

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Пензенский государственный технологический университет»

Дипломный проект на тему: «Информационная система голосового управления персональным компьютером»

Выполнил: студент гр. 10И2 Дамаев С.С.

Руководитель: к.т.н., доцент кафедры «Информационные
технологии и системы» Юрманов В.А.

Актуальность и цель работы

- повышение эффективности и удобства персонального компьютера;
- ограничение доступа к данным на персональном компьютере за счет голосовой аутентификации;
- интеллектуализация и автоматизация процесса обработки голосовых команд.
- Целью работы является разработка информационной системы, обеспечивающей сбор, обработку, анализ речевых команд, полученных от пользователя и управление персональным компьютером.

Задачи дипломного проектирования

- анализ предметной области и выявление в ней существующих проблем;
- оценка современного состояния исследований в области распознавания голосовых команд;
- анализа существующих решений в области голосового управления, отвечающих требованиям моделируемой системы;
- разработка подсистемы ввода информации;
- разработка подсистемы цифровой обработки сигналов;
- разработка подсистемы идентификации на базе нейронной сети;
- разработка имитационной модели информационного процесса передачи информации, планирования и проведения имитационных экспериментов;
- расчет рентабельности предприятия по производству ИС ;
- рассмотрение вопросов информационной

Современные программные продукты для распознавания ГОЛОСОВЫХ КОМАНД

Название системы	Назначение	Структурная единица	Обучение	Поддержка русского языка	Дикторо-независимость	Кэфф. расп., зависит от словаря	Стоимость
Voice Digger	Поиск ключевых слов	Слово, фраза	Да	Да	Да	до 97% ,	–
Тuple Premium 2.0	Голосовое управление ПК	Слово, фраза	Да	Да	Да	до 95%	0\$-13\$
Google Voice Search	Голосовой поиск	Слово, фраза	Нет	Да	Да	–	–
Dragon NaturallySpeaking	Голосовое управление, распознавание речи	–	Нет	Нет	Нет	до 98%	85-100\$
ViaVoice	Голосовое управление, распознавание речи	–	Нет	Нет	Да	до 95%	100-110\$

Диаграмма вариантов использования

ИСПОЛЬЗОВАНИЯ

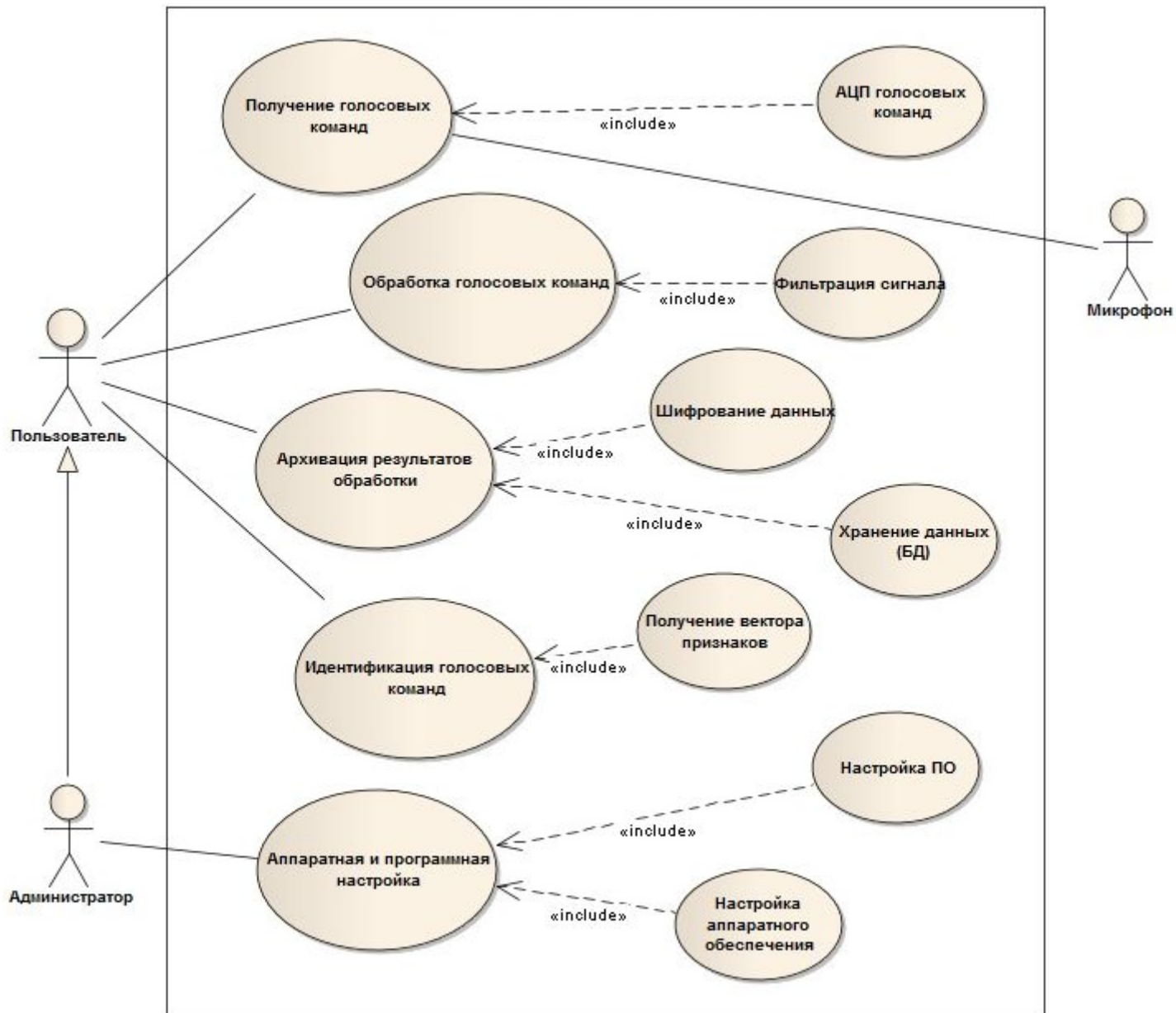


Диаграмма состояний

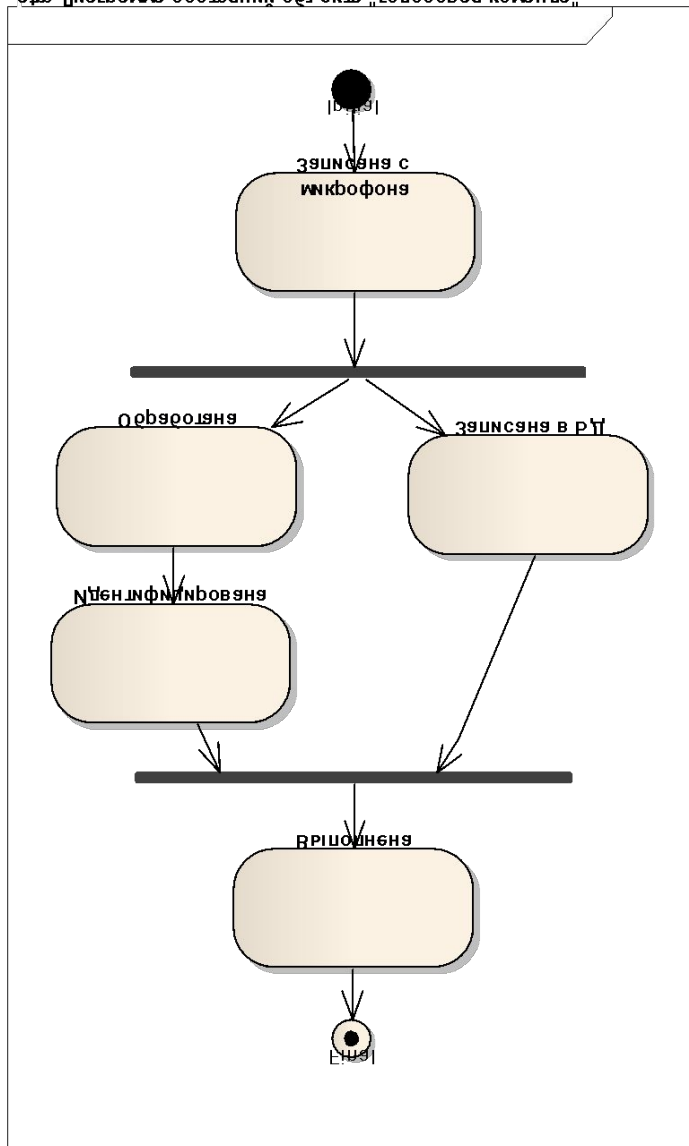
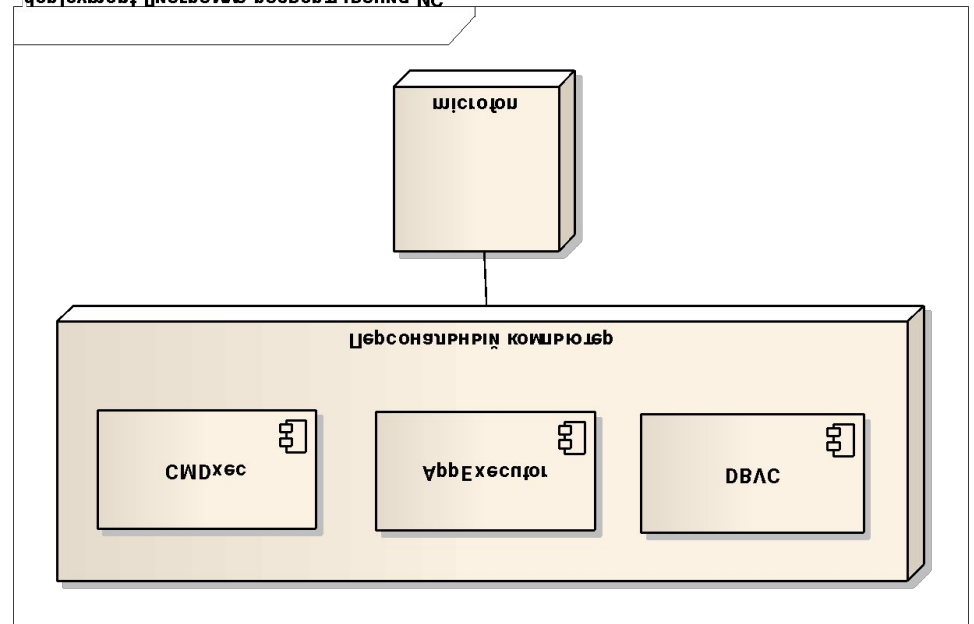
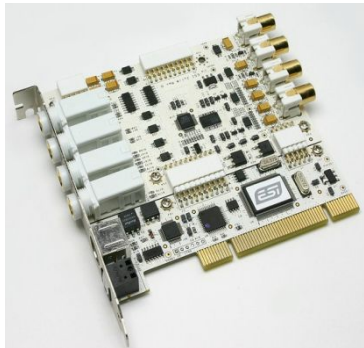


Диаграмма развертывания



Подсистема ввода информации

Технические характеристики звуковых плат



Внешний вид аудио карты Esi Juli

Модель	Разрядность ЦАП/АЦП, бит.	Максимальная частота АЦП, кГц.	Отношение сигнал/шум ЦАП/АЦП, дБ.	THD ЦАП/АЦП, %	Тип подключения	Кол-во входных аналоговых каналов	Другие интерфейсы	Цена, руб.
<i>Esi Juli</i>	24 / 24	192	112 / 114	0.0004 / 0.0004	PCI 2.2	2	S/PDIF	4400
ASUS Xonar Essence STX	24 / 24	192	124 / 118	0.0003 / 0.0002	PCI-E	2	S/PDIF	5800
M-Audio Audiophile 2496	24 / 24	96	104/100	0.002/ 0.002	PCI 2.1	2	S/PDIF MIDI	3700
INFRASONIC Quartet	24 / 24	192	97 / 100	0.0012 / 0.0012	PCI 2.2	4	S/PDIF	5000
M-Audio Audiophile 102	24 / 24	192	108 / 113	0.0023 / 0.0006	PCI 2.2	2	S/PDIF	7 4800

Выбор моделей микрофонов

Модель	Чувствительность	Частотный диапазон	Направленность	Сопротивление	Габаритные размеры, мм	Вес, г.	Цена, руб
<i>Genius MIC-01A</i>	-40 дБ	100 Гц - 10 кГц	всенаправленный	2.2 кОм	26 x 100 x 105	40	250
Speed-Link SL-8691-SBK	-58 дБ	100 Гц - 15 кГц	однаправленный	2.2 кОм	-	7	500
Dialog M-101W	-58 дБ	20 Гц - 16 кГц	всенаправленный	3 кОм	-	118	400
Ritmix RDM-111	-60 дБ	100 - 16000 Гц	всенаправленный	2.2 кОм	25 x 5 x 250	230	520
Panasonic RP-VC151	-48 дБ	100 - 20000 Гц	всенаправленный	1000 Ом	32 x 12 x 21	14	480
Sony ECM-PC50	-42 дБ	100 - 15000 Гц	всенаправленный	-	8 x 19 x 250	10	1 050

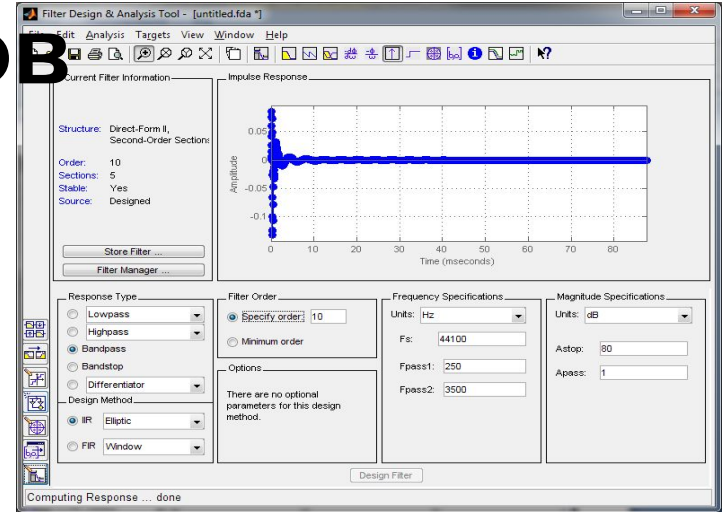
Требования к характеристикам микрофона:

- чувствительность: не хуже 40 дБ;
- частотный диапазон: от 100 Гц до 10 кГц – полностью перекрывает возможный диапазон голосовых команд пользователя ИС;
- направленность: всенаправленный.

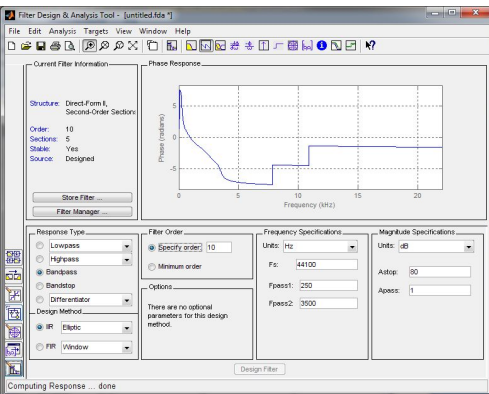
Подсистема цифровой обработки сигналов

Характеристики ППФ

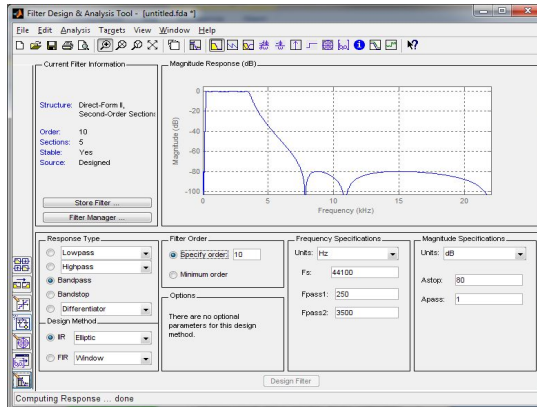
Наименование характеристики	Значение
Тип фильтра	ППФ
Тип импульсной характеристики	БИХ
Порядок фильтра	10
Метод построения	Эллиптический
Частота дискретизации	44100 Гц
Полоса среза F_{c1}	250 Гц
Полоса среза F_{c2}	3500 Гц
Коэффициент подавления	80



ИХ ППФ



ФЧХ ППФ



АЧХ ППФ

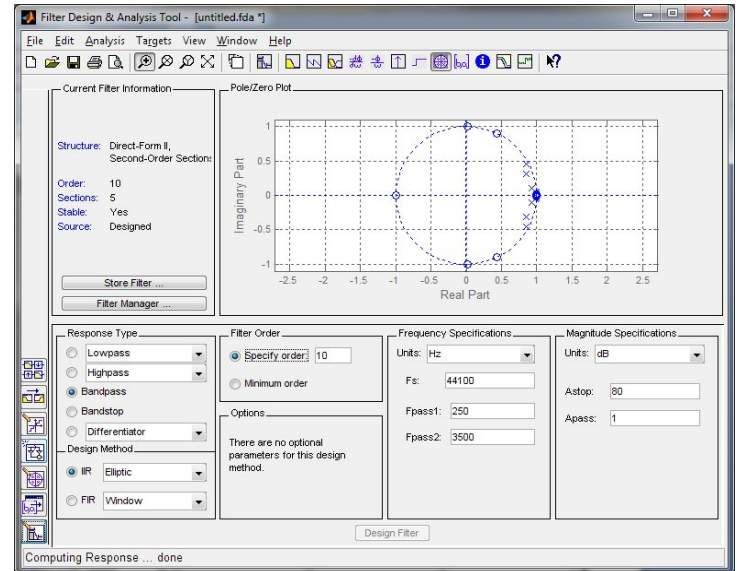
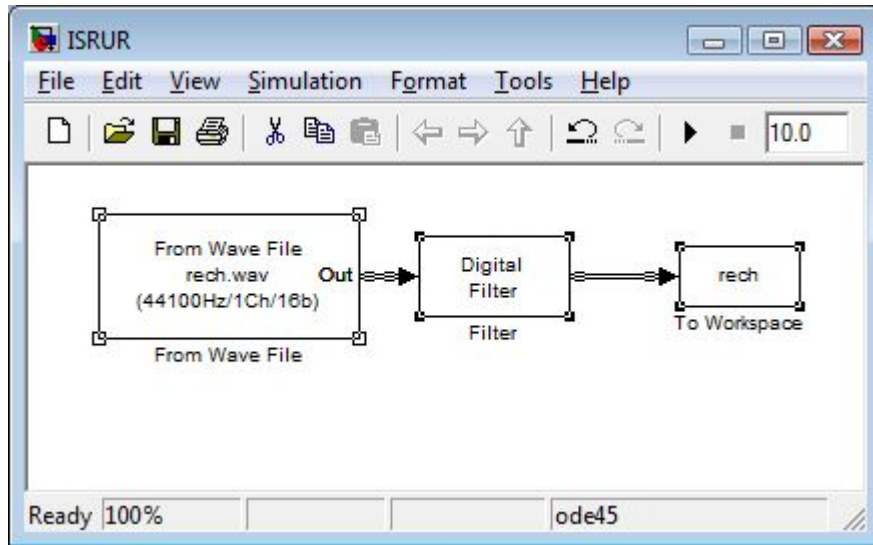
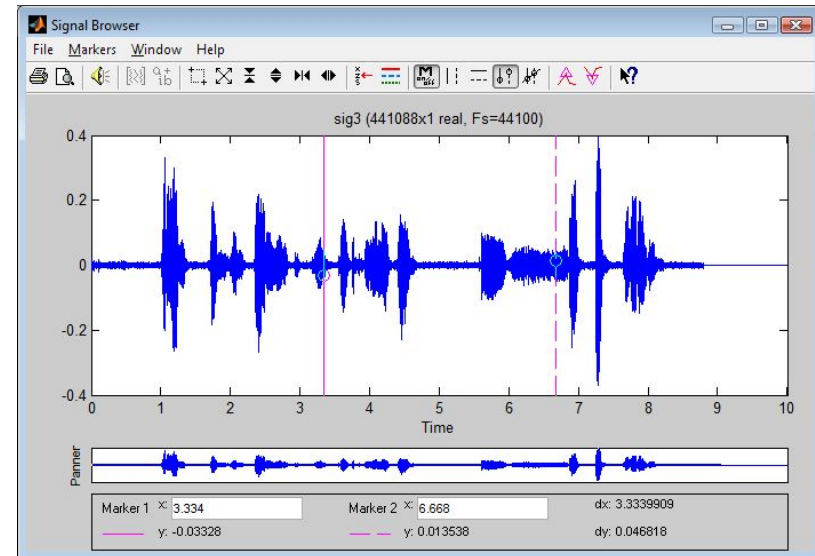


Диаграмма нулей и полюсов ППФ общий вид

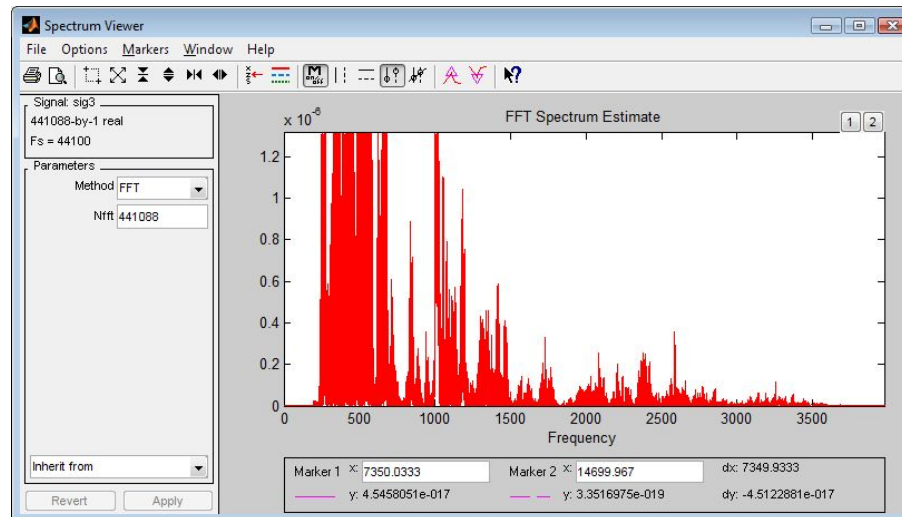
Фильтрация сигнала



simulink модель фильтрации



Речевой сигнал, прошедший фильтрацию



Спектр речевого сигнала

Подсистема идентификации

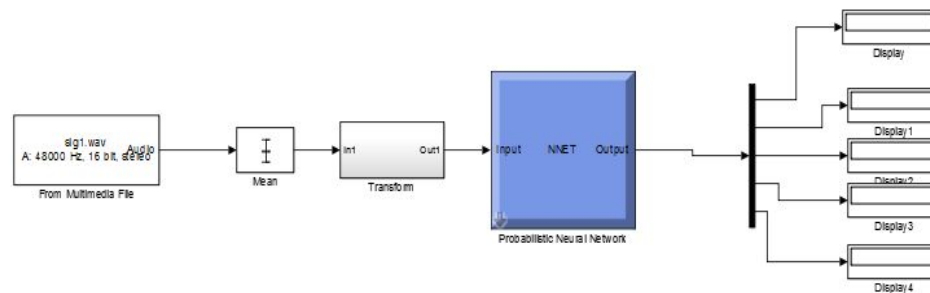
Структурная схема



Модель подсистемы идентификации

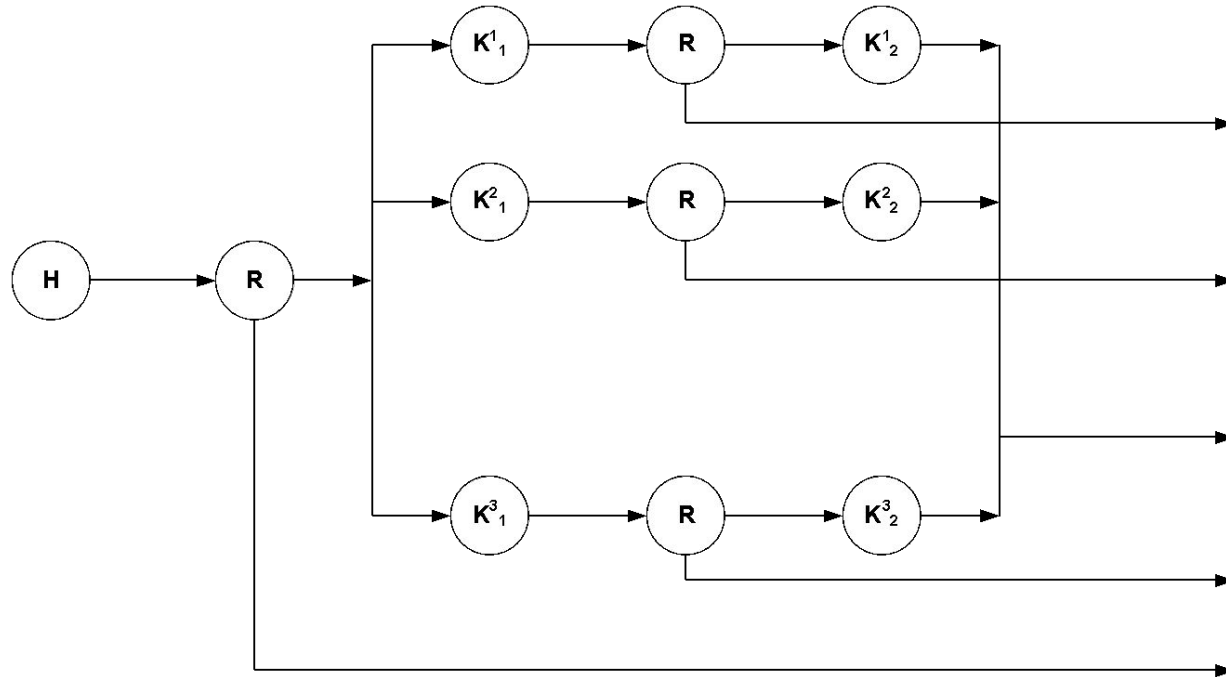
Список типичных команд клиента

Команда	Код
Да	1000000000000000
Вверх	0000000000001000
Вниз	0000000000000100
Вправо	0000000000000010
Влево	0000000000000001



Оптимизация информационных процессов

Структурная схема имитационной модели в символической Q - схеме



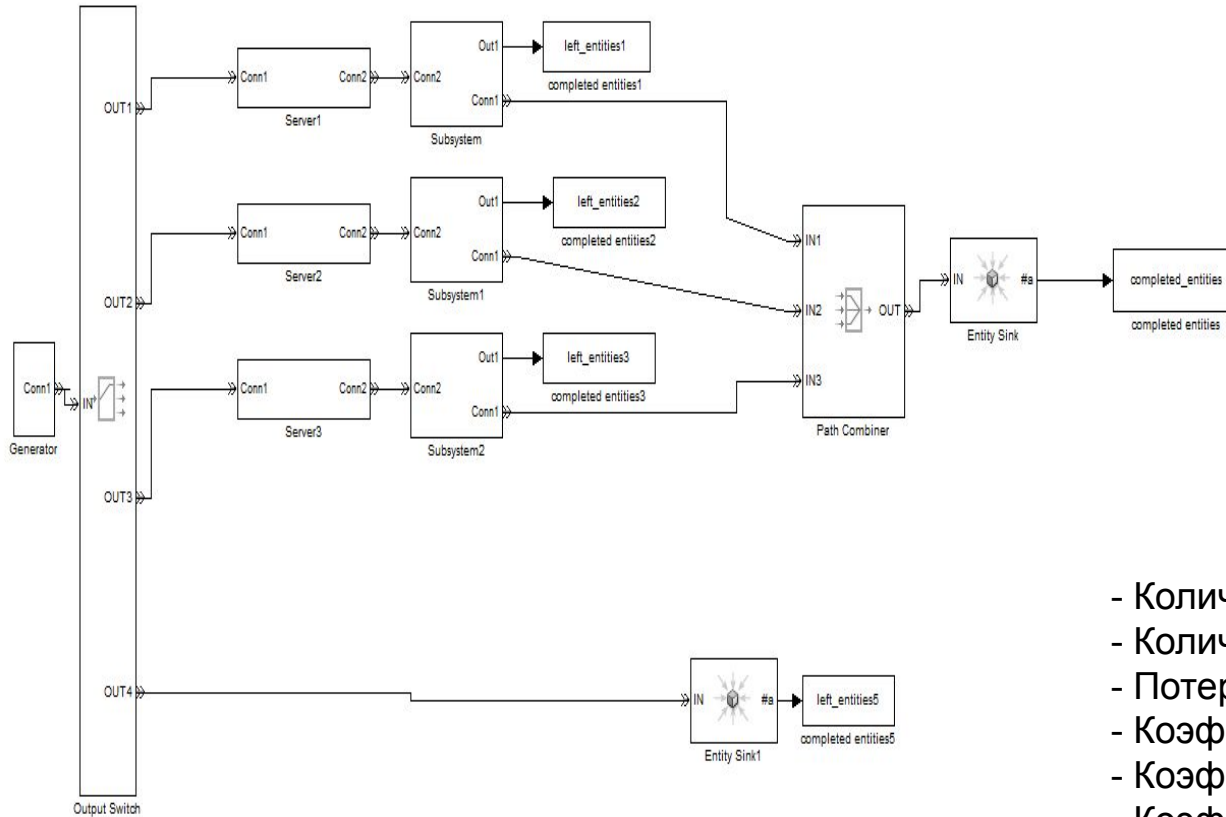
Здесь H (источник заданий) - имитирует поступление команд голосового управления;

R - распределитель – имитирует распределение команд по каналам и приборам обработки;

K_1 - имитирует процесс предварительной обработки команды управления;

K_2 - имитирует процесс идентификации команды управления.

Оптимизация информационных процессов



Имитационная модель процесса голосового управления

Результаты моделирования

- Количество созданных сущностей: 1230
- Количество обслуженных сущностей: 1224
- Потери на первом этапе проверки: 3
- Коэффициент загрузки 1 прибора: 1
- Коэффициент загрузки 2 прибора: 0.95408
- Коэффициент загрузки 3 прибора: 0.86639
- Потери на втором этапе проверки: 3
- Коэффициент загрузки 1 прибора: 0.87051
- Коэффициент загрузки 2 прибора: 0.84473
- Коэффициент загрузки 3 прибора: 0.81753

Планирование эксперимента

	-	0	+
X1	20 п/мин	40 п/мин	60 п/мин
X2	0,3± 0,1сек	0,6± 0,1сек	0,9± 0,1сек
X3	0,3± 0,15сек	0,6± 0,15сек	0,9± 0,15сек

Результаты эксперимента

Интенсивность вх. потока, ед/мин.	Потери на этапе обработки	Потери на этапе идентификации	Общие потери
35,44	3	3	6
30,88	0	2	2
26,32	2	2	4
21,76	1	1	2
17,2	1	2	3
12,64	2	3	5
8,08	2	0	2
3,52	1	1	2

Разработка корпоративной ИС предприятия

Ставка внутренней рентабельности

Точка безубыточности

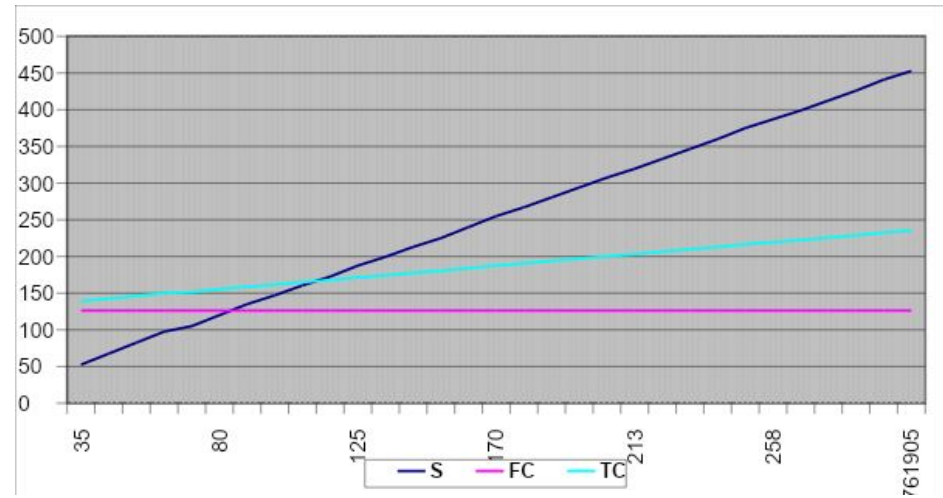
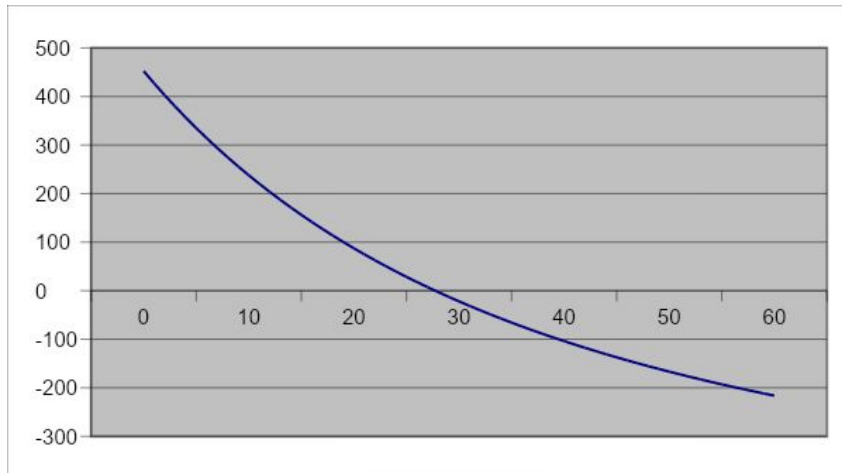
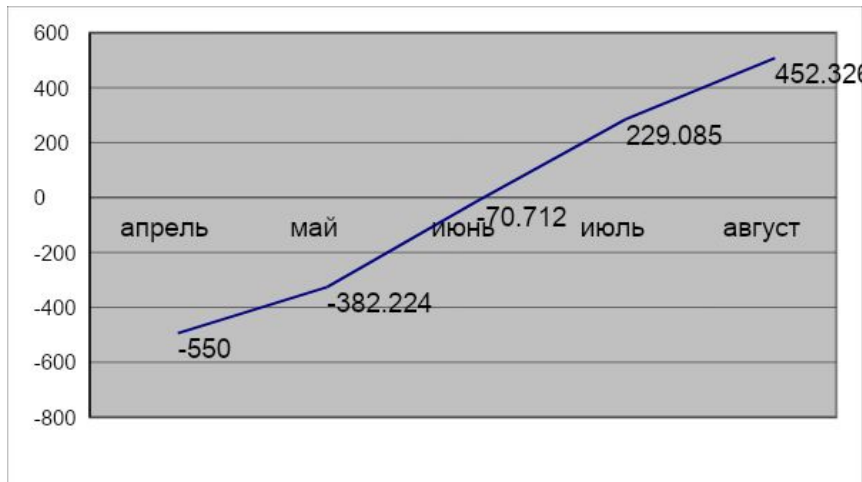
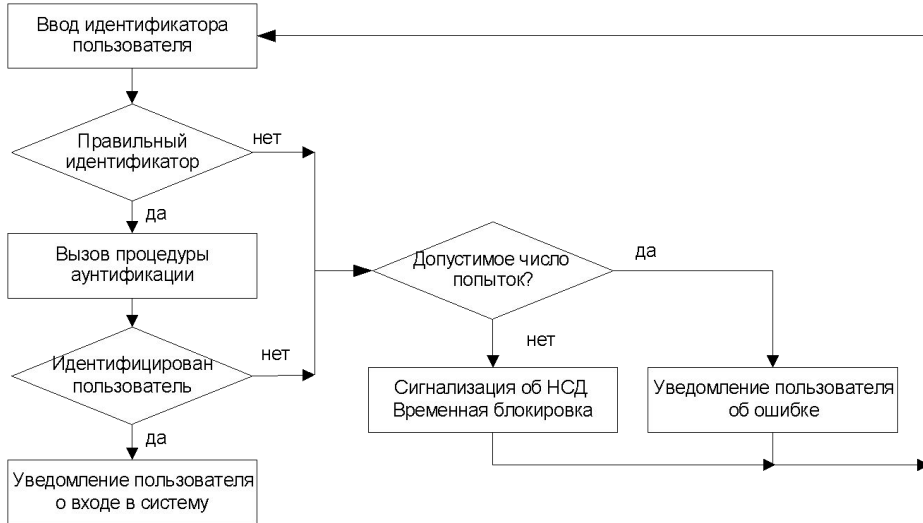


График окупаемости



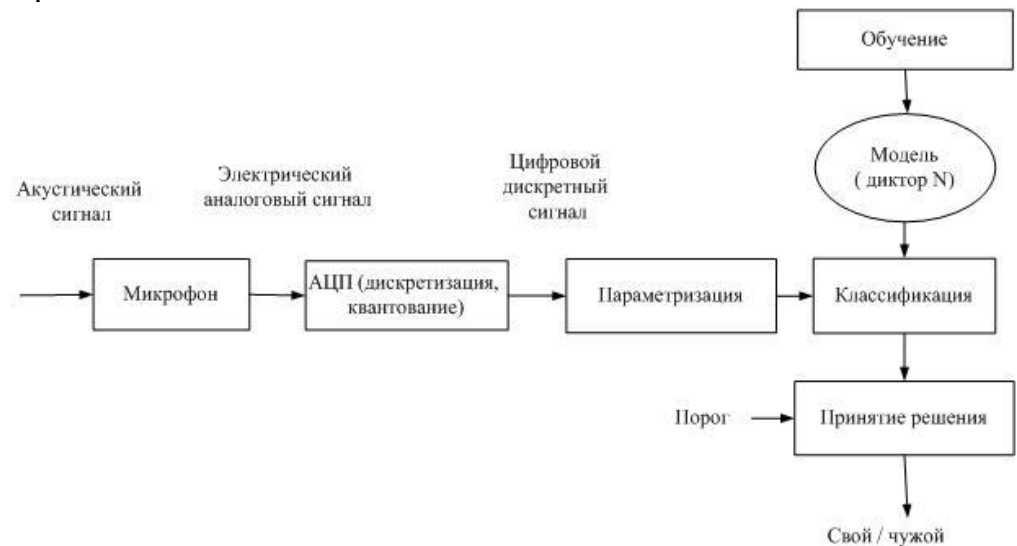
Показатель	Значение
Количество продукции, Q (шт./мес)	150
Цена единицы продукции, тыс. руб.	4000
Валовая выручка, $S(Q)$, тыс. руб.	600 000
Валовые издержки, TC , тыс. руб.	525 000
Критический объем продаж, Q^* , шт./мес.	115
Рентабельность проекта (показатель ROI)	0,4461
Норма прибыли	18%
Абсолютная эффективность капиталовложений, $NPV_i(10\%)$	317,498
Ставка внутренней рентабельности, IRR	27,73%
Индекс доходности, PI	1,335

Информационная безопасность



Классическая процедура идентификации и аутентификации

Схема аутентификации по голосу



Основные результаты и

Выводы

- проведен анализ предметной области;
- проведен анализ современного состояния исследований в области распознавания голосовых команд;
- проанализированы существующие решения в области голосового управления, отвечающие требованиям моделируемой системы;
- разработана подсистема ввода информации;
- разработан полосно-пропускной фильтр;
- разработана подсистема идентификации на базе нейронной сети;
- разработана имитационная модель информационного процесса передачи информации, планирования и проведения имитационных экспериментов;
- рассчитана рентабельность предприятия по производству ИС ;
- рассмотрены вопросы информационной безопасности.

Спасибо за внимание!