



Утверждаю

Ректор Технологический  
университет Таджикистана  
доктор технических наук, доцент  
Амонзода И. Т.

«06» 02 2023г.

### Отзыв

ведущей организации на диссертационную работу **Ашурзода Бахром Хайридин** на тему «Методы и модели поиска ключевых слов в речи на таджикском языке (спектральный анализ – особенности)», представленную на соискание ученой степени кандидата технических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей

Диссертация и автореферат Ашурзода Б.Х. поступили на кафедру «Программирование и компьютерная инженерия» Технологического университета Таджикистана от диссертационного совета 6D.КОА-049 по защите кандидатских диссертаций при Таджикском техническом университете имени академика М.С. Осими.

Решением указанного совета кафедра «Программирование и компьютерная инженерия» Технологического университета Таджикистана утверждена в качестве ведущей организации по защите кандидатской диссертации соискателя Ашурзода Б.Х.

Основные сообщения с экспертизой материалов исследования представлены кандидатом технических наук, и.о. доцента, заведующим кафедрой «Программирование и компьютерная инженерия», Технологического университета Таджикистана Гуломсафдаровым Абдулназаром Гуломназаровичем и кандидатом физико-математических наук, и.о. доцента Зариповым Саидахмадом Асроровичем.

**Актуальность темы.** Настоящее исследование посвящено анализу и разработке моделей и методов распознавания речи на таджикском языке. Актуальность исследования заключается в необходимости разработки новой модели распознавания речи, направленной на практическую точность систем распознавания слитной речи.

Вопросы взаимодействия человека и машины являются одним из самых важных при создании новых компьютеров. Многие из наиболее эффективных средств взаимодействия человека с машиной были бы для него естественными: через визуальные образы и речь.

**Личный вклад автора** заключается в определении основных задач исследования и основного направления исследований под руководством научного руководителя. Алгоритмы и программные продукты, а также

окончательные результаты исследования были получены автором самостоятельно.

**Объем и структура диссертации.** Диссертационная работа состоит из введения, 4 глав, заключения, списка использованной литературы и приложений. Основной текст размещен на 132 страницах, включает 18 таблиц, 32 рисунков. Список литературы, включает 124 наименований.

Во **введении** обоснована актуальность выбранной темы, определены цель, задачи, объект и предмет исследования, раскрыты новизна, теоретическая и практическая значимость работы, приведены положения, вынесенные на защиту, описаны методологическая база и материал исследования, изложены сведения об апробации работы и ее структуре.

В **главе 1** рассматриваются модели представления речевых сигналов, анализируются существующие методы поиска ключевых слов в речи и программные системы поиска ключевых слов в слитной речи.

Граница между данными задачами является условной: распознать слитную речь – это и значит выделить какие-то сегменты, при совокупности и сложении которых на выходе система распознает речь как одно целое. Все системы, в первую очередь, подразумевают поиск ключевых слов или лексических элементов в слитной или раздельной речи. Поиск зависит от различных факторов и условий, и должен определяться несколькими составляющими для качественного поиска ключевого слова.

Во **главе 2** рассматриваются особенности обработки речевого сигнала на таджикском языке, анализируются способы разработки структуры и состава системы поиска ключевых слов, а также представлена модель представления речевого сигнала.

Модель распознавателя речи представлен в виде блок-схемы, где реализуется основная идея построения синтезатора речи по тексту. В нем использованы обозначения:  $p_s$  - пауза между слогами при произношении слова;  $p_w$  - пауза между словами при чтении предложения (соответствует пробелу между словами);  $p_i$  - пауза, отмечающая внутренний знак препинания (запятая, двоеточие, точка с запятой, тире);  $p_e$  - пауза, отмечающая внешний знак препинания (точка, многоточие, знак вопросительный, знак восклицательный);  $p_a$  - пауза, отмечающая конец абзаца (в диссертации представлена блок-схема).

**Глава 3** посвящена для описания процессов разработки алгоритмов и методов их реализации. В ней автор описывает разработку расчета параметров вероятностей графической модели на примере, скрытой марковской модели и условных случайных полей.

Вместе с разработкой модели следует разработать также алгоритм обучения модели. Обучение гибридной модели заключается в оценке параметров как марковской цепи, так и весов нейронной сети. Пока не существует алгоритма, который бы позволил одновременно оценить оба множества параметров и для СММ, и для нейронной сети. Кроме того, поскольку для нейронной сети

используется обучение «с учителем», то требуется значительный объем акустических данных, размеченных вручную, который в настоящее время отсутствует.

В главе 4 представлена архитектура программного комплекса с учетом современных стандартов.

В параграфе 4.1. описываются следующие блоки программного обеспечения такие как: класс «TMatrix», класс «TWaveFile», класс «TWaveHeader», класс «TAnalyzer», класс «TDescriptor», класс «TRecognizer», класс «TRecognizerResult», класс «TMFCC» в виде отдельных диаграмм.

В параграфе 4.2 осуществляется проектирование вычислительных модулей программного комплекса, куда входят: структура WAVE файла (структура анализируемых сигналов), деление анализируемого сигнала на фреймы, mel-кепстральные коэффициенты (MFCC), разложение в ряд Фурье, расчёт mel-фильтров, применение фильтров, логарифмирование энергии спектра, косинус-ное преобразование и сравнение сигналов, алгоритм динамической трансформации временной шкалы.

Экспериментальное исследование программной системы описывается в параграфе 4.3. В ходе исследования были описаны и пройдены следующие этапы: принцип работы системы при обучении и заполнении словаря, запись варианта произношения слова, обработка звукового сигнала в вектор, нормализация вектора, удаление тишины в начале и в конце вектора (сигнала), обработка дескриптора и нормализация вектора, поиск эквивалентов вычисленных дескрипторов в словаре и база данных.

В параграфе 4.4 представлены основные выводы четвертой главы.

**Научная новизна работы.** В ходе исследования предлагается новый подход к созданию акустической модели ключевых слов с использованием акустических моделей фонем, отличающихся от известных моделей, в том числе и в языковом направлении. Впервые решена задача качественного и точного распознавания таджикских слов на основании сравнительного фонемного анализа:

- реализован новый метод распознавания ключевых слов на таджикском языке;
- реализован новый метод представления ключевых слов с применением скрытой Марковской модели и случайного поля;
- проведён сравнительный анализ результатов работы предлагаемых методов на коллекции из 20 дикторов 300 слов, подтверждающий их эффективность;
- создан комплекс алгоритмов и программ для обработки базы данных большого объема, реализующий описанные в данной работе методы.

**Оценка содержания диссертации.** Диссертация написана четким и ясным языком с большим количеством графического материала, поясняющего и

иллюстрирующего соответствующие результаты научных положений и технических решений. По содержанию работы можно сделать следующие замечания:

1. В тексте диссертации и автореферата замечены некоторые грамматические и стилистические ошибки, необходимо обратить внимание на строгую идентичность приведенных цитат, ссылок и сносок с опубликованными вариантами официальных и научных документов.

2. Не соблюден алфавитный порядок аббревиатур и сокращений, обратить внимание на фонетические особенности таджикских слов в части произносимых и непроезносимых букв в словах.

3. В полной мере раскрыть сущность применения скрытых марковских методов в исследовании при решении данной задачи.

4. В главе 4 под названием «Разработка программ поиска ключевых слов в речи» в страницах 88-97 рекомендуется убрать методический материал, не уместный содержанию главы. Здесь целесообразно размещение именно научных результатов диссертанта.

5. В части научной новизны (пункт 3) автореферата говорится о сравнительном анализе 300 слов, озвученных 20 дикторами, но содержание и результаты данного анализа не приводятся.

6. В описании алгоритмов, созданных программ и модулей обнаружены несоответствия международным стандартам составления и описания блок-схем и других компонентов.

7. В тексте автореферата и диссертации упоминается о 39 параметрах идентификации голосовой информации, но отсутствует их список и описание каждого параметра.

8. Созданная соискателем электронная база данных (словарь) содержит 300 слов, озвученная 4 вариантами, что не подтверждает её репрезентативность. Это компенсируется возможностью автоматического восполнения БД новыми озвученными словами.

9. Каждый полученный и представляемый к защите результат должен быть закреплен ссылкой на конкретный научный источник.

10. В работе не отражены научные результаты таджикских ученых, таких как Усманов З.Д., Исмаилов М.А., Зарипов С.А. и др. по грамматической (морфологической, фонетической, синтаксической) особенности таджикского языка.

11. Вышеуказанные замечания не умаляют научно-практическую ценность работы.

**Анализ диссертационной работы в целом позволяет сделать следующие выводы:**

1. Представленная Ашурзода Б.Х. кандидатская диссертация отвечает паспорту специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей» (п.2, 4 и 9).

2. Представленная диссертационная работа Ашурзода Б.Х. является самостоятельной, законченной научной квалификационной работой, обладающей признаками актуальности, новизны, внутреннего единства, научной и практической значимости.

3. Диссертационная работа Ашурзода Бахром Хайриддин на тему «Методы и модели поиска ключевых слов в речи на таджикском языке (спектральный анализ – особенности)» соответствует требованиям ВАК при Президенте Республики

Таджикистан, предъявляемым к диссертациям на соискание ученой степени кандидата технических наук, а ее автор заслуживает присуждения ему ученой степени кандидата технических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Отзыв обсуждён и утверждён на расширенном заседании научного семинара кафедры «Программирование и компьютерная инженерия» Технологического университета Таджикистана от 02.02.2023, протокол №9.

Председатель:

заведующий кафедрой «Программирование и компьютерная инженерия»,

кандидат технических наук, и.о. доцента

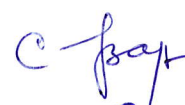


Гуломсафдаров А.Г.

Эксперт:

Зарипов Саидахмад Асрорович, кандидат

физико -математический наук, и.о.доцента



Зарипов С.А.

Уч. секретарь:

доктор Phd, и.о. доцента



Хамидова Д.Н.

Заверяю:

Начальник отдела кадров

и спец. работ ТУТ



Н.А.Бухориев