домен-К" реализуются функции:

1. Вычисление и проверка электронно-цифровой подписи в соответствии с ГОСТ 34.10-94 "Информационная технология. Криптографическая защита информации. Процедуры выработки и проверки электронной цифровой подписи на базе асимметричного криптографического алгоритма" и ГОСТ 34.10-2001 "Информационная технология. Криптографическая защита информации. Процессы формирования и проверки электронной цифровой подписи". Вычисление ЭЦП производится с использованием секретных ключей.
2. Хэширование информации в соответствии с ГОСТ 34.11-94 "Информационная технология. Криптографическая защита информации. Функция хэширования". Представляет собой однонаправленное преобразование свертки, независимое от ключа. Используется в алгоритмах ЭЦП, а также для подтверждения достоверности данных (в частности, паролей).
3. Размер секретных ключей шифрования, секретных ключей ЭЦП и секретных ключей Диффи-Хелмана (секретных ключей для открытого распределения ключей) составляет 256 битов (32 байта). Размер открытых ключей ЭЦП и открытых ключей для протокола Диффи-Хелмана составляет 1024 бита (128 байтов) при использовании ГОСТ 34.10-94 и 512 бит (64 байта) для ГОСТ 34.10-2001.

Литература

1. www.ancud.ru
2. ospkibook.sourceforge.net
3. www.pki-page.org
4. www.cryptopro.ru
5. www.infotecs.ru
6. www.digitalsign.ru

ТРЕНАЖЕР ДЛЯ ПОДГОТОВКИ ОПЕРАТОРОВ ПОДВОДНЫХ АППАРАТОВ

Н. Буйда, М. Григорьев, А., С, Медведев, Кошель Р
Московский Государственный Технический Университет им. Н.Э. Баумана,
кафедра СМ11

SIMULATOR FOR PREPARATION OF OPERATORS OF UNDERWATER DEVICES

N. Buyda, M. Grigoriev, A.S., Medvedev, R.Koschel
Bauman Moscow State Technical University,
Department SM11

Аннотация

В работе рассматриваются вопросы технического обеспечения подводных тренажеров на основе опыта разработок подводных телеуправляемых комплексов (ТПК), выполняемых в течение ряда лет на кафедре СМ-11 МГТУ им Баумана.

Abstract

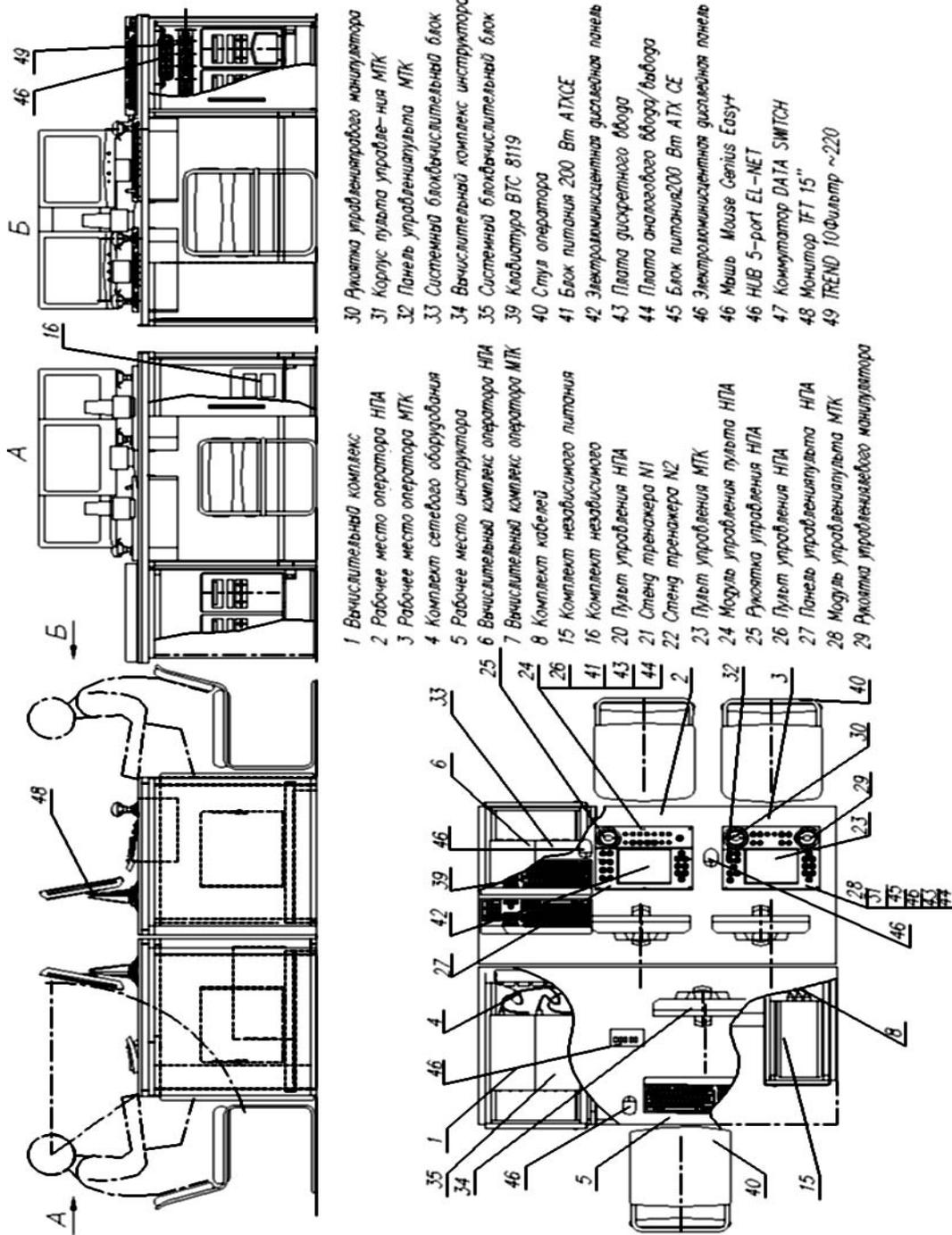
In work questions of technical maintenance of underwater simulators are considered on the basis of experience of development of the underwater long-distance complexes which is carried out for a number of years on faculty of Sm - 11 of MGTU n.a. Bauman.

На кафедре СМ11 был создан ряд программно-технических систем имитационного моделирования (тренажеров) для подготовки операторов телеуправляемых для осмотровых и рабочих подводных аппаратов. Тренажер предназначен как для использования в учебных центрах подготовки экипажей, так и для условий судового применения для тренировок без запуска подводного аппарата.

Типовой тренажерный комплекс (ТК) содержит пультовые средства управления и вычислительный моделирующий комплекс (ВК). Производительность ВК определяется задачами моделирования, которые в свою очередь определяются конструкцией подводного комплекса, функционированием его систем, определяемыми техническими условиям эксплуатации. Все процессы моделирования должны осуществляться в режиме реального времени.

Организация вычислительного процесса и структура программно-аппаратных средств тренажерного комплекса определяется следующими требованиями:

- структура ВК технических средств должна обладать гибкостью в использовании технического (ТО) и программного обеспечения (ПО);
- при разработке программных средств ТК необходимо обеспечить модульность в организации вычислительного процесса, с целью обеспечения гибкости в разделении задач моделирования;
- вычислительный комплекс ТК должен обеспечивать возможность наращивания новых программных модулей и дополнительных вычислительных
- ВК тренажера должен обеспечивать связь с другими моделирующими комплексами, например, судовым тренажером базового судна;
- количество используемых ЭВМ определяется качеством моделирования задач и субъективным восприятием выполняемых действий при работе оператора за пультами управления.



Состав ТК

Общий состав ТК (технические средства, рабочие места, ВК, пультовые устройства управления) ясен из рис 1.

Количественный состав технических средств ТК содержит следующие компоненты:

- ПЭВМ (4 шт.) + пультовые ЭВМ (2 шт.);
- монитор (3 шт.);
- клавиатура, мышь (3 шт.);
- пульт управления подводным аппаратом;

Рис 1 Общий вид тренажерного комплекса

- пульт управления манипуляционным комплексом;
- акустические колонки;
- сетевой коммутатор (HUB);
- независимый источник питания (2 шт.).

Стенд тренажера

Стенд тренажера (рис 2) представляет собой конструкцию на основе стальных коробов, облицованных мебельными плитами. Передвижные приборные шкафы стенда, используемые для размещения вычислительной техники, средств коммуникации и источников питания, обеспечивают свободное обслуживание и достаточную защищенность от неосторожного обращения с ВК. Шкафы обеспечены средствами принудительной вентиляции и электрической защиты от перегрузок.

Мониторы и средства управления размещаются на столе стенда, Для выполнения функций обучения и визуальными средствами наблюдения ВК дополняются платами графического акселератора и звуковыми средствами пояснения.

Пульты управления

Конструкция пультов полностью совпадает по своему назначению и использованию со штатными средствами управления реального подводного аппарата. Некоторые изменения связаны только в части организации обмена через комплекс сопряжения и питания пультов. В составе программно-аппаратного комплекса имитационной аппаратуры включается комплекс сопряжения с основным ВК аппаратуры пультов управления и встроенных ЭВМ. Все остальные системы ТПК, включающие системы подводного аппарата, системы спуско-подъемных устройств, навигационное и гидроакустическое оборудование, системы энергетики имитируется с помощью вычислительного комплекса (ВК), моделирующих технических средств и систем управления тренажера.

При работе на ТК обучаемые операторы и инструкторы занимают свои рабочие места. Предусмотрено два рабочих места обучаемых и рабочее место инструктора. Рабочие места располагаются друг относительно друга на расстоянии в пределах 20 м.



Рис 2 Внешний тренажера для подготовки операторов подводных аппаратов

Габариты тренажерного комплекса

Габариты рабочего места для двух операторов и инструктора

- 1600 x 1300 x 1140 мм.

Состав моделирующего ПО

В программное обеспечение включены следующие модули:

Модуль ИНСТРУКТОР
Модуль ДИНАМИКА АППАРАТА
Модуль ДИНАМИКА МАНИПУЛЯТОРОВ
Модуль ЭНЕГЕТИКА
Модуль ГИДРОАКУСТИКА
Модуль ВИЗУАЛИЗАЦИЯ
Модуль ПУЛЬТ НПА
Модуль ПУЛЬТ МТК

Операционная система

В качестве основной операционной системы используется Windows 2000.

Вычислительный комплекс

Общее число моделирующих ЭВМ составляет - 6 Количество узлов в подсети тренажера составляет - 6

В ВК применены ЭВМ достаточной производительности в следующей конфигурации:

- процессор AMD Athlon 1333 ;
- графический акселератор 32 мБ GIGABYTE GA-GF1280 ;
- накопитель HDD IBM DESKSTAR 20GB 7200 об/мин ;
- оперативная память MICRON DDR 128MB;

Энергопитание ТК

Электропитание ТК осуществляется от однофазной сети переменного тока напряжением ~220 В, частотой - 50 Гц, через два вторичных источника питания(UPS), имеющийся в составе тренажера. Потребляемая мощность не более 850 Вт.

ВОПРОСЫ ВИЗУАЛЬНОГО МОДЕЛИРОВАНИЯ СИСТЕМ В ТРЕНАЖЕРАХ ПОДВОДНЫХ ТЕЛЕУПРАВЛЯЕМЫХ КОМПЛЕКСОВ

А.Гнеломедов, С Медведев

Московский Государственный Технический Университет им. Н.Э. Баумана,
кафедра СМ11

QUESTIONS OF VISUAL MODELLING OF SYSTEMS IN SIMULATORS OF UNDERWATER LONG-DISTANCE COMPLEXES

A.Gnelomedov, S.Medvedev

Bauman Moscow State Technical University,
Department SM11

Аннотация

В докладе описывается метод моделирования систем управления процессами связанными с эксплуатацией подводных телеуправляемых комплексов (ПТК), примененный при разработке тренажерного комплекса в кафедры СМ-11 МГТУ им. Н.Э.Баумана.

Abstract

In the report the method of modelling of control systems by processes connected with operation of the underwater long-distance complexes (ПТК), applied is described at development of a training complex in faculties of Sm - 11 of MGTU n.a. N.E.Baumana.

В основе метода лежит разделение уровней разработки: уровень математического моделирования и уровень организации информационного обмена с целью максимально возможного исключения рутинного, непроизводительного труда. Предлагаемое решение состоит в разработке программного обеспечения (ПО) управления информационным обменом внутри и между системами, позволяющее:

- быстро и гибко сформировать широкий спектр систем управления, имеющих место в реальном подводном телеуправляемом комплексе
- использовать объектные структуры в задачах моделирования систем управления движения, систем управления манипуляционным комплексом, датчикового оборудования, систем энергетики и гидравлики и т.п.

Упростив написание сложных многосвязных программ, которые трудно поддаются диагностированию. Далее рассматривается метод формирования структуры систем управления применительно к системе управления движением (СУД) подводного телеуправляемого комплекса.

Объект моделирования

Для решения задачи моделирования систем необходимо определить понимание того, что это такое, основные решаемые задачи, компоненты системы, приоритетные области для моделирования.

СУД предназначена для ручного, автоматизированного и автоматического дистанционного управления положением центра масс и угловой ориентацией подводного аппарата (ПА).

СУД представляет собой программно-аппаратный комплекс, конструктивно разделенный на отдельные блоки связанные между собой информационными связями. В процессе функционирования СУД находится в состоянии информационного обмена.

Обобщенно, задачу СУД можно определить как наиболее оптимальное и быстрое преобразование целевых указаний оператора заданных в общем случае в виде конечных значений или изменений обобщенных параметров движения в сигналы управления для исполнительных органов ПА с учётом параметров состояния ПА получаемых с датчиков обратной связи.

Для целей моделирования можно выделить три уровня функционирования СУД:

- макроуровень, когда по указаниям оператора, решаются общие задачи, учитывающие интересы пользователя: куда плыть, что снимать, что обрабатывать;
- алгоритмический уровень – внутренние управляющие и преобразующие алгоритмы управления;
- аппаратный уровень обмена и соответствия реальных сигналов, поступающих с СУД на конкретные исполнительные органы ПА.

Наиболее приоритетными для решаемой задачи являются два последних уровня, однако задачи моделирования аппаратного уровня обмена должны решаться совместно с аналогичными задачами на этапе моделирования движительно-рулевого комплекса и структуры систем управления ПА, поэтому детальное рассмотрение этих вопросов не включено в рамки данной статьи.

Моделирование

Выбрав основной задачей моделирования – моделирование внутренних алгоритмов систем, определим возможные виды целевых систем для разработчика СУД:

- математические системы – наиболее обобщенное моделирование логики блоков систем, подразумевающее идеальные связи. Погрешности таких систем – это погрешности применяемых математических методов: приближения и упрощения задач, а также погрешности дискретизации, определяемые аппаратными средствами. Обрабатываемые сигналы системы – числовые значения, не ассоциированные с какими-либо физическими аналогами;
- электрические системы – в рамках решаемой задачи являются частным случаем моделирования математических систем. Внимание уделяется моделированию реально существующих электрических устройств, моделированию токоведущих соединений. Числовые значения сигналов этой системы – являются аналогами электрических сигналов, параметров устройств;
- гидравлические системы – также как и предыдущий тип систем являются частным случаем моделирования математических систем. Внимание уделяется моделированию реально существующих гидравлических устройств, моделированию трубопроводов. Числовые значения сигналов этой системы – являются аналогами гидравлических сигналов, параметров устройств.

В число погрешностей систем двух последних типов следует добавить моделируемые погрешности обусловленные неидеальностью реальных устройств. Реализация таких систем является открытой, пользователь может модернизировать их, наблюдать и анализировать работу отдельных участков.

Математические системы

ПО тренажерного комплекса содержит:

- встраиваемый интерпретатор команд – уровень кода программы;
- внешние модули для редактирования – файла описания СУД.

Встраиваемый интерпретатор – библиотека классов в общем случае реализующая методы разбора и анализа команд, а также базовые математические алгоритмы их обработки, способы обмена информацией, интерфейсную часть. Подключается к проекту на этапе разработки и отладки. Открытость библиотеки на уровне исходных кодов предоставляет разработчику возможность модернизации, наращивания классов и методов интерпретатора под конкретные задачи.

Разработка конечной СУД производится в специальном программном модуле, предназначенном для визуальной разработки структурной схемы и её отладки. Конечный результат - файл содержащий структурированное текстовое описание СУД на встроенном языке моделирования.

Модернизация и разработка систем может производиться способом непосредственного редактирования кода файла, описывающего СУД, при этом используется либо редактор языка моделирования, либо любой доступный текстовый редактор, также для создания и

корректировки систем предусмотрен специальный модуль, позволяющий осуществлять визуальное проектирование и отладку. Состав такого модуля имеет следующие элементы:

- панель выбора компонент (блоков);
- интерпретатор, позволяющий решать задачу в интерактивном режиме;
- графический редактор для проектирования систем и просмотра процессов;
- генератор кода описывающего СУД для использования разработанных моделей в реальном тренажере.

На рис. 1 показан внешний вид диалоговое окно редактора.

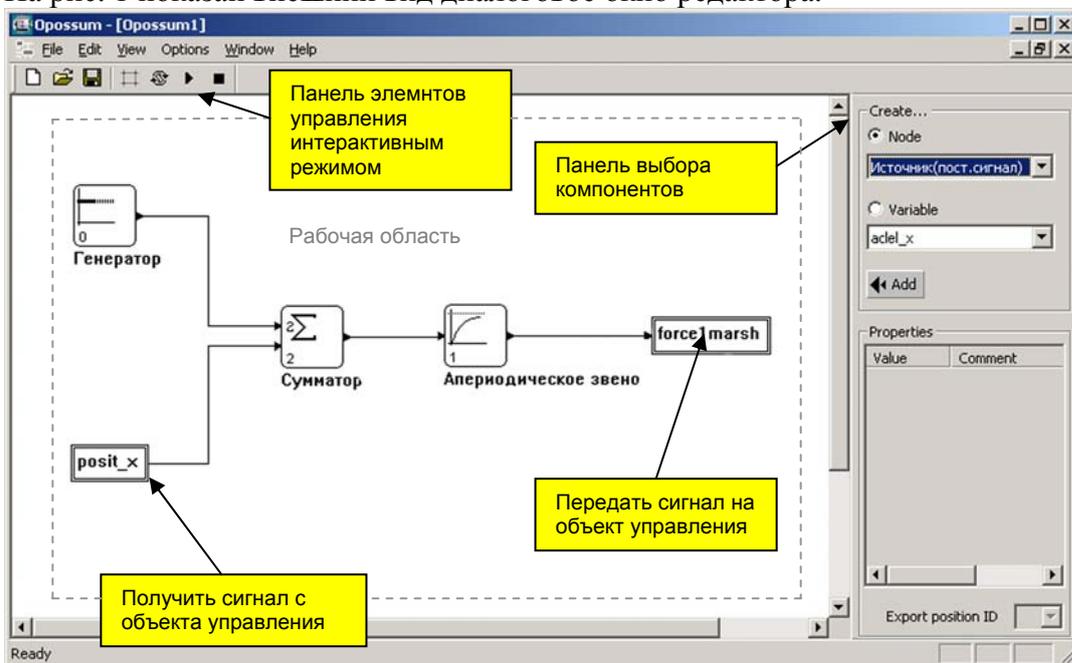


Рис.1. Диалоговое окно редактора структурных схем

Разработка системы производится путём составления графической схемы внутри рабочей области редактора. Панель выбора компонентов содержит доступные блоки и ссылки на параметры объекта.

Пример описания СУД показан ниже:

b0	марш	12	0.0	0							
b9	АЗПогр	0	1.00	100.00	-100.00						
b10	Сумм(6)_1	10	-1.00	-1.00	1.00	1.00	-1.00	0.00	100.00	-100.00	
b16	Сумм(6)_7	10	1.00	0.00	0.00	0.00	0.00	0.00	100.00	-100.00	
w0	omeg_kren	b7	i0								
w10	forcelag	b16	o0								
w11	Wire_12	b0	o0	b10	i0	b11	i0	b12	i0	b13	i0
w20	Wire_21	b9	o0	b14	i3	b15	i3				

Выделим основные свойства такой системы:

- Универсальность - возможность моделирования любой системы на базе имеющихся библиотек, возможность наращивания библиотек на уровне кода программы, что обеспечивает создание наиболее быстрых и оптимальных алгоритмов.
- Гибкость - возможность модернизации системы во внешних модулях без корректировки кодов, благодаря тому, что разделены уровень программиста и разработчика СУД.
- Скорость переналадки - модернизация может осуществляться путём визуального моделирования, либо в любом текстовом редакторе.
- Совместимость - общий язык описания систем.

Таким образом, мы определили основную область разработки - алгоритмический уровень управления, выбрали в качестве основной модели математическую модель, разделили уровни программиста и разработчика систем, предложив язык структурированного описания систем и методы их усовершенствования.

ВОПРОСЫ МОДЕЛИРОВАНИЯ ОБЪЕКТОВ ВИРТУАЛЬНОГО ПОДВОДНОГО МИРА В ТРЕНАЖЕРАХ ПОДВОДНЫХ ТЕЛЕУПРАВЛЯЕМЫХ КОМПЛЕКСОВ

Н.Буйда, Кошель Р., М. Григорьев

Московский Государственный Технический Университет им. Н.Э. Баумана,
кафедра СМ11

QUESTIONS OF MODELLING OF OBJECTS OF THE VIRTUAL UNDERWATER WORLD IN SIMULATORS OF UNDERWATER LONG-DISTANCE COMPLEXES

N.Buyda, R.Koschel, M.Grigoriev

Bauman Moscow State Technical University,
Department SM11

Аннотация

В работе представлены результаты разработки тренажерных комплексов, созданных на кафедре СМ-11 СМ МГТУ им. Н.Э.Баумана.

Abstract

In work results of development of the training complexes created on faculty of Sm - 11 of SM of MGTU n.a. N.E.Baumana.

Проблематика моделирования водной среды

В работе описывается подход к формированию моделей виртуального пространства подводного мира применительно к тренажерам для обучения операторов подводных аппаратов. Модели подводного мира в общем случае представляют собой упорядоченную совокупность объектов, имеющих физические параметры, оптические и гидроакустические свойства, геометрические представления, совокупность математических и логических зависимостей, отражающих структуру и отношения между элементами мира, а также между движущимися объектами. Математические и геометрические модели являются обобщенными и предназначены для описания определенного класса объектов. Специфика выполняемых задач в тренажерном комплексе обуславливает необходимость одновременного поддержания трех типов моделей виртуального мира.

Рассматриваются модели следующего вида:

- *модель визуализации* для формирования телевизионных изображений. Качество и точность модели определяется субъективными требованиями восприятия оператором картины подводного мира и скорости перерисовки кадров;
- *модель гидроакустической обстановки* подводного мира. Качество определяется требованиями по визуализации гидроакустических изображений подводных объектов;
- *модель механического взаимодействия* для решения задач динамики движения с учетом возможности коллизий подводных объектов в водной среде.

При требованиях общности геометрии модельного пространства в этих трех моделях возникают ограничения по степени дискретизации представления в виде сеток и ограничивающих объемов с заданными физическими свойствами составляющих элементов.

Модель визуализации

Основной критерий качества модели визуализации – реалистичность синтезируемого изображения.

Возможности систем машинной графики:

- передача представление о глубине сцены с использованием кинематографических приёмов;
- формирование очертаний предмета и его характерных линий, передающих его объемность, экранирование одних предметов другими, и, как следствие, наложение

изображений переднего плана на изображение предметов в глубине, передача перспективы;

- визуализация геометрии объектов, аппроксимированной полигональной сетью. Точность аппроксимации существенно влияет на качество синтезируемого изображения;
- возможность создания оптических и световых эффектов в виде ореолов источников света, возникновения замутнения, взрывов и т.п.;
- возможность наложения текстур на поверхность объектов;
- учёт оптических и светотехнических свойств объектов, участвующих в формировании сцены.

При визуализации водной среды учитываются:

- характеристики оптических свойств воды;
- освещенность водной среды естественным светом;
- освещенность искусственным светом подводного аппарата и других источников;
- характеристики пропускания и обратного светорассеяния;
- характеристики светорассеяния при освещенности донных объектов искусственным светом.

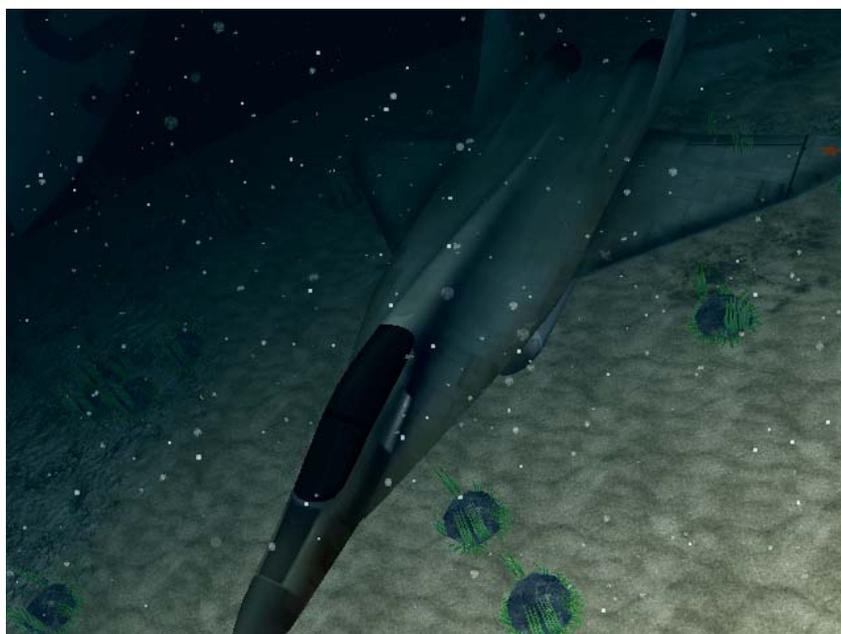


Рис. 1. Визуализация модели дна с затонувшим объектом

Модель морского дна

Модель подводного мира включает в себя модель поверхности морского дна (см. рис. 1).

Модель донной поверхности предполагает:

- построение генеральной поверхности дна на основе методов фрактальной геометрии, что позволяет моделировать районы морского дна практически неограниченного размера;
- использование отметок глубин, полученных с геоинформационных цифровых карт морского дна;
- применение интерполяции малых областей дна с повышенной степенью детализации;
- применение мультитекстурирования дна.

Модель гидроакустической обстановки

Гидроакустическая модель учитывает излучающие характеристики гидролокаторов, гидроакустические распределенные характеристики водной среды и отражающие свойства объектов гидролокации (поглощение, отражение, рассеяние). Эти характеристики сложным образом зависят от формы объектов.

По этой причине для задач расчета эхосигнала рационально заменять сложные по конфигурации цели эквивалентными упрощенными по отражающей способности телами стандартной формы, такими, как сфера, цилиндр и т.п. Гидроакустические свойства цели со сложной пространственной конфигурацией различны для различных положений источника и приемника гидроакустического сигнала относительно условного центра объекта локации. Отражающие свойства описываются пространственной диаграммой отражающей способности объекта, которая должна быть задана для локальных объектов гидролокации. При расчете отражающих свойств дна необходимо задание пространственных распределений отражающих характеристик.

При решении задач эхолокации необходимо учитывать аномалии распространения звука от излучателя к приемнику в зависимости от скорости звука и свойств водной среды.

Учет наложения прямого сигнала и сигнала, отраженного от поверхности моря в направлении зеркальном, относительно падающей волны значительно усложняет расчетную модель, и поэтому расчёт приходится проводить в приближении Ламберта. В ряде случаев при отражении звука от неровной поверхности моря происходит рассеивание сигнала, которое может быть учтено только на основе упрощений. Значительный разброс значений сигнала связан с флуктуациями акустических сигналов вдоль трассы распространения.

Из гидроакустических свойств в модели учитываются:

- акустические характеристики дна;
- акустические и аномальные свойства водной среды;
- диаграмма излучения эхолокатора.

Модель механического взаимодействия

Механическое взаимодействие объектов сцены при симуляции работы подводного комплекса может быть сведено к рассмотрению следующих задач:

- функционирование манипуляторов при работе со связанными объектами среды;
- динамическое движение подводного аппарата в условиях механического взаимодействия с дном и другими объектами.

Поскольку объекты, участвующие в механических контактах, представляют собой геометрические тела с сеточным полигональным описанием, возникает два конфликтных фактора, касающихся точности представления геометрии:

- требование увеличения числа элементов разбиения для более правильного определения механического взаимодействия;
- требование уменьшения числа элементов с целью сокращения объёма вычислений.

Последний фактор особенно влияет на качество моделирования ударов, которые могут возникать между аппаратом и объектами. Поэтому при рассмотрении коллизий необходимо иметь упрощенную модель взаимодействий, полученную из условия допустимого проникновения объектов взаимодействия друг в друга при ударе.

Выводы

Для достаточного представления виртуального подводного мира в тренажерах подводных телеуправляемых комплексов необходимо поддерживать три вида геометрических моделей - визуализационная модель, модель гидроакустики, модель механического взаимодействия движущихся объектов и окружающей обстановки.

МЕТОД ФОРМИРОВАНИЯ МОДЕЛИ ДИНАМИКИ ПОДВОДНОГО АППАРАТА В ТРЕНАЖЕРНЫХ СИСТЕМАХ

С.Медведев, А.Гнеломедов, Климов Е.

Московский Государственный Технический Университет им. Н.Э. Баумана,
кафедра СМ11

METHOD OF FORMATION OF MODEL OF DYNAMICS OF THE UNDERWATER DEVICE IN TRAINING SYSTEMS

S.Medvedev, A.Gnelomedov, E.Klimov
Bauman Moscow State Technical University,
Department SM11

Аннотация

В тезисах рассматриваются вопросы представления динамики движения подводного аппарата в водной среде в тренажерном комплексе для обучения операторов.

Abstract

In theses questions of representation of dynamics of movement of the underwater device in the water environment in a training complex for training operators are considered.

Тренажёрные системы на сегодняшний день являются интерактивными системами реального времени, поэтому модель динамики аппарата в условиях возможности возникновения новых связей и коллизий, должна обеспечивать достаточное быстродействие при вычислении в процессе выполнения тренировки. Машинный такт решения уравнений динамики не должен превышать 50мс.

В тренажерах подводный аппарат и другие динамические объекты представляются в виде моделей абсолютно твёрдых тел с заданными геометрическими описаниями. Предполагается, что форма этих тел не изменяется в процессе взаимодействия. В отличие от решений, использующих метод конечных элементов, описанный в [6], предлагаемое решение легче реализуемо в программном обеспечении. Для моделирования гибких объектов, например кабеля или троса, используется метод Виттенберга: представление в виде совокупности твёрдых тел, соединённых упругими связями.

Существует несколько подходов для моделирования движения твердых тел в задачах динамики реального времени. Известно, что для системы твёрдых тел широко используется метод уравнения Лагранжа второго рода. Он состоит в нахождении значений независимых параметров системы из уравнений Лагранжа. Значения всех остальных параметров определяются из уравнений связей между телами. Такой метод подходит для построения моделей систем тел с заранее определёнными связями. Поэтому, в нашем случае, когда в процессе симуляции могут возникать новые связи, использование известных подходов представляет значительные трудности при реализации в интерактивном приложении.

В теории автоматического управления для описания объектов используются переменные состояния. Изменение этих переменных определяется из дифференциальных уравнений первого порядка.

На основе такого подхода в работе [4] освещается метод моделирования движения свободных твёрдых тел. В качестве переменных состояния рассматриваются: положение и ориентация тела в пространстве, также их первые и вторые производные - линейные и угловые скорости и ускорения. Модель включает методы нахождения значений этих переменных:

- изменение ускорений объекта под действием внешних воздействий описывается уравнениями на основе второго закона Ньютона и следствий из него;
- скорости, положение и угловая ориентация тела находятся численными методами из соответствующих дифференциальных уравнений.

Модель свободного твёрдого тела можно использовать для описания движения подводного аппарата или другого объекта динамики, если воздействие водной среды и других тел описывать с использованием терминов внешних воздействий.

Описание воздействие водной среды на аппарат включает модель гидродинамики. В ней используются уравнения из [1], упрощённые за счёт исключения слагаемых зависящих от ускорения тела. Эти уравнения раскрывают зависимость гидродинамических сил от параметров тела: линейной и угловой скорости объекта, матриц присоединённых масс и моментов. Также на тела в водной среде действуют сила тяжести и сила Архимеда. Их расчёт описан в [5].

Из всех случаев взаимодействия между объектами среды рассматриваются два случая:

- взаимодействие продолжительного характера, например, между подводным аппаратом и поддерживающим его тросом;

- ударные взаимодействия - при ударах аппарата или при падении объекта на дно.

Продолжительные взаимодействия могут быть описаны связями. Связь – это условие, налагаемое на движение объекта. Например, шарнирная связь ограничивает перемещение точки объекта, к которой она приложена. В основе моделирования связей лежит процесс минимизации ошибки связи, которая показывает величину нарушения условия связи. Управление ошибкой происходит за счёт действия сил реакции связи на объект, которые определяются из матричного уравнения $-C = dC/dR * \delta R$,

где C – текущая ошибка связи,

δR – вектор определяющий изменение сил реакции,

dC/dR – якобиан связи, который отражает влияние изменения ошибки связи на изменение сил реакции.

Модель удара строится на основе положений классической стереомеханики, представленной в [2]. Такая модель позволяет рассматривать не прямой удар, и неявно учитывать пластическую деформацию поверхностей тел через коэффициент восстановления Ньютона.

Необходимым условием ударного взаимодействия тел является наличие области контакта их поверхностей. В некоторых случаях в районе этой области скорость сближения тел равна нулю или пренебрежимо мала. В таких ситуациях модель удара заменяется на модель контакта покоя, описанная в [4].

Литература

1. Привязные подводные системы. Прикладные задачи статики и динамики.- СПб.: Изд-во С.- Петерб. ун-та, 2000.-324с.,ил.
2. Динамика систем с механическими соударениями. – М.: Международная программа образования, 1997. – 336 с. – ISBN 5-7781-0031-0
3. Добронравов В.В., Никитин Н.Н. Курс теоретической механики: Учебник для машиностроит. спец. вузов. – 4-е изд., перераб. И доп. – М.; Высш. школа, 1983. – 575с., ил
4. David Barraff : Physically Based Modeling: Principles and Practice. (C)1997, Robotics Institute, Carnegie Mellon University. Документ с сайта <http://www-2.cs.cmu.edu/~baraff/>.
5. Designing a class library for interactive simulation of rigid body dynamics / by Bart G.B. Barenbrug. - Eindhoven : Eindhoven University of Technology, 2000.
6. Complex Flexible Multibody Systems with Application to Vehicle Dynamics. Jorge AMBRÓSIO and João GONÇALVES. *Instituto de Engenharia Mecânica (IDMEC), Instituto Superior Técnico, Av. Rovisco Pais, 1041-001 Lisboa, Portugal*
Документ с сайта http://www.dem.ist.utl.pt/~NATO_ARW/.

СИСТЕМНОЕ КОГНИТИВНОЕ МОДЕЛИРОВАНИЕ КТР ЭВС

Кутаков С. В.

Научный руководитель: к.т.н., доцент Резчикова Е. В.
Кафедра ИУ4, МГТУ имени Н.Э.Баумана, г. Москва, Россия

SYSTEM COGNITIVE MODELLING THE DTD OF ED

Kutakov S. V.

Scientific adviser: candidate of technical science, senior lecturer Rezchikova E. V.
IU4 Department, Bauman Moscow State Technical University, Russia
KutakoffSV@mtu-net.ru

Аннотация

В статье рассматривается возможность применения креативной графики для целей конструирования ЭВС

Abstract

The possibility of applying the creative graphics for ED design properties is described in this article

Инженерная мысль имеет форму не только слов, но и образов. Не секрет, что большинство инженеров во всем мире имеет визуальную репрезентацию. Действительно подавляющая часть людей скорее обнаружат скрытый смысл в рисунке, чем в словесном описании физического явления. Левое полушарие привыкло оперировать словами и другими подобными символами. Что же касается правого полушария, то оно работает за счет необъятных ресурсов нашей памяти. Именно на этой бессознательной стороне мозга и происходит интерпретация всевозможных визуальных образов. Фигуры и рисунки способны натолкнуть исследователя на определенные ассоциации, подсказать новое направление решения технической задачи. Таким образом, создание системы когнитивного моделирования будет способствовать увеличению творческих способностей инженера.

Отыскание решения технической задачи с помощью творческих когнитивных стратегий поиска условно можно изобразить как воздействие на техническую задачу когнитивного оператора:

$$O = [\Gamma, \Xi, B]$$

где Γ — генерация поля технических решений;

Ξ — экспертная оценка полученных решений;

B — визуализация поля полученных решений, постановки задачи, наиболее перспективных решений.

Рассмотрение процесса отыскания технического решения, как воздействия на техническую задачу когнитивного оператора приводит к появлению в системе структурного синергетического эффекта. Когнитивный оператор позволяет наиболее полно рассматривать описание технической задачи, позволяет наиболее эффективно производить постановку задачи, формулирование общих и смежных задач. Затем оператор позволяет генерировать поле технических решений и производить оценку полученных решений. Причем оценка может осуществляться как непосредственно самим исследователем, так и с помощью базы знаний, составленной другими экспертами. Такой комплекс средств рассмотрения технической задачи способствует активации творческой составляющей исследователя, что приводит к появлению полезного эффекта, который нельзя было бы получить при использовании когнитивных стратегий поиска по отдельности.

Важной составляющей когнитивного оператора является визуализация. Визуализацию постановки задачи, поля технических решений можно осуществить с помощью когнитивных

карт. При разработке ЭВС широко применяется модульный принцип конструирования, базирующийся на системном подходе. Для визуализации системного подхода применяется когнитивная карта — Системный оператор (Рис.1). Системный оператор позволяет на девяти и более экранах описать всю структуру, разрабатываемого изделия. В центре оказывается разрабатываемый прибор, сверху и снизу отображается входимость объекта в более крупные системы и разбиение на более мелкие элементы. Слева отображается прототип изделия, что делает заимствование более эффективным. Справа отображается положение дел в будущем, помогая тем самым исследователю осуществить более полный анализ исходных данных и предсказать дальнейшее развитие объекта или технологии. По мере надобности в когнитивную карту может быть добавлены дополнительные экраны, описывающие над-над-систему и под-под-системы. Такой инструментарий позволяет наглядно отображать процедуру разукрупнения схем или конструкций на функциональные группы разных уровней сложности для реализации их конструктивными модулями. Здесь хорошо отражены внутрисистемные связи всех уровней, а также тенденции развития технических систем во времени.



Рис. 1. Программа, реализующая системный оператор.

Важным этапом разработки электронной аппаратуры является подготовительный этап, на котором оцениваются потребности рынка и народного хозяйства, для формулирования технического задания. Визуализация уточнения постановки задачи и формулирования направления поиска технического решения осуществляется за счет применения когнитивной карты — Метод фокальных объектов (Рис.2). Метод основан на выборе случайных объектов, описания свойств случайных объектов и далее группировании выбранных свойств с искомым фокальным объектом. Программа, реализующая его, позволяет просмотреть полученный варианты и отобрать наиболее интересные. Эта когнитивная карта применяется для увеличения творческого потенциала исследователя. Она позволяет создавать целый спектр концептуальных технических идей.

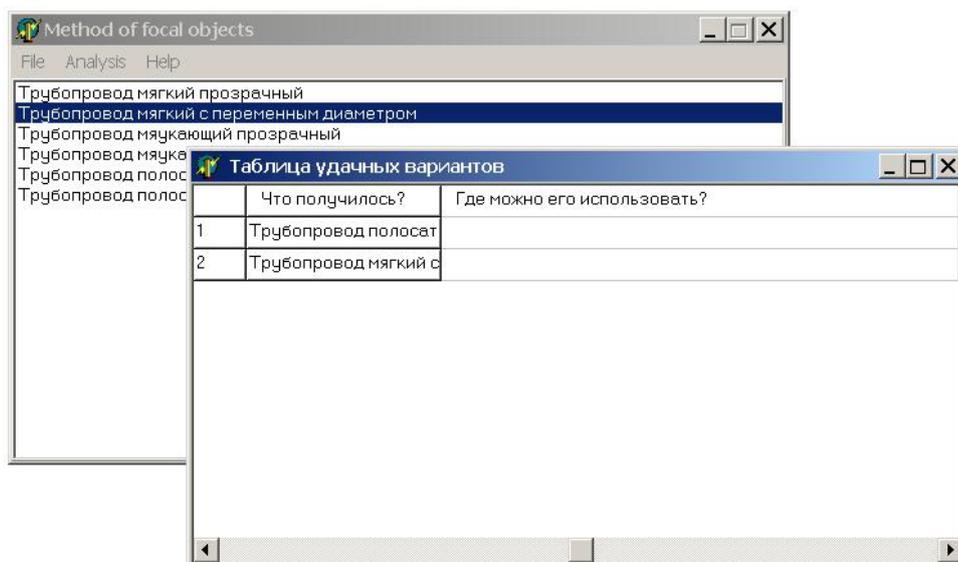


Рис. 2. Программа, реализующая метод фокальных объектов.

Для визуализации предположительного пути анализа и решения задачи, а также на этапе разработки структурных схем и технических предложений целесообразно применять когнитивную карту — Майнд-мэп. С помощью Майнд-мэпа составляется сценарий анализа исходных данных либо структурная схема изделия. Также Майнд-мэп предоставляет широкие возможности по схематическому представлению наиболее удачных вариантов решений и методические процедуры, призванные донести до исследователя методики генерации и оценки дивергентного поля технических решений.

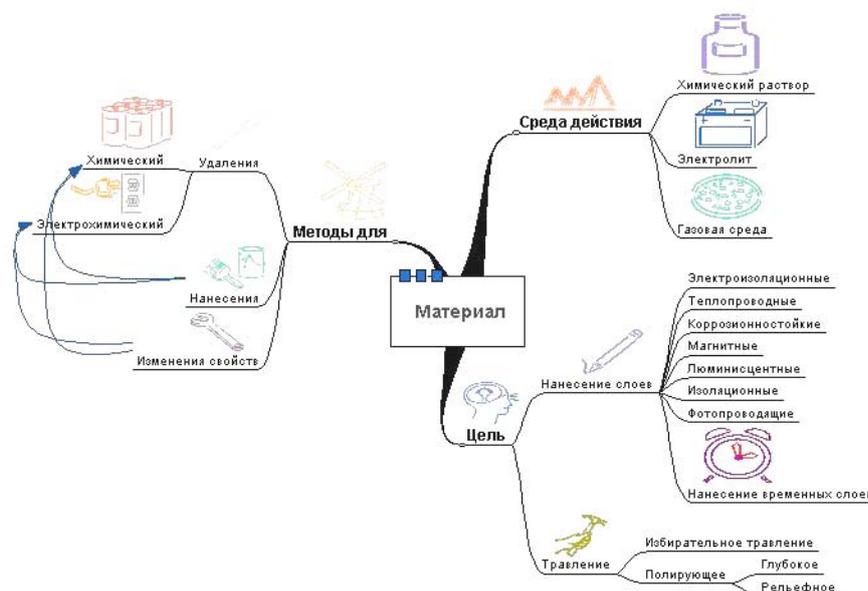


Рис. 3. Майнд-мэп «Нанесение покрытий»

Визуализировать этапы выбора системы элементов и разработки системы элементов позволяет когнитивная карта — Цветущий лотос. Эта система отображения данных позволяет детально описать проблему, основные её составляющие, вычленить более мелкие подзадачи для дальнейшей детальной проработки.

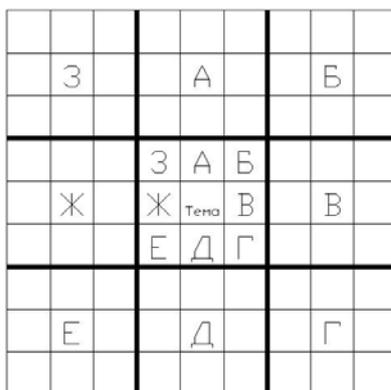


Рис. 4. Цветущий лотос.

Таким образом, использование системы когнитивного моделирования ЭВС позволяет увеличить творческие способности исследователя, организовать коллективную творческую деятельность и обеспечить опережающее создание инновационных баз данных в области проектирования электронной аппаратуры.

Литература

1. Конструкторско-технологическое проектирование электронной аппаратуры: Учебник для вузов / К.И. Билибин, А.И. Власов, Л.В. Журавлева и др. Под общ. ред. В.А. Шахнова. — М.: Изд-во МГТУ им. Н.Э. Баумана, 2002. — 528 с.: ил.
2. Алдер Г. НЛП-графика. Мышление в рисунках и образах. — СПб.: Питер, 2003. — 192 с.: ил.