

ОТЗЫВ

официального оппонента – заведующего лабораторией Федерального государственного бюджетного учреждения науки «Институт проблем управления им. В.А. Трапезникова РАН» (г.Москва), доктора технических наук, старшего научного сотрудника Фархадова Маиса Паша Оглы о диссертации «Разработка средств исследования и повышения помехоустойчивости систем автоматического распознавания голосовых команд в телефонии» Левина Евгения Калмановича, представленной на соискание ученой степени доктора технических наук по специальности 05.12.13 «Системы, сети и устройства телекоммуникаций».

Актуальность темы. Внедрение систем автоматического распознавания голосовых команд (САРГК) в телефонию сдерживается высокой трудоемкостью их разработки. Трудоемкость, во многом, обусловлена необходимостью обеспечения требуемой устойчивости САРГК к воздействию помех. Воздействие помех приводит к снижению достоверности распознавания команд, которая характеризуется вероятностями пропуска команды и ложного срабатывания системы на помеху.

При исследовании помехоустойчивости САРГК приходится проводить большое количество процедур тестирования системы объемными выборками звукозаписей для контроля достоверности распознавания при отладке алгоритмов подавления помех. Большой объем тестовой выборки звукозаписей обусловлен высокой изменчивостью произнесения команд, разнообразием помех, а также малыми значениями вероятностей ложного срабатывания и пропуска команды.

С расширением сферы использования САРГК увеличивается степень разнообразия помех и растут требования к повышению достоверности распознавания, следовательно, растет и трудоемкость разработки САРГК. Поэтому решаемая в диссертации задача создания средств повышения помехоустойчивости, а также средств исследования помехоустойчивости, снижающих трудоемкость тестирования САРГК, является **актуальной**.

Научная новизна работы характеризуется следующими результатами.

1. Предложен метод сравнения результатов тестирования САРГК, полученных при исследовании влияния различных помех на достоверность распознавания команд. При оценке вероятности ошибки распознавания команды в процессе тестирования САРГК определяются метрика сопоставления текущего произнесения команды со «своей» моделью и наибольшая метрика сопоставления данного произнесения с «чужой» моделью. Определяется разность метрик.

Величина разности является случайной величиной, что обусловлено случайным характером изменчивости произнесения команды. Используя

отношение выборочного среднего к выборочному среднеквадратическому отклонению в качестве критерия достоверности распознавания при сравнении различных вариантов обработки сигналов в САРГК, можно сократить число произнесений этой команды в тестовой выборке звукозаписей по сравнению с известным методом при сохранении неизменной вероятности ошибки сравнения. Известный метод в качестве критерия достоверности распознавания использует относительную частоту пропуска команды.

Получены выражения, с помощью которых определяются требуемые объемы тестовых выборок для предложенного и известного методов. Справедливость полученных выражений подтверждена результатами имитационного моделирования. В частности, результаты моделирования показывают, что, если при подавлении помехи вероятность пропуска команды уменьшилась с 13% до 10%, то объем тестовой выборки сокращается с 580 до 400 звукозаписей команды при сохранении прежней вероятности ошибки сравнения равной 5%. Целесообразность использования метода подтверждена результатами экспериментов.

2. Предложен метод оценки вероятности ложного срабатывания системы на произнесения слов, которые не являются командами. В отличие от известного метода оценка определяется с учетом того, насколько велика степень соответствия произнесения акустической модели команды. Предложенный метод позволяет снизить зависимость оценки от состава тестовой выборки звукозаписей. Повышается степень объективности оценки вероятности ложной тревоги для данной САРГК.

Получены выражения, которые позволяют оценить изменение метрики сопоставления какого-либо произнесения с акустическими моделями помех и команд при изменении степени указанного выше соответствия.

Для экспериментальной проверки метода САРГК была протестирована двумя объемными выборками звукозаписей с разным составом произнесений слов, которые не являлись командами. Доверительные интервалы оценок вероятностей ложной тревоги для одинаковых степеней отличия произнесений от акустических моделей команд в большинстве случаев перекрывались, что свидетельствует о слабой зависимости оценок от состава тестовой выборки звукозаписей (Рисунок 2.2).

3. В результате анализа алгоритма подавления квазипериодической помехи получены выражения, отражающие зависимость степени подавления квазипериодической помехи методом неадаптивной компенсации от соотношения уровней помехи, шума и речевого сигнала. Под квазипериодической помехой понимается сумма шума и периодического компонента, период которого меняется незначительно на интервале длительности команды. Целесообразность использования неадаптивной компенсации обусловлена малой длительностью команды.

На интервале паузы определяется период помехи. Путем сдвига во времени копии помехи, присутствующей в паузе речевого сигнала, на

интервалы, кратные периоду, и последующего вычитания копии из смеси сигнала с помехой осуществляется подавление помехи на интервале длительности команды.

Степень подавления помехи определяется погрешностью определения периода. Получены выражения, которые позволяют определить зависимость погрешности от частоты дискретизации и уровня шума. Справедливость выражений подтверждена имитационным моделированием.

С помощью проанализированного алгоритма была обработана запись с телефонной линии речевого сигнала, пораженного помехой. Величина подавления помехи составила 27,3 дБ.

4. На основе анализа системы адаптивной компенсации широкополосных радиопомех получены выражения, отражающие зависимость степени подавления радиопомех совместно с их отражениями от погрешности реализации требуемых значений задержек в адаптивных трансверсальных фильтрах. Фильтры используются в условиях пространственно-разнесенного приема. Адаптивная компенсация осуществляется на основе алгоритма Хоуэлса-Эпплбаума. Рассмотрена зависимость степени подавления помехи от значения вектора управляющих коэффициентов. Справедливость полученных выражений проверена имитационным моделированием.

5. Проанализирован предложенный алгоритм подавления влияния фазоманипулированного сигнала на работу цепей адаптации. Влияние сигнала подавляется путем использования обратной связи по решению. Построены графики зависимости отношения сигнал-помеха на выходе компенсатора от отношения сигнал-помеха на входе. Увеличение отношения сигнал-помеха обеспечивается и в том случае, когда отношение сигнал-помеха на входе компенсатора больше единицы.

Достоверность и обоснованность полученных научных результатов обеспечивается применением математического аппарата теории вероятностей, теории матриц, теории цифровой обработки сигналов, теории адаптивной компенсации помех, а также использованием имитационного моделирования и проведением экспериментов.

Теоретическая значимость работы заключается в следующем.

Проанализировано изменение метрики сопоставления произнесения с акустическими моделями помех и команд.

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

Проанализировано влияние аддитивного шума на результат оценки периода помехи при компенсации помех с изменяющимся периодом.

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

Исследовано влияние обратной связи по решению на работу адаптивного компенсатора помех.

Практическая ценность диссертации заключается в следующем.

1. Использование предложенного метода сравнения результатов тестирования САРГК, полученных при настройке алгоритмов подавления помех, позволяет уменьшить требуемый объем тестовой выборки звукозаписей, что сокращает объем экспериментальных исследований САРГК. Объем тестовой выборки определяется с помощью полученных выражений исходя из допустимой вероятности ошибки сравнения и ожидаемой вероятности ошибки пропуска команды.

2. Разработана методика выявления помехи и голосовой команды, которым соответствует наименьшая достоверность распознавания. Использование введенного критерия достоверности распознавания команды в методике позволяет снизить объем тестирования САРГК по сравнению с известным подходом.

3. Разработанные методика и соответствующие программные средства оценки вероятности ложного срабатывания САРГК на произнесения слов, не являющихся командами, позволяют снизить объем экспериментов при настройке системы на конкретные условия ее эксплуатации.

4. Полученные выражения позволяют определить минимально возможную частоту дискретизации, а также максимально допустимый уровень шума, при которых обеспечивается необходимый уровень подавления квазипериодической помехи.

5. Разработаны программные средства для исследования компенсации помех, как в абонентской линии, так и в радиоканале.

6. Предложено устройство адаптивной компенсации широкополосных радиопомех, защищенное авторским свидетельством на изобретение.

7. Разработанные программно-аппаратные средства, обеспечивающие сбор данных о появлении ошибок распознавания команд при подключении САРГК к телефонной линии, упрощают настройку системы при вводе ее в эксплуатацию.

8. Разработанные структурные схемы комплекса аппаратуры для испытаний устройств компенсации радиопомех позволяют составить комплект аппаратуры для проведения соответствующих испытаний. Аппаратные решения, связанные с изменением частоты имитатора помех, защищены авторскими свидетельствами на изобретения.

Научные положения, выводы и рекомендации, сформулированные в диссертации, полностью обоснованы. Результаты диссертации достаточно полно отражены в публикациях соискателя и прошли апробацию на научно-технических конференциях различного уровня.

Автореферат диссертации соответствует основным положениям диссертации и отражает практическую и теоретическую значимость

достигнутых результатов работы, позволяет понять суть решаемой научной проблемы.

К числу **недостатков диссертации** можно отнести следующее.

1. В подразделе 3.4. рассмотрена неадаптивная компенсация квазипериодических помех, но нет детального сравнения с возможностями адаптивной компенсации.

2. На стр.184 приведен график зависимости критерия достоверности распознавания команд от отношения сигнал-шум. Форма кривых на графике достаточно сложная, но комментарии по поводу этих графиков очень краткие.

Указанные недостатки не являются определяющими и не оказывают существенного влияния на научную и практическую значимость полученных в диссертации результатов.

Заключение. Диссертация является законченной научно-квалификационной работой, в которой решена актуальная научная проблема по разработке средств исследования и повышения помехоустойчивости систем автоматического распознавания голосовых команд в телефонии, имеющая важное социально-экономическое, хозяйственное значение.

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».

Заведующий лабораторией
автоматизированных систем массового
обслуживания и обработки сигналов
ФГБУН Института проблем управления
им. В.А. Трапезникова РАН (г.Москва),
доктор технических наук,
старший научный сотрудник


Фархадов Маис Паша Оглы

14 октября 2014г.

