



Санкт-Петербургский государственный университет

Кафедра системного программирования

# Решение противоречия в дипломном проекте (голосовой помощник врача «СТОММИС») с применением ТРИЗ

Альшаеб Басель, группа 24.М71-мм

Курс ТРИЗ

Преподаватель: к.ф.-м.н. Старший преподаватель С.С. Сысоев

Санкт-Петербург  
2025

## Описание задачи

- В существующей медицинской информационной системе **СТОММИС** врачи взаимодействуют с интерфейсом с помощью традиционных средств ввода — клавиатуры, мыши или сенсорного экрана. Эти методы могут быть неэффективны во время медицинских процедур или в ситуациях, где требуется быстрое и «безрукое» управление.
- **Цель данной работы:** улучшить систему, внедрив интерфейс голосового управления, который позволит врачам выполнять ключевые действия при помощи голосовых команд. Такое решение повышает удобство, снижает нагрузку на пользователя и обеспечивает более быструю и безопасную работу с системой в клинических условиях.

## Описание технического противоречия

Основная сложность заключается в проектировании подсистемы распознавания речи:

- Непрерывное прослушивание обеспечивает высокую точность, но требует значительных вычислительных ресурсов.
- Ограниченнное прослушивание экономит ресурсы, но может привести к потере важных голосовых команд.
- Если слушаем *всё* аудио, то растут расходы (время, вычисления, сеть).
- Если слушаем *не всё*, то рискуем пропустить команду врача.
- Хотим: **минимальные ресурсы и высокая точность распознавания**.

**Техническое противоречие:** Улучшение надёжности распознавания  $\Rightarrow$  рост энергопотребления/сложности.

# Параметры ТРИЗ

Улучшаемый параметр

**Надёжность распознавания (#27)**

Ухудшаемые параметры

- Энергопотребление (#21)
- Сложность устройства/системы (#36)

21: Энергопотр.

27: Надёжность

36: Сложность

# Матрица противоречий (Часть 1)

27 (Надёжность) vs 21 (Энергопотребление)

Рекомендуемые принципы: **21, 11, 26, 31**

- **21. Принцип проскока (Skipping):** обработка наиболее важных сегментов аудио и пропуск неинформативных аудиосигналов.
- **11. Принцип "заранее подложенной подушки" (Cushion in advance):** предусмотреть буфер, чтобы не потерять начало команды.
- **26. Принцип копирования (Copying):** использовать упрощённую «копию» основной модели для предварительного анализа.
- **31. Принцип применения пористых материалов (Porous materials)**

# Матрица противоречий (Часть 2)

27 (Надёжность) vs 36 (Сложность)

Рекомендуемые принципы: **13, 35, 1**

- **13. Принцип "наоборот" (The other way round)**
- **35. Принцип изменения параметров (Parameter change):** динамически менять частоту дискретизации.
- **1. Принцип дробления (Segmentation):** разделить систему на фазы: Idle → Trigger → Recognition.

# Выбранные изобретательские принципы и идея решения

- ① 26 – **Принцип копирования:** простая модель для предварительного распознавания.
- ② 21 – **Принцип проскока:** игнорирование неинформационных аудиосигналов и фоновых шумов.
- ③ 11 – **Принцип "заранее подложенной подушки":** предусмотреть буфер, чтобы не потерять начало команды.
- ④ 35 – **Принцип изменения параметров:** адаптивная смена настроек (частота дискретизации).
- ⑤ 1 – **Принцип дробления:** разделить систему на фазы.

**Изобретательская идея:** создать адаптивную систему распознавания речи, в которой простая модель предварительно фильтрует аудио и активирует полную модель нейросети только при необходимости.

# Карта соответствия: принципы ТРИЗ → архитектура голосового помощника

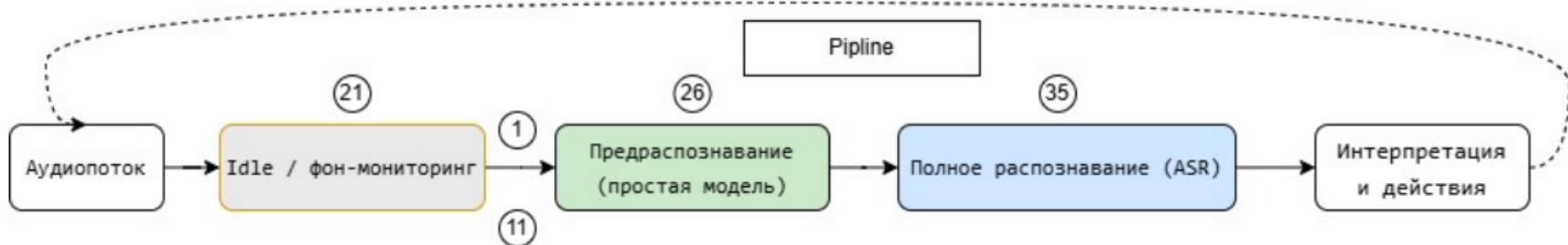
1) Дробление:  
Idle -> Trigger -> ASR

21) Проскок:  
игнорирование фона/тишины

26) Копирование:  
простая модель для предраспознавания

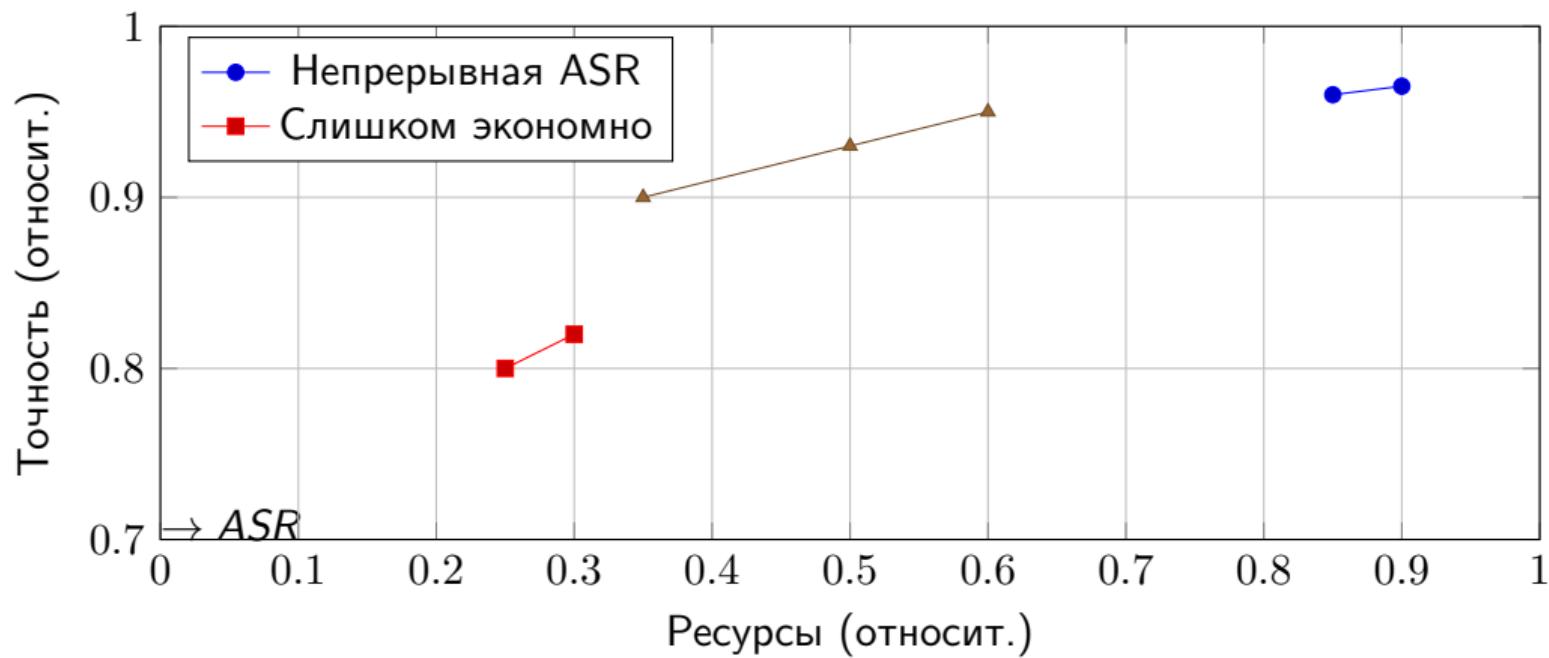
11) "Заранее подложенной подушки":  
наличие буфера

35) Изменение параметров:  
частота дискретизации



Числа в кружках — номера изобретательских принципов ТРИЗ: 1, 11, 21, 26, 35.

## Компромисс «ресурсы–точность»: эффект адаптивности



Адаптивная схема достигает  $\approx 60\text{--}70\%$  экономии ресурсов при точности 90–95%.

# Выводы

- ① Противоречие: точность ↑ vs ресурсы/сложность ↑.
- ② Выбранные принципы ТРИЗ: **26, 21, 11, 35, 1.**
- ③ Изобретательская идея: VAD/KWS → ASR.