

## ОТЗЫВ

официального оппонента - профессора кафедры «Многоканальные телекоммуникационные системы» Московского технического университета связи и информатики, доктора технических наук Шаврина Сергея Сергеевича 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.

В настоящее время в телефонии интенсивно развиваются системы голосового самообслуживания (СГС), важной составной частью которых являются системы автоматического распознавания голосовых команд (САРГК). С развитием СГС расширяется область их использования и увеличивается число пользователей системы, поэтому, с одной стороны, повышаются степень разнообразия и интенсивность помех, воздействующих на САРГК, а с другой - растут требования к достоверности распознавания команд. Усложнение помеховой обстановки и рост требований к достоверности распознавания голосовых команд (ГК) обуславливают растущую трудоемкость разработки САРГК, что сдерживает развитие СГС.

Высокая трудоемкость разработки САРГК во многом обусловлена проведением большого количества экспериментов - процедур тестирования системы объемными выборками звукозаписей ГК, пораженных помехами. По результатам тестирования определяются оценки вероятностей ложного пропуска команды и ложного срабатывания системы при отладке алгоритмов подавления помех. Большой объем тестирования определяется высокой степенью изменчивости произнесения ГК и разнообразием помех.

В диссертации решается актуальная задача создания средств исследования и повышения помехоустойчивости САРГК, обеспечивающих снижение объема тестирования, проводимого при их разработке.

Научная новизна работы характеризуется следующими результатами.

1. На основе введенного критерия достоверности распознавания команд предложен метод сравнения результатов распознавания команд, полученных при изменении степени влияния помех на САРГК.

При оценке вероятности ошибки распознавания некоторой команды в процессе тестирования САРГК определяются метрика сопоставления текущего произнесения команды со «своей» моделью и наибольшая метрика сопоставления данного произнесения с «чужой» моделью. Если разность данных метрик положительна, то результат распознавания считается правильным. В противном случае имеет место ошибка распознавания.

Величина разности является случайной величиной, что обусловлено изменчивостью произнесения команды. Так как на произнесение оказывает влияние множество факторов, то распределение вероятностей случайной величины считается нормальным.

Используя для данного распределения отношение оценок математического ожидания и среднеквадратического отклонения 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

Полученные выражения (2.10), (2.11), (2.13), (2.14) - позволяют определить требуемые объемы тестовой выборки звукозаписей и вероятность ошибки сравнения для известного и предложенного методов в зависимости от ожидаемой вероятности ошибки распознавания. Справедливость полученных выражений подтверждена результатами имитационного моделирования и результатами экспериментов.

Результаты моделирования, в частности, показывают, что, если вероятности распознавания команды сравниваемых результатов тестирования равны 0.1 и 0.13, то при заданной вероятности ошибки сравнения, равной 0.05, использование критерия позволяет сократить объем тестовой выборки звукозаписей голосовой команды с 580 до 400 звукозаписей (на 31%).

2. Предложен метод оценки вероятности ложного срабатывания системы от произнесений посторонних слов, которые являются помехами, возникающими вне интервала длительности голосовой команды. Оценка определяется с учетом того, насколько произнесения соответствуют акустической модели команды. Предложенный метод по сравнению с известным позволяет снизить зависимость оценки от состава тестовой выборки звукозаписей. Повышается степень объективности оценки качества САРГК.

Во втором разделе диссертации выведены соотношения, которые позволяют оценить изменение метрики сопоставления какого-либо произнесения с акустическими моделями помех и команд при изменении степени указанного выше соответствия.

Проведена экспериментальная проверка метода с применением двух объемных тестовых выборок с разным составом звукозаписей. Доверительные интервалы оценок, полученные при тестировании одной и той же САРГК данными выборками для заданного списка команд, в подавляющем большинстве случаев перекрываются (Рисунок 2.2 и таблица ПЗ.1.3).

3. Выведены соотношения, описывающие зависимость степени подавления квазипериодической помехи методом неадаптивной компенсации от соотношения уровней помехи, шума и речевого сигнала. В качестве помехи рассматривается зашумленное периодическое напряжение, период которого медленно меняется на интервале команды. Возможность компенсации обусловлена наличием пауз в речевом сигнале. Так как длительность команды мала и нет возможности провести адаптивную компенсацию, принято решение о применении неадаптивной компенсации.

На интервале паузы определяется период помехи. Сдвигая во времени копию помехи, созданную в паузе, поочередно на интервалы, кратные периоду, можно осуществить подавление помехи на интервале длительности команды. Степень подавления помехи определяется погрешностью определения периода, которая, в свою очередь, зависит от частоты дискретизации и уровня шума. Полученные в третьем разделе выражения позволяют определить указанную зависимость. Справедливость математических выкладок подтверждена имитационным моделированием.

Эффективность неадаптивной компенсации сложной по форме квазипериодической помехи продемонстрирована на примере обработки записи с телефонной линии речевого сигнала, пораженного помехой. Величина подавления помехи составила 27,3 дБ.

4. Выведены соотношения, отписывающие зависимость степени подавления широкополосных радиопомех методом адаптивной компенсации при пространственно-разнесенном приеме от погрешности реализации требуемых значений задержек в адаптивных трансверсальных фильтрах (АТФ). Проанализирована ситуация наличия отражений помехи.

Степень подавления помехи зависит от погрешности реализации требуемых величин задержек в АТФ. Рассмотрены потенциальные возможности компенсации, а также возможности адаптивной компенсации при наличии погрешностей. Проанализирована адаптивная компенсация на основе процессора Хоуэлса-Эплбаума, и рассмотрена зависимость степени подавления помехи от значения вектора управляющих коэффициентов. Справедливость полученных выражений подтверждена результатами имитационного моделирования.

5. Предложен алгоритм адаптивной компенсации радиопомех с подавлением влияния сигнала на процесс адаптации путем использования обратной связи по решению. Проанализировано влияние фазоманипулированного сигнала на работу цепей адаптации. С помощью полученных выражений и аналитического моделирования получены графики зависимости отношения сигнал-шум на выходе компенсатора от аналогичного отношения на входе.

Использование алгоритма позволяет подавить помеху даже в случае, когда отношение сигнал-помеха на входе компенсатора близко к единице. В частности, при отношении сигнал-помеха на входе, равном 4 дБ, отношение сигнал-помеха на выходе равно 8.2 дБ.

Достоверность полученных научных результатов обеспечивается применением математического аппарата теории вероятностей, теории матриц, теории цифровой обработки сигналов, теории скрытых марковских процессов, 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 помехи «своей» и «чужой» команд.

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

Проанализировано влияние аддитивного шума на результат оценки периода помехи при компенсации квазипериодических помех.

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

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 метода сравнения результатов тестирования САРГК, полученных в ходе отладки алгоритмов подавления помех, позволяет уменьшить требуемый объем тестовой выборки звукозаписей, что сокращает трудоемкость тестирования и формирования выборки.

2. Выведенные соотношения позволяют определить необходимый объем тестовой выборки в зависимости от заданной вероятности ошибки сравнения.

2. Разработанная методика выявления по результатам тестирования помехи и голосовой команды, которым соответствует наименьшая достоверность распознавания, 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, позволяет снизить объем экспериментов при настройке системы на конкретные условия ее эксплуатации.

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

5. Разработанный набор функций системы Matlab обеспечивает исследования алгоритмов адаптивной и неадаптивной компенсации помех. Результатом исследования являются значения погрешности задержки, которую требуется установить в трансверсальных фильтрах, а также значения управляющих коэффициентов при адаптивной компенсации помех. В случае неадаптивной компенсации определяется уровень подавления помехи.

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

7. 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

Основные научные результаты диссертации достаточно полно отражены в публикациях соискателя, прошли апробацию на международных, российских и региональных конференциях.

Автореферат диссертации соответствует основным положениям диссертации, достаточно полно отражает практическую и теоретическую значимость достигнутых результатов работы, позволяет понять суть решаемой научной проблемы.

К числу недостатков диссертации можно отнести следующее.

1. Не рассмотрено влияние речевых кодеков на достоверность распознавания голосовых команд.

2. При анализе системы адаптивной компенсации помех не проанализировано ее быстродействие, которое является важной характеристикой системы.

3. В выражение 2.12 (стр.88) входит случайная величина  $\beta$ , распределение которой считается нормальным, однако это допущение подтверждено лишь моделированием, аналитическое доказательство отсутствует.

Указанные недостатки не являются определяющими и не оказывают существенного влияния на научную и практическую значимость 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».

Профессор кафедры «Многоканальные телекоммуникационные системы» Московского технического университета связи и информатики доктор технических наук

С. С. Шаврин

Подпись Шаврина С. С. заверяю  
Ученый секретарь  
Ученого совета МТУСИ  
10.10.14г.



Т.В. Зотова