



открытое акционерное общество

ВЛАДИМИРСКОЕ КОНСТРУКТОРСКОЕ БЮРО РАДИОСВЯЗИ

Россия, 600009, г. Владимир, а/я 68

тел. (4922)43-15-54 • факс (4922)53-06-33 • E-mail: vkbrs@vkbrs.elcom.ru

«УТВЕРЖДАЮ»

Генеральный директор

ОАО «Владимирское конструкторское
бюро радиосвязи», к.т.н.



А.Е. Богданов
22.10.2014 г.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

на диссертацию «Разработка средств исследования и повышения помехоустойчивости систем автоматического распознавания голосовых команд в телефонии» Левина Евгения Калмановича, представленную на соискание ученой степени доктора технических наук по специальности 05.12.13 «Системы, сети и устройства телекоммуникаций».

Главным сдерживающим фактором использования систем автоматического распознавания голосовых команд (САРГК) в телефонии является высокая трудоемкость их разработки. Трудоемкость во многом определяется необходимостью обеспечения устойчивости САРГК к воздействию помех. Высокая степень изменчивости произнесения команд и разнообразие помех обуславливают проведение большого количества экспериментов при исследовании помехоустойчивости САРГК в процессе ее разработки. Эксперименты проводятся путем тестирования САРГК объемными выборками звукозаписей, в которых учитывается воздействие помех на систему. По результатам эксперимента оценивается достоверность распознавания в условиях воздействия помех - определяются относительные частоты пропуска команды и ложного срабатывания системы.

С расширением области использования САРГК в телефонии увеличивается разнообразие помех, и растут требования к достоверности распознавания команд, что ведет к увеличению объема экспериментов и, следовательно, к росту трудоемкости разработки систем.

Указанный подход к оценке достоверности распознавания используется для исследования помехоустойчивости в научных подразделениях. Но при разработке САРГК для практического применения такие трудозатраты становятся неприемлемыми. Целесообразно проанализировать возможность создания средств оценки помехоустойчивости САРГК, предназначенных для конкретного практического использования, которые позволят снизить трудоемкость разработки. Поэтому решаемая в диссертации задача создания средств исследования и повышения помехоустойчивости САРГК, обеспечивающих снижение объема экспериментов, проводимых при разработке указанных систем, является **актуальной**.

Научная новизна диссертации характеризуется следующими результатами.

1. Предложен критерий достоверности распознавания команды, который определяется при тестировании САРГК выборкой звукозаписей. В процессе тестирования определяются разности метрик сопоставления произнесений данной команды со «своей» и ближайшей «чужой» акустическими моделями – формируется выборка разностей метрик. В качестве критерия используется отношение выборочного среднего к выборочному среднеквадратическому отклонению.

2. Разработан метод сравнения результатов тестирования САРГК по достоверности распознавания команд при изменении параметров речевых сигналов и акустических моделей. Изменения параметров речевых сигналов обусловлены действием помех, а изменения параметров акустических моделей осуществляются в процессе разработки системы. Метод использует указанный выше критерий достоверности распознавания команды. Применение метода позволяет сократить объем тестовой выборки звукозаписей по сравнению с использованием в качестве критерия относительной частоты пропуска команды. Получены выражения, позволяющие определить минимально допустимый объем тестовой выборки звукозаписей при заданной вероятности ошибки сравнения результатов тестирования по достоверности распознавания команды.

3. Разработан метод оценки вероятности ложного срабатывания САРГК на произнесения слов, не являющихся командами, в зависимости от степени соответствия «некомандных» произнесений акустическим моделями команд. Метод обеспечивает снижение зависимости оценки от состава тестовой выборки звукозаписей.

4. Получены выражения, отражающие зависимость степени подавления квазипериодической помехи с помощью неадаптивной компенсации от частоты дискретизации сигналов, а также от соотношения уровней аддитивного шума и периодического компонента с меняющимся периодом. Компенсация осуществляется вычитанием задержанной во времени копии помехи, находящейся в паузе речевого сигнала, из суммы сигнала и помехи.

5. Получены выражения, определяющие зависимость степени подавления широкополосных радиопомех совместно с их отражениями от погрешности

реализации требуемых задержек в адаптивных трансверсальных фильтрах компенсатора помех.

6. Предложен алгоритм снижения влияния фазоманипулированного сигнала на работу цепей адаптации компенсатора помех. Алгоритм использует обратную связь по решению.

Теоретическая значимость полученных в диссертации результатов заключается в следующем.

1. Проведен анализ метрики сопоставления произнесения с акустическими моделями помехи, «своей» и «чужой» команд. Выделены основные типы метрик сопоставления произнесенного звука с монофонными и трифонными акустическими моделями звуков. Тип метрики зависит от степени соответствия произнесенного звука и его окружения акустической модели.

2. Оценена вероятность принятия ошибочного решения при сравнении результатов тестирования САРГК по достоверности распознавания команд с использованием предложенного критерия достоверности распознавания.

3. Проанализировано влияние аддитивного шума и частоты дискретизации сигнала на результат оценки периода помехи при компенсации квазипериодических помех.

4. Получено выражение оптимального вектора управляющих коэффициентов для процессора Хоуэлса-Эпплбаума, используемого для адаптивной компенсации широкополосных радиопомех с отражениями;

5. Исследовано влияние обратной связи по решению, используемой для уменьшения влияния сигнала на работу цепи адаптации, на степень подавления помехи с помощью адаптивной компенсации.

Практическая значимость работы заключается в следующем.

1. Использование предложенного критерия сравнения результатов тестирования САРГК, по достоверности распознавания команд позволяет уменьшить объем тестовой выборки звукозаписей.

2. Разработана методика выявления помехи и голосовой команды, которым соответствует нижняя граница достоверности распознавания. Методика использует предложенный критерий достоверности распознавания команды, что позволяет снизить объем тестовой выборки звукозаписей по сравнению с известным методом.

3. Разработана методика оценки вероятности ложного срабатывания САРГК на произнесения слов, не являющихся командами, которая позволяет снизить объем экспериментов при необходимости повторной оценки, когда меняется состав «некомандных» слов в условиях эксплуатации системы. Разработано программное обеспечение методики.

4. Разработан набор функций системы Matlab для исследования алгоритмов компенсации помех на основе имитационного моделирования.

5. Разработаны программно-аппаратные средства, обеспечивающие сбор данных о появлении ошибок распознавания команд в процессе эксплуатации САРГК на телефонной линии.

6. Предложены устройства компенсации радиопомех, защищенные авторскими свидетельствами на изобретения.

7. Разработаны структурные схемы комплекса аппаратуры для испытаний устройств компенсации радиопомех.

Структура и объем работы

Диссертация состоит из введения, шести разделов, заключения, списка литературы, имеющего 207 наименований отечественных и зарубежных источников, в том числе 70 работ автора, и приложений. Общий объем диссертации составляет 257 страниц, в том числе 201 страницу основного текста, 33 страницы списка литературы, 50 рисунков, 16 таблиц и 22 страницы приложений.

Во введении обоснована актуальность работы, сформулированы цели и задачи исследований, дана характеристика научной новизны и практической значимости результатов работы.

В первом разделе диссертации дана краткая характеристика основных этапов обработки речевых сигналов (РС) при их распознавании. Подавление помех, создаваемых самим диктором, его акустическим окружением и каналом связи, осуществляется как на стадии предварительной обработки РС, так и на стадии сопоставления параметров РС с акустическими моделями команд, пауз и помех. Для подавления помех на стадии предварительной обработки сигналов наиболее перспективно использование компенсации помех, так как здесь в наименьшей степени изменяются параметры РС, сопоставляемые далее с акустическими моделями.

Подавление помех на стадии сопоставления параметров РС с акустическими моделями осуществляется использованием акустических моделей отдельных звуков, из которых состоят произнесения слов, не являющихся командами, а также отдельные модели помех.

Помехоустойчивость САРГК оценивается степенью снижения достоверности распознавания - увеличения относительных частот пропуска команды и ложного срабатывания системы при воздействии помех. Указанные частоты определяются тестированием САРГК объемными выборками звукозаписей речевых сигналов. С расширением области использования САРГК растут требования к достоверности распознавания команд, и увеличивается разнообразие помех, что приводит к необходимости увеличения объема тестовых выборок звукозаписей. Трудоемкость разработки САРГК увеличивается.

Во втором разделе проведен анализ метрик сопоставления параметров РС с акустическими моделями, входящими в состав сети моделей. Рассматривается несколько типов метрик сопоставления отдельного звука произнесения с различными акустическими моделями. Акустической модели команды соответствует несколько комбинаций типов метрик, которые обуславливают ложное срабатывание САРГК в виде этой команды на произнесение, не являющееся командой. Каждая комбинация обозначается кодом, для которого и определяется относительная частота ложного срабатывания.

Если меняется состав «посторонних» произнесений, то меняется только часть состава кодовых комбинаций, вызывающих ложное срабатывание, соответствующее данной команде. Это обстоятельство позволяет уменьшить тестовую выборку звукозаписей при определении относительной частоты ложной тревоги для нового состава «посторонних» произнесений. Используются в выборке только те звукозаписи, которые соответствуют новым кодовым комбинациям. Проведены объемные эксперименты, показывающие, что при изменении состава «посторонних» произнесений относительная частота ложного срабатывания для данной кодовой комбинации и данной команды меняется мало. Данный вывод справедлив не только для отдельной команды, но и для целого списка команд.

Предложен критерий достоверности распознавания команды, который определяется при тестировании САРГК выборкой звукозаписей. В процессе тестирования определяются разности метрик сопоставления произнесений данной команды со «своей» и ближайшей «чужой» акустическими моделями – формируется выборка разностей метрик. В качестве критерия используется отношение выборочного среднего к выборочному среднеквадратическому отклонению.

Использование критерия позволяет уменьшить объем тестовой выборки звукозаписей при сравнении результатов тестирования по достоверности распознавания команды, когда меняются значения параметров РС и параметры акустических моделей. В частности, из результатов моделирования следует возможность снижения объема выборки на 31% при заданной вероятности ошибки сравнения 5%, если вероятности пропуска команды сравниваемых результатов тестирования равны 10% и 13%.

В третьем разделе экспериментально исследуется влияние шумовых, импульсных и гармонических помех, присутствующие на интервале длительности команды на достоверность ее распознавания. Использование предложенного критерия достоверности распознавания позволяет определить влияние помехи даже в том случае, когда использование в качестве критерия относительной частоты ошибки распознавания не позволяет заметить это влияние.

Проведены эксперименты по исследованию влияния различных помех на огибающую кратковременного спектра синтезированного сегмента РС. Отмечена немонотонность зависимости искажений огибающей от отношения сигнал-помеха. Проведен эксперимент для оценки эффективности метода подавления помех, который используется в стандарте Европейского института стандартов в области телекоммуникаций ETSI ES 202 212 V1.1.2 (2007-01).

Рассмотрена возможность подавления зашумленной периодической помехи с изменяющимся периодом методом компенсации. Получены выражения, отражающие зависимость степени подавления квазипериодической помехи с помощью неадаптивной компенсации от частоты дискретизации сигналов, а также от соотношения уровней аддитивного шума и периодического компонента с меняющимся периодом. Компенсация осуществляется вычитанием задержанной во времени копии помехи, находящейся в паузе речевого

сигнала, из суммы сигнала и помехи.

Справедливость выражений подтверждена имитационным моделированием. Эффективность компенсации проверена экспериментально. Подавление помехи, соответствующей по форме сигналу «занято», составило 27дБ.

В четвертом разделе рассмотрены возможности адаптивной компенсации широкополосных радиопомех совместно с их отражениями, присутствующих на магистральном участке телефонного канала связи. В каналах с селективно-частотными замираниями применяют разнесенный прием, что позволяет осуществить адаптивную компенсацию помех. Используются адаптивные трансверсальные фильтры, которые устанавливаются в каналах приема. Адаптация осуществляется с помощью процессора Хоуэлса-Эпплбаума.

Получены выражения, определяющие зависимость степени подавления широкополосных радиопомех совместно с их отражениями от погрешности реализации требуемых задержек в адаптивных трансверсальных фильтрах компенсатора помех. Справедливость выражений подтверждена имитационным моделированием. Предложено устройство компенсации помех, защищенное авторским свидетельством на изобретение.

Предложен алгоритм снижения влияния фазоманипулированного сигнала на работу цепей адаптации компенсатора помех. Алгоритм использует обратную связь по решению.

В пятом разделе рассмотрен функциональный состав программных средств, используемых при проектировании САРГК. Рассмотрены разработанные программные средства, предназначенные для исследования помехоустойчивости САРГК. В частности, рассмотрены возможности исследования алгоритма адаптивной компенсации помех. Разработан программный комплекс с возможностью графического программирования различных видов исследования помехоустойчивости САРГК. Рассмотрена работа программы для оперативного редактирования сети акустических моделей.

В системе Matlab разработан набор функций, обеспечивающий исследование помехоустойчивости с использованием предложенного критерия достоверности распознавания команд.

В шестом разделе рассмотрены аппаратные средства, необходимые для исследования помехоустойчивости САРГК. Разработана система сбора информации об ошибках распознавания при организации прямого доступа абонента телефонной сети общего пользования к абоненту учрежденческой АТС.

При разработке средств адаптивной компенсации широкополосных помех, возникает задача определения степени подавления помех. Для решения данной задачи разработана схема двухканального устройства компенсационной обработки сигналов. Разработана схема испытаний компенсатора помех в составе станции связи, которая во время испытаний работает в режиме кольцевой проверки.

Достоверность результатов, полученных в диссертации, обусловлена применением математического аппарата теории матриц, теории вероятностей, математической статистики, теории адаптивной фильтрации, а также проведением моделирования и экспериментов. Научные положения, выводы и рекомендации, сформулированные в диссертации, достаточно полно обоснованы.

Основные положения диссертации отражены в публикациях автора, а также прошли апробацию на научных конференциях разного уровня.

Текст автореферата полностью отражает содержание диссертации, позволяет оценить результаты работы.

Результаты диссертации целесообразно использовать при разработке систем голосового самообслуживания, при построении систем мониторинга служебных телефонных переговоров, систем поиска необходимой информации в архивах звукозаписей телефонных обращений в различные организации, а также при подготовке специалистов в области компьютерной телефонии.

К недостаткам диссертации относится следующее.

1. При анализе возможности подавления квазипериодической помехи с помощью неадаптивной компенсации не рассмотрены возможности адаптивной компенсации.
2. Не проанализировано влияние искажений речевого сигнала, возникающих при сжатии потока данных речи, на достоверность распознавания.
3. Нет подробного анализа зависимостей критерия достоверности распознавания команд от отношения сигнал-шум, полученных в результате моделирования (Рисунок 5.9).

Отмеченные недостатки имеют частный характер и не сказываются существенно на научной и практической значимости полученных в диссертации результатов. Диссертация является законченной научно-квалификационной работой, в которой решена актуальная научная проблема по созданию средств исследования и повышения помехоустойчивости систем автоматического распознавания голосовых команд в телефонии.

Диссертация удовлетворяет требованиям п.п. 9, 10, 11, 13, 14 Положения о присуждении ученых степеней, а ее автор Левин Евгений Калманович заслуживает присуждения ученой степени доктора технических наук по специальности 05.12.13 «Системы, сети и устройства телекоммуникаций».

Отзыв обсужден и одобрен на секции НТС, протокол № 1 от 21.10.2014 г.

Главный инженер ОАО «Владимирское
конструкторское бюро радиосвязи»,
к.т.н.



В.Н. Никонов