

ОТЗЫВ

на диссертацию “Разработка средств исследования и повышения помехоустойчивости систем автоматического распознавания голосовых команд в телефони” Левина Евгения Калмановича, представленную на соискание ученой степени доктора технических наук.

Системы голосового самообслуживания (СГС) представляют собой одно из перспективных направлений повышения эффективности функционирования современных телекоммуникаций и сетей телефонной связи как их части за счет внедрения средств вычислительной техники. Факторами, сдерживающими объемы их внедрения, являются помехи и высокая степень изменчивости произнесения голосовых команд (ГК), что сопровождается большим количеством ошибок распознавания. Это значительно усложняет разработку систем автоматического распознавания голосовых команд (САРГК), которые являются неотъемлемой частью СГС.

Трудоёмкость разработки САРГК в значительной степени определяется большим количеством экспериментов по оценке достоверности распознавания команд в условиях воздействия помех. Известные методы оценки достоверности распознавания и помехоустойчивости рассматриваемых систем были созданы, в основном, в интересах проведения научных исследований и не рассчитаны на решение практических задач.

Совокупность этих факторов приводит к тому, что положенная в основу научного исследования диссертационной работы проблема разработки средств исследования и повышения помехоустойчивости систем распознавания голосовых команд, является актуальной.

Диссертация состоит из введения, шести глав, заключения, списка литературы и четырех приложений. Материал, представленный в содержательной части работы, с необходимой степенью полноты рассматривает саму проблему и методы ее решения, а также содержит научную новизну и практическую ценность.

Научная новизна работы заключается в следующем.

1. Предложен метод сравнения результатов тестирования САРГК выборкой звукозаписей, полученных при воздействии различных помех. В основу метода положен подход по выявлению той помехи, которая обуславливает большую вероятность ошибки распознавания произнесенной команды с привлечением в качестве численной меры разности метрик сопоставления произнесений команды со “своей” и с ближайшей “чужой” акустическими моделями. Достоверность метода увеличивается привлечением предложенного автором критерия достоверности распознавания, который определяется на основе оценок математического ожидания и среднеквадратического отклонения выборки значений разности.

Предложенный метод позволяет снизить объем тестовой выборки звукозаписей, что подтверждается результатами экспериментов. Получены со-

ответствующие математические выражения, справедливость которых проверена имитационным моделированием.

2. Предложен метод оценки вероятности ложного срабатывания системы на посторонние слова, основанный на отражении последней цифровой кодовой комбинацией и учитывающий близость произнесения акустической модели команды. По сравнению с известным методом значительно снижается зависимость оценки от состава тестовой выборки звукозаписей и повышается степень объективности оценки качества собственно системы.

Корректность предложенного метода подтверждена тестированием САРГК двумя различными выборками звукозаписей. Доверительные интервалы оценок для одинаковых кодовых комбинаций, как правило, перекрываются, что свидетельствует об инвариантности оценки вероятности ложной тревоги к составу тестовой выборки звукозаписей.

3. Получены выражения, отражающие возможность подавления квазипериодических помех методом неадаптивной компенсации. Квазипериодической считалась помеха в виде зашумленного напряжения, период которого медленно меняется во времени.

Компенсация осуществляется путем поочередного сдвига во времени и последующего вычитания фрагмента помехи, расположенного в паузе речевого сигнала, из суммы сигнала и помехи. Величина сдвига кратна периоду помехи. Степень подавления помехи зависит от погрешности оценки периода в паузе. Полученные выражения описывают зависимость степени подавления помехи от уровней шума и речевого сигнала, а также от частоты дискретизации сигналов.

Справедливость полученных выражений подтверждена имитационным моделированием и результатами обработки записи смеси речевого телефонного сигнала и помехи. Подавление помехи отличается высокой эффективностью и достигает 27,3 дБ.

4. Получены аналитические выражения, отражающие

а) зависимость степени компенсации широкополосных радиопомех при пространственно-разнесенном приеме от погрешности реализации требуемых задержек в трансверсальных адаптивных фильтрах.

б) степень подавления помех от выбора вектора управляющих коэффициентов, используемого при адаптивной компенсации по алгоритму Хоуэлса-Эплбаума.

Справедливость полученных выражений подтверждена результатами имитационного моделирования.

5. Предложен алгоритм адаптивной компенсации радиопомех, обеспечивающий подавление влияния сигнала на работу цепей адаптации. Получены выражения, позволяющие оценить степень подавления помех в зависимости от отношения сигнал-помеха на входе компенсатора. Методом аналитического моделирования определена зависимость отношения сигнал-помеха на выходе компенсатора от аналогичного отношения на его входе.

Теоретическая значимость работы характеризуется следующими результатами.

1. Проанализирована зависимость метрики сопоставления произнесений слов с акустическими моделями команд и помех от степени соответствия произнесений модели.

2. Введен критерий достоверности распознавания команды, позволяющий снизить вероятность ошибки сравнения результатов тестирования САРГК.

3. Доказана возможность подавления квазипериодических помех, воздействующих на речевой сигнал, путем использования неадаптивной компенсации.

4. Проанализировано влияние погрешности установки величин задержки сигнала в адаптивных трансверсальных фильтрах на степень подавления радиопомех при наличии их отражений.

5. Продемонстрирована возможность уменьшения влияния сигнала на работу цепей адаптации в компенсаторе помех за счет использования обратной связи по решению.

Практическая ценность полученных в диссертации результатов заключается в следующем.

1. Предложенный метод позволяет значительно сократить объем тестирования САРГК при настройке ее на конкретные условия эксплуатации. Разработаны программные средства, позволяющие оценить вероятность ложной тревоги в зависимости от кода помехи.

2. На основе предложенного метода сравнения результатов тестирования САРГК разработана методика выявления критически важного для безошибочной работы системы сочетания помехи и команды. Методика позволяет снизить объем тестовых выборок звукозаписей по сравнению с известным методом.

3. Предложен алгоритм подавления квазипериодической помехи. Полученные выражения обеспечивают расчет степени подавления помехи и выбор частоты дискретизации сигналов.

4. Разработаны программные средства, позволяющие определять основные параметры устройств компенсации помех путем имитационного моделирования.

5. Разработаны программно-аппаратные средства, обеспечивающие сбор информации об ошибках распознавания при подключении системы распознавания команд к телефонной линии.

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

7. Разработан аппаратный комплекс для испытаний компенсаторов радиопомех.

Недостатками диссертации считаю следующее.

1. Достаточно общая формулировка некоторых результатов работы, которые характеризуют научную новизну. Например, говорится о предложен-

ном критерии достоверности распознавания, но не указывается, в чем состоит его новизна; говорится о методе сравнения результатов распознавания команд, полученных в разных сеансах тестирования, но не указывается, чем отличаются данные сеансы (С.9).

2. В разделе 2.1 справедливость полученных выражений метрик сопоставления произнесений с акустическими моделями не проверялась моделированием. Представлены лишь результаты экспериментов.

3. Утверждение, что закон распределения разности метрик $\Delta M_{t-m,ii}$ близок нормальному закону, не доказано аналитически (С.61).

В работе использованы современные методы исследований, основанные на применении математического аппарата теории матриц, теории вероятностей, математической статистики, теории статистической радиотехники, а также проведением моделирования и экспериментов.

Научные положения, выводы и рекомендации, сформулированные в диссертации, обоснованы с необходимой степенью полноты.

Основные положения диссертации с необходимой степенью полноты отражены в публикациях автора, в т.ч. в ведущих периодических изданиях, включенных в перечень ВАК, а также апробированы на научных конференциях разного уровня. Текст работы написан простым и понятным языком.

Текст автореферата полностью отражает содержание диссертации, позволяет понять логику рассуждений автора и оценить результаты работы.

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, а полученные в ней результаты обладают необходимой степенью научной новизны и практической ценностью.

Диссертация полностью 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”.

Официальный оппонент,

Директор по развитию ООО
“РдМ Дистрибьюшн”, доктор
технических наук

А.Б. Семенов

15.10.2014

Подпись А.Б. Семенова заверяю:

Генеральный директор ООО
«РдМ Дистрибьюшн»

Р.А. Чистяков