

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ИМЕНИ Н.Э. БАУМАНА  
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)

# Метод автоматического определения ритмического рисунка и темпа цифровой музыкальной записи на основе байесовского иерархического моделирования

Студент: Петрова Анна Алексеевна

Группа: ИУ7-86Б

Научный руководитель: Кивва Кирилл Андреевич

**Цель:** реализовать метод автоматического определения переменных темпа и ритма музыки на основе байесовского иерархического моделирования.

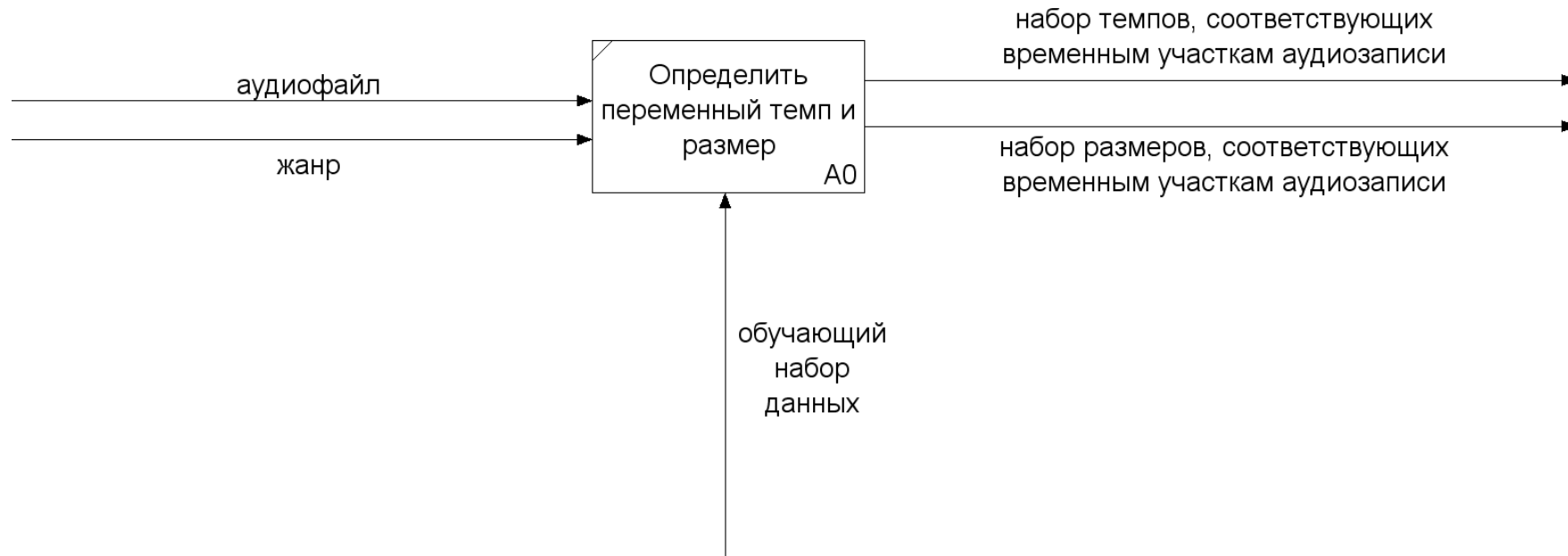
**Задачи:**

- провести анализ предметной области и сформулировать проблему;
- проанализировать и сравнить основные методы определения темпа и ритма;
- разработать метод решения поставленной задачи;
- спроектировать архитектуру разрабатываемого программного обеспечения;
- реализовать разработанный метод;
- протестировать и сравнить результаты работы реализованного метода с результатами, полученными с помощью известных аналогов.

# Постановка задачи

Ограничения:

- входной аудиофайл в формате mp3;
- знаменатели размеров равны 4.



# Методы определения ритма и темпа

Метод	Точность результатов	Переменный темп и ритм	Формат входного аудиофайла	Размер обучающего датасета*
ДВП	~ 65 % (13 верных из 20)	Не определяются	Нет ограничений	Обучение не нужно
Скрытые марковские модели	~ 80 %	Могут определяться при модификации метода	MIDI	88
На основе БИМ**	~ 82%	Не определяются	Нет ограничений	100
Сверточная нейросеть	До 92%	Не определяются	Нет ограничений	8596

*\*На основе данных из исследований*

*\*\*БИМ – байесовское иерархическое моделирование*

# Байесовское моделирование

$$P(Q|X) = \frac{P(X|Q)P(Q)}{P(X)}, \quad (1)$$

$X$  – собранные данные с известным определяемым параметром,

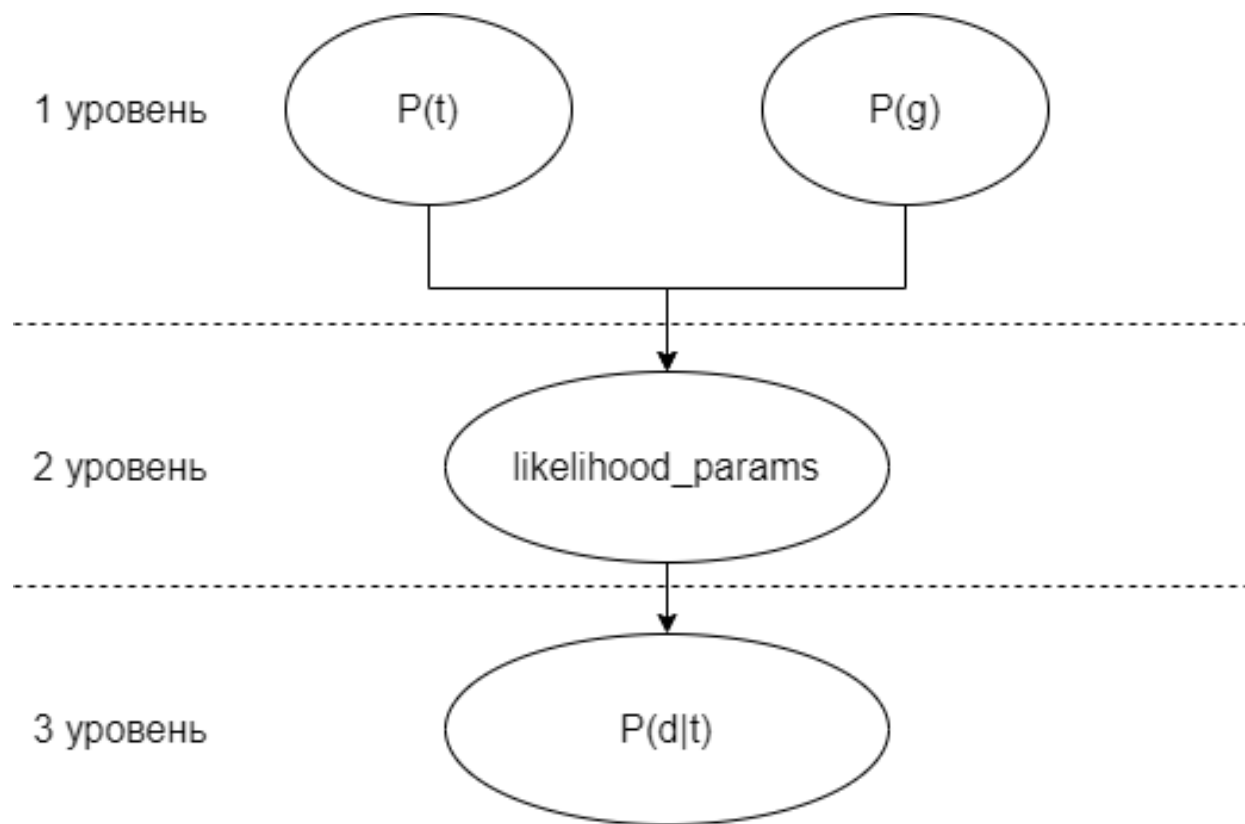
$Q$  – искомый параметр,

$P(Q)$  – априорная вероятность с заданным распределением,

$P(X|Q)$  – функция правдоподобия ( $X$  фиксированное),

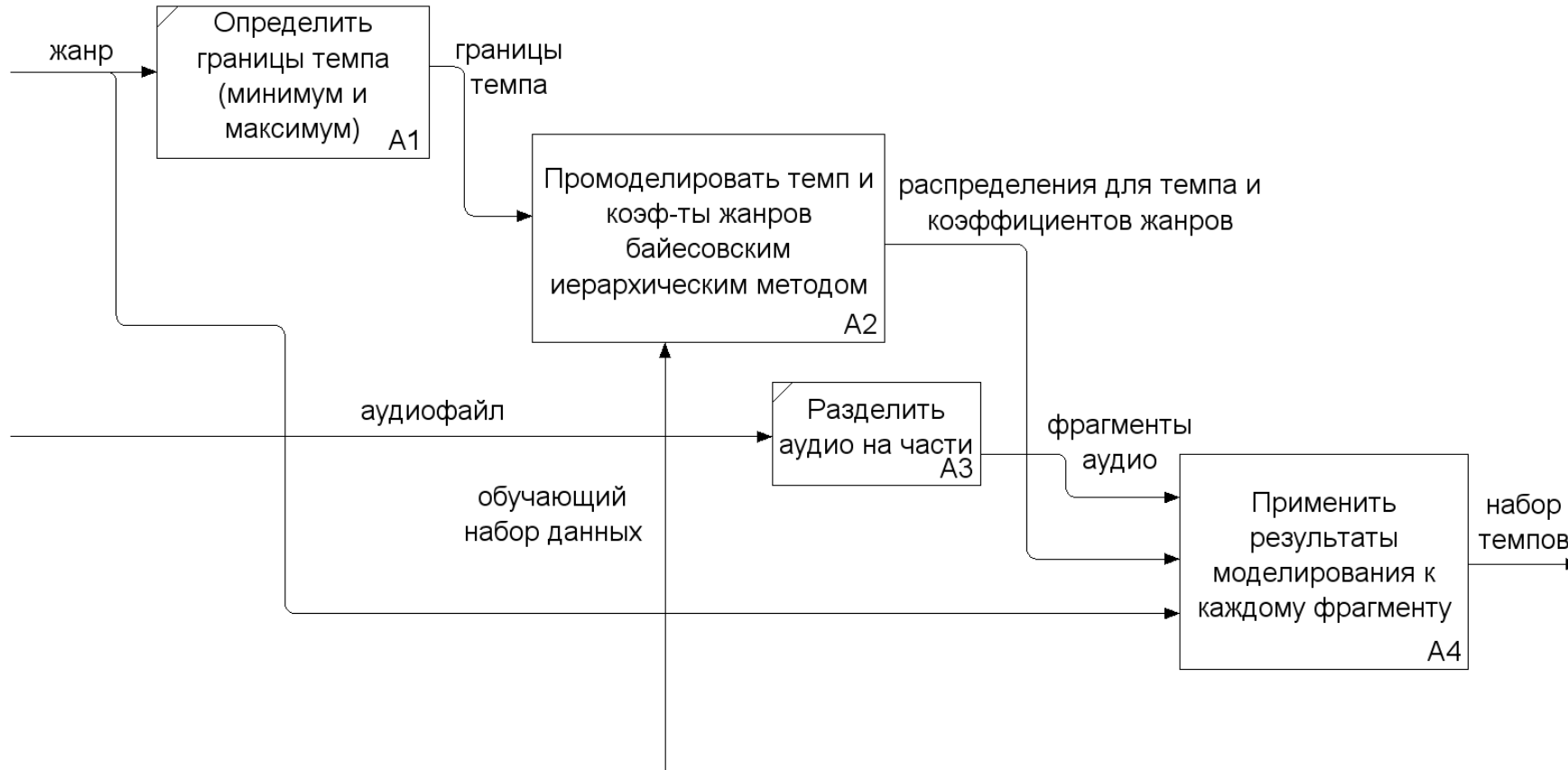
$P(X)$  – нормирующее число, гарантирующее, что  $P(Q|X) \leq 1$ .

# Иерархический подход

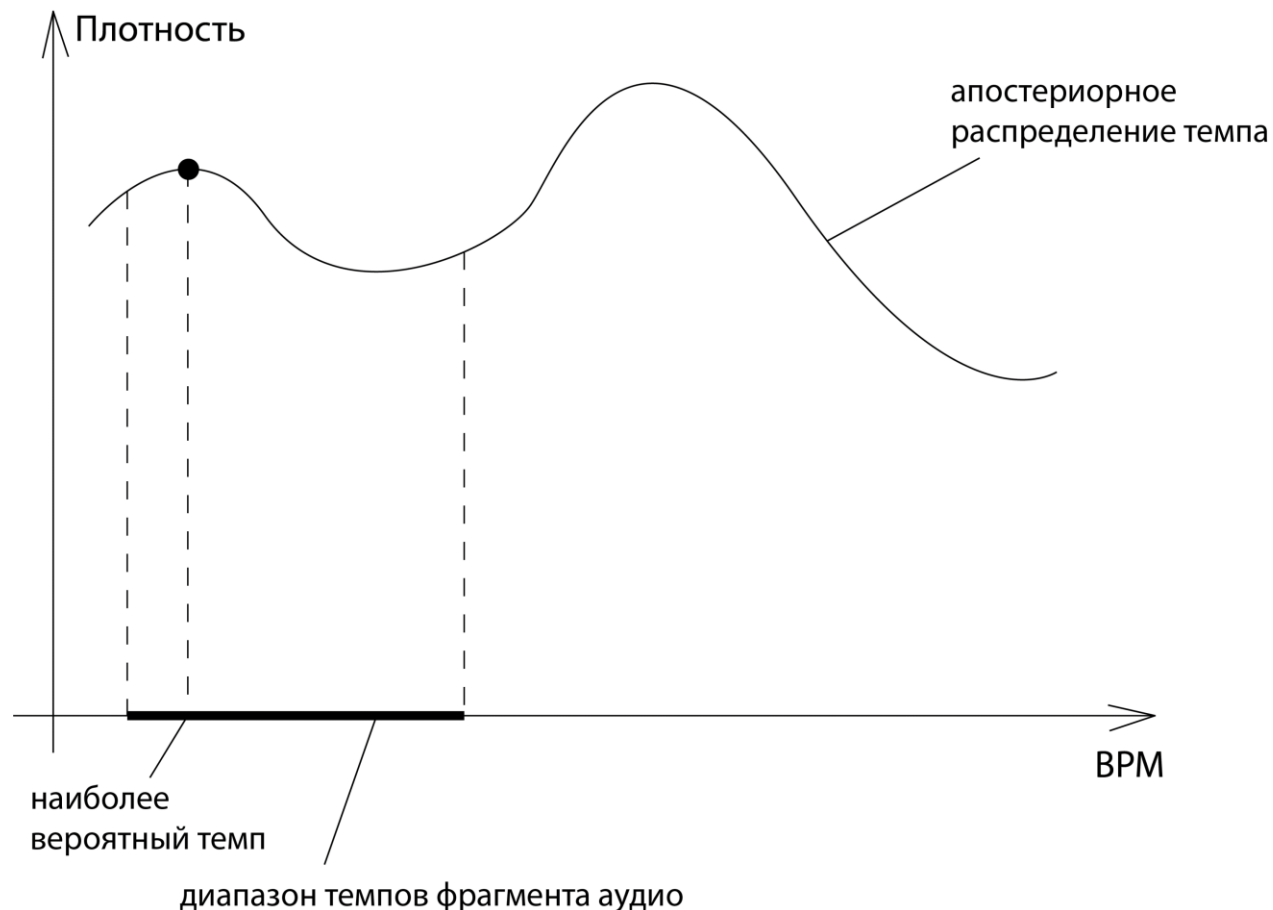


$t$  – искомый темп,  
 $g$  – жанр анализируемой музыки,  
 $P(t)$ ,  $P(g)$  – априорные  
распределения темпа и жанра,  
 $\text{likelihood\_params}$  – параметры  
распределения функции  
правдоподобия,  
 $P(d|t)$  – функция правдоподобия.

# Алгоритм оценки переменного темпа



# Применение результатов моделирования к аудио фрагменту



Корректировка темпа  
коэффициентом жанра:

$$t = tempo + c * \sigma, \quad (2)$$

$t$  – итоговый темп,

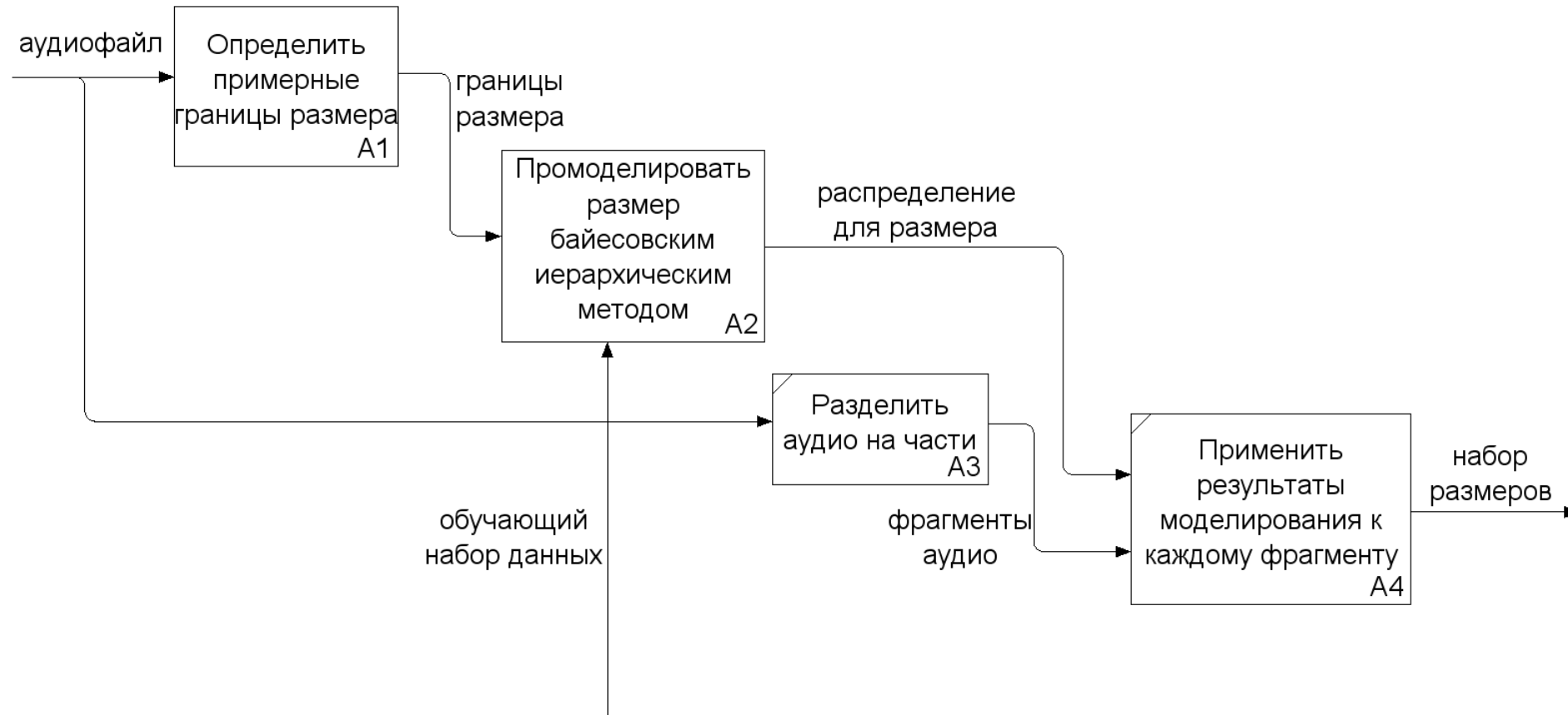
$tempo$  – темп, найденный  
применением результатов  
моделирования к фрагменту,

$c$  – коэффициент жанра,

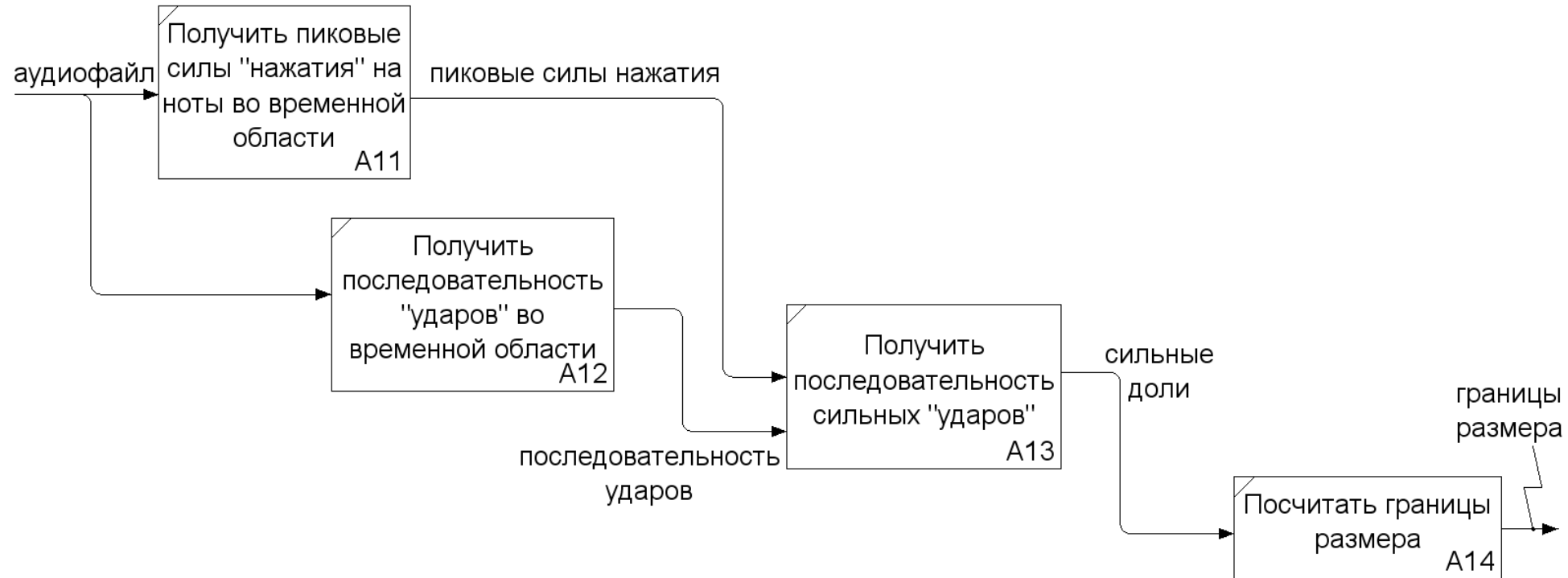
$\sigma$  – дисперсия априорного  
распределения темпа.



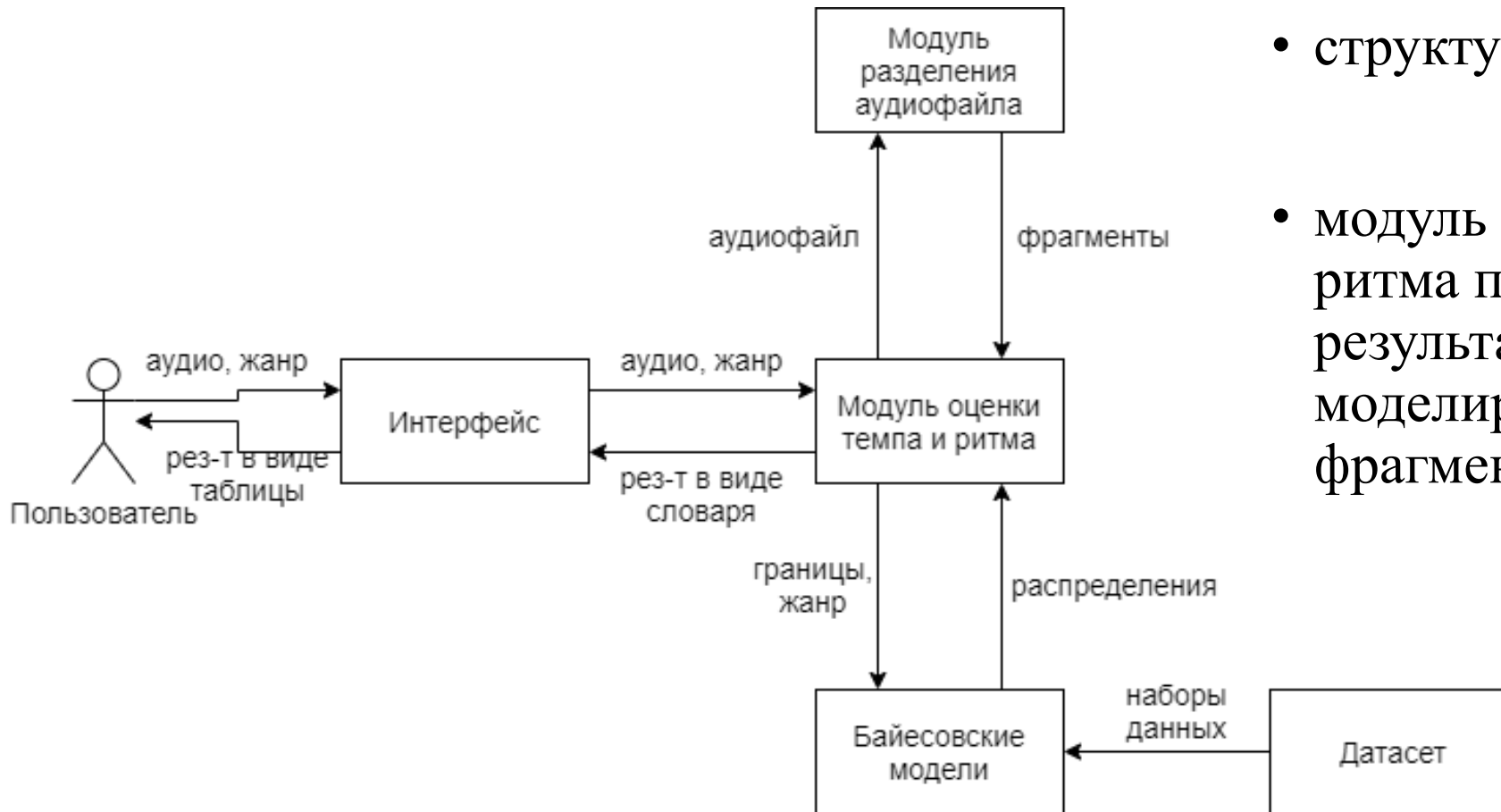
# Алгоритм оценки переменного ритма



# Определение границ размера



# Структура программы

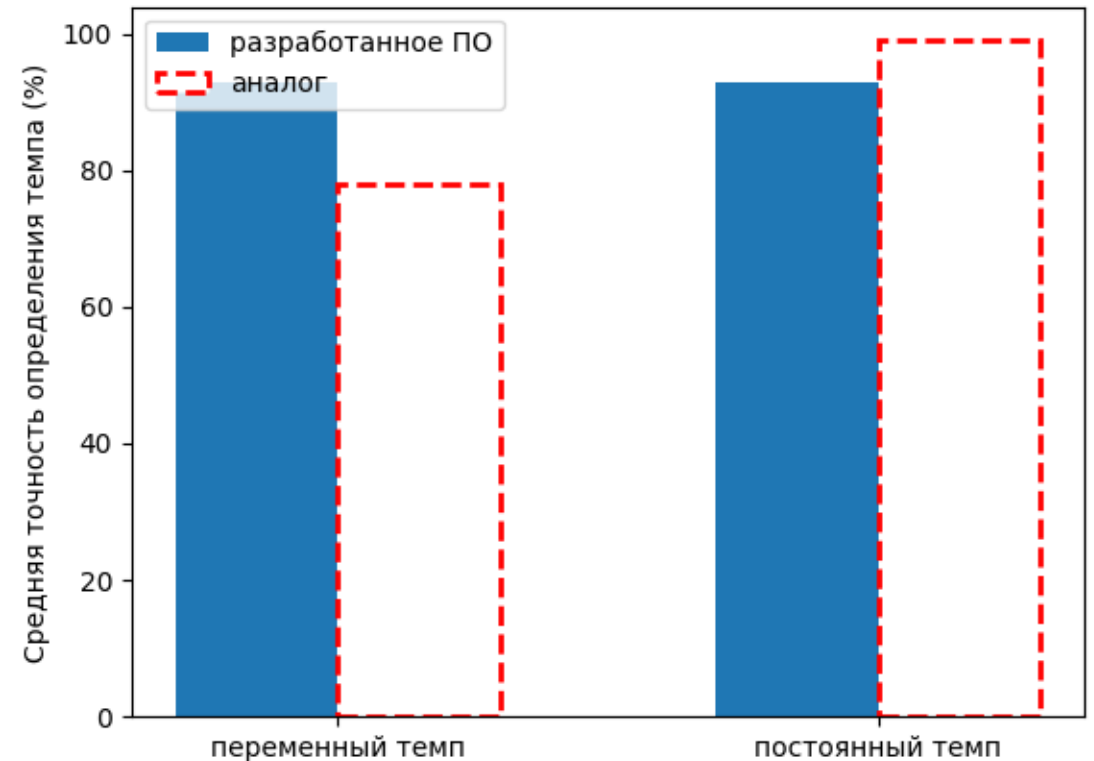


- структурный подход;
- модуль оценки темпа и ритма применяет результаты байесовского моделирования к аудио фрагментам.

# Сравнение результатов работы с аналогом

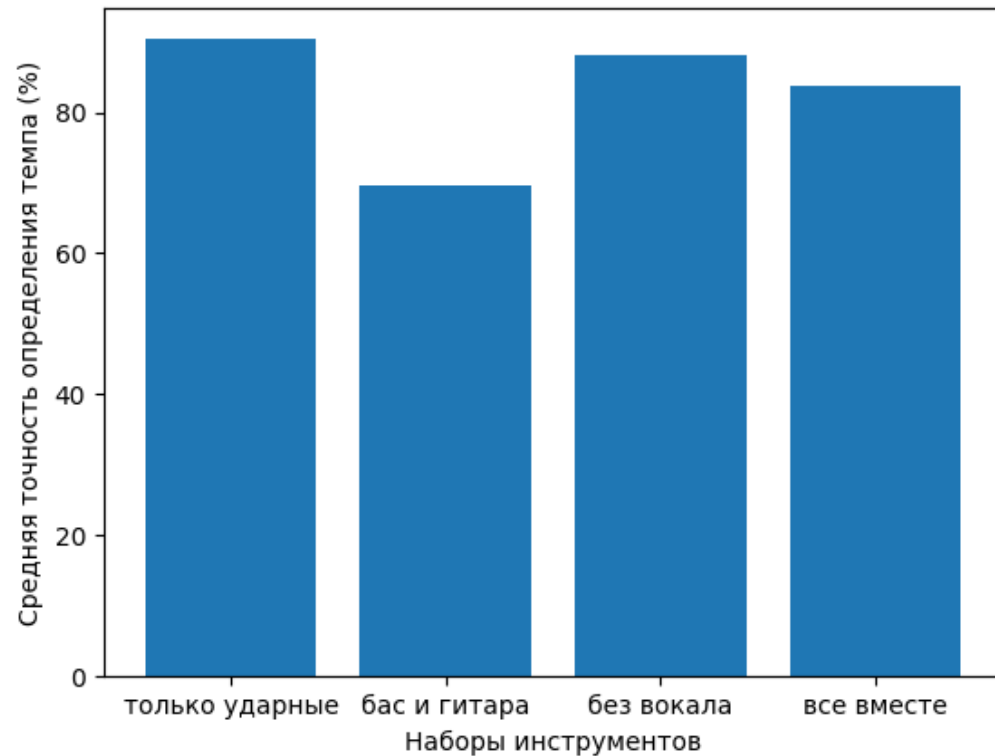
	Ошибка метода (%)	Ошибка аналога* (%)
Постоянный темп	7	1
Переменный темп	7	22

*\*Метод `beat_track()` из библиотеки `librosa`*

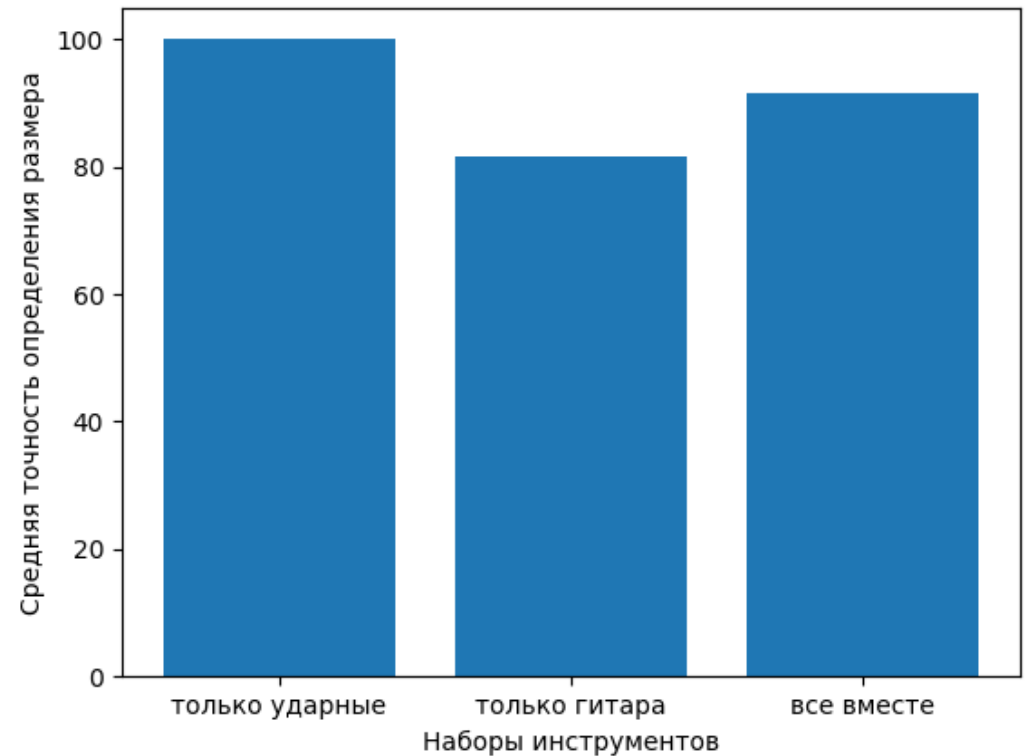


# Применимость метода для разных наборов инструментов

Точность определения темпа (%):

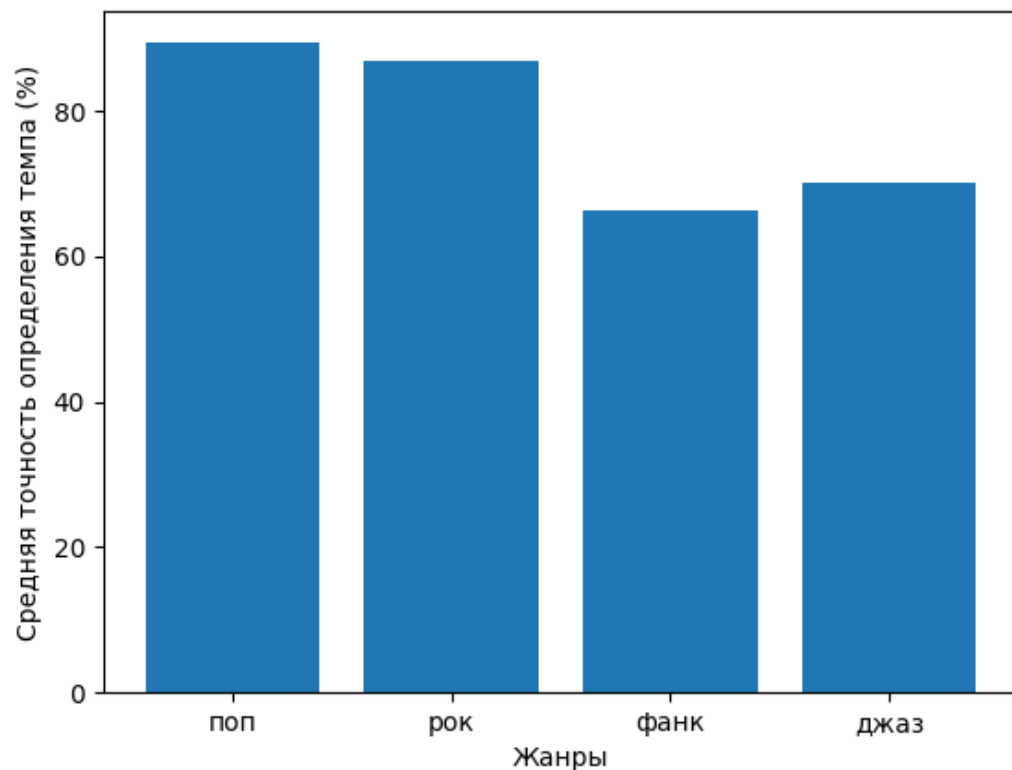


Точность определения ритма (%):

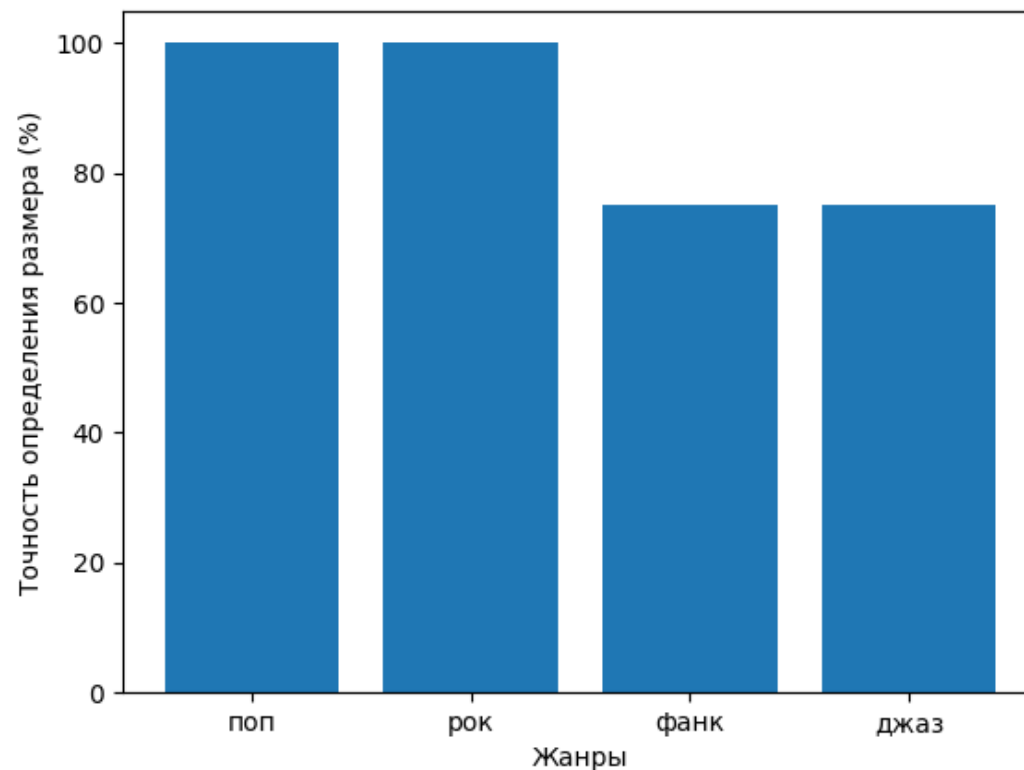


# Применимость метода для разных жанров

Точность определения темпа (%):



Точность определения ритма (%):



# Заключение

Был разработан метод автоматического определения переменных темпа и ритма музыки на основе байесовского иерархического моделирования.

Были решены следующие задачи:

- проведен анализ предметной области и сформулирована проблема;
- проанализированы основные методы определения темпа и ритма и проведено их сравнение;
- разработан метод решения поставленной задачи;
- спроектирована архитектура разрабатываемого программного обеспечения;
- реализован разработанный метод;
- проведено сравнение результатов работы реализованного метода с результатами, полученными с помощью аналога.

# Дальнейшее развитие

- Улучшение оценки постоянного темпа
- Увеличение точности определения темпа и ритма при работе с гитарной музыкой
- Повышение точности результатов при работе с более сложными и нестандартными жанрами (фанк, джаз, классика)
- Определение размеров с другими знаменателями
- Добавление работы с иными форматами аудиофайлов