



**Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)**

**Выпускная квалификационная работа бакалавра на тему:
«Метод анализа активности пользователей системы
автоматизированного проектирования с использованием
поиска последовательных шаблонов.»**

Студент: Пронин Арсений Сергеевич

Группа: ИУ7-82Б

Руководитель: Никульшина Татьяна Александровна

2023 г.

Актуальность

- Оптимизация пользовательского опыта
- Улучшение процесса разработки
- Выявление необычного или нежелательного поведения

Цель и задачи

Цель: разработать и программно реализовать метод анализа активности пользователей САПР с использованием поиска последовательных шаблонов.

Задачи:

- 1) рассмотреть существующие решения в области анализа активности пользователей, выбрать для них критерии оценки и сравнить;
- 2) формализовать задачу в виде IDEF0-диаграммы;
- 3) разработать метод анализа активности пользователей САПР с использованием поиска последовательных шаблонов;
- 4) разработать программное обеспечение, реализующее описанный метод;
- 5) исследовать характеристики разработанного метода.

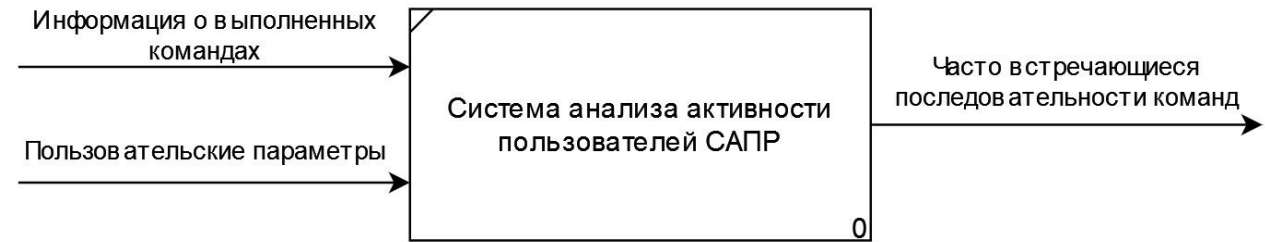
Сравнение рассмотренных методов

Метод	Требование к входным данным	Учет времени транз-ий	Сложность алгоритма
Мат. модель пользов. актив. ПО	Множество событий, функция классификации событий, множество сессий, множество последовательных шаблонов	Нет	$O(n \cdot m)$, где n – кол-во шаблонов, m – кол-во сессий
Apriori	Транзакции с набором элементов и минимальный уровень поддержки	Нет	$O(D \cdot I \cdot 2^{ I })$, где $ D $ – кол-во транзакций, $ I $ – общее число предметов
GSP	База данных с полями: id последовательности, id и время транзакции, набор элементов и минимальный уровень поддержки	Да	$O(I ^l)$, где $ I $ – общее число предметов, l – длина наибольшей ЧВП
GOMS	Последовательность действий	Нет	$O(n)$, где n – число действий в послед-ти

Формализованная постановка задачи

Ограничения на входные данные:

- Данные о выполненных командах должны содержать информацию об их последовательности
- Пользовательские параметры:
 - Минимальный уровень поддержки от 0 до 1
 - Минимальный и максимальный разрыв между командами от 0 до INT_MAX (2147483647)



Уровень поддержки

Пусть есть сессия:

Время транзакции	Команда
5	1
8	2
8	3

Сессия не поддерживает пос-ть $\langle 3, 2, 1 \rangle$

При $\text{min_gap} = 0$, $\text{max_gap} = 3$ сессия поддерживает пос-ть $\langle 1, 2, 3 \rangle$

При $\text{min_gap} > 0$ или $\text{max_gap} < 3$ сессия не поддерживает пос-ть $\langle 1, 2, 3 \rangle$

Коэффициент зависимости

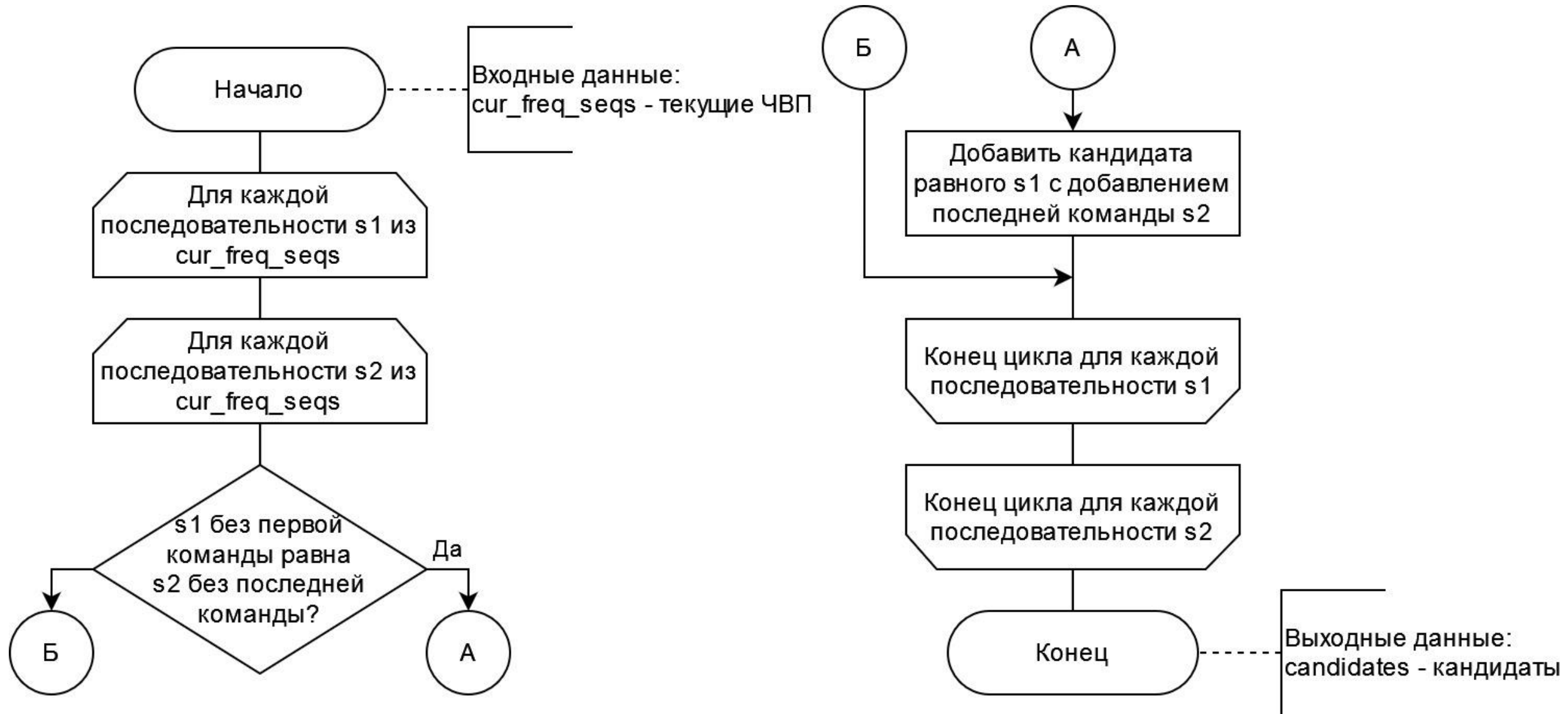
Пусть поддержка последовательности $\langle 1, 2, 3 \rangle$ равна 0.5, а последовательностей $\langle 1 \rangle$, $\langle 2 \rangle$ и $\langle 3 \rangle$ равны 0.6, 0.8 и 1 соответственно

В таком случае, коэффициент зависимости $= \frac{0.5}{0.6 \times 0.8 \times 1} \approx 1.042$

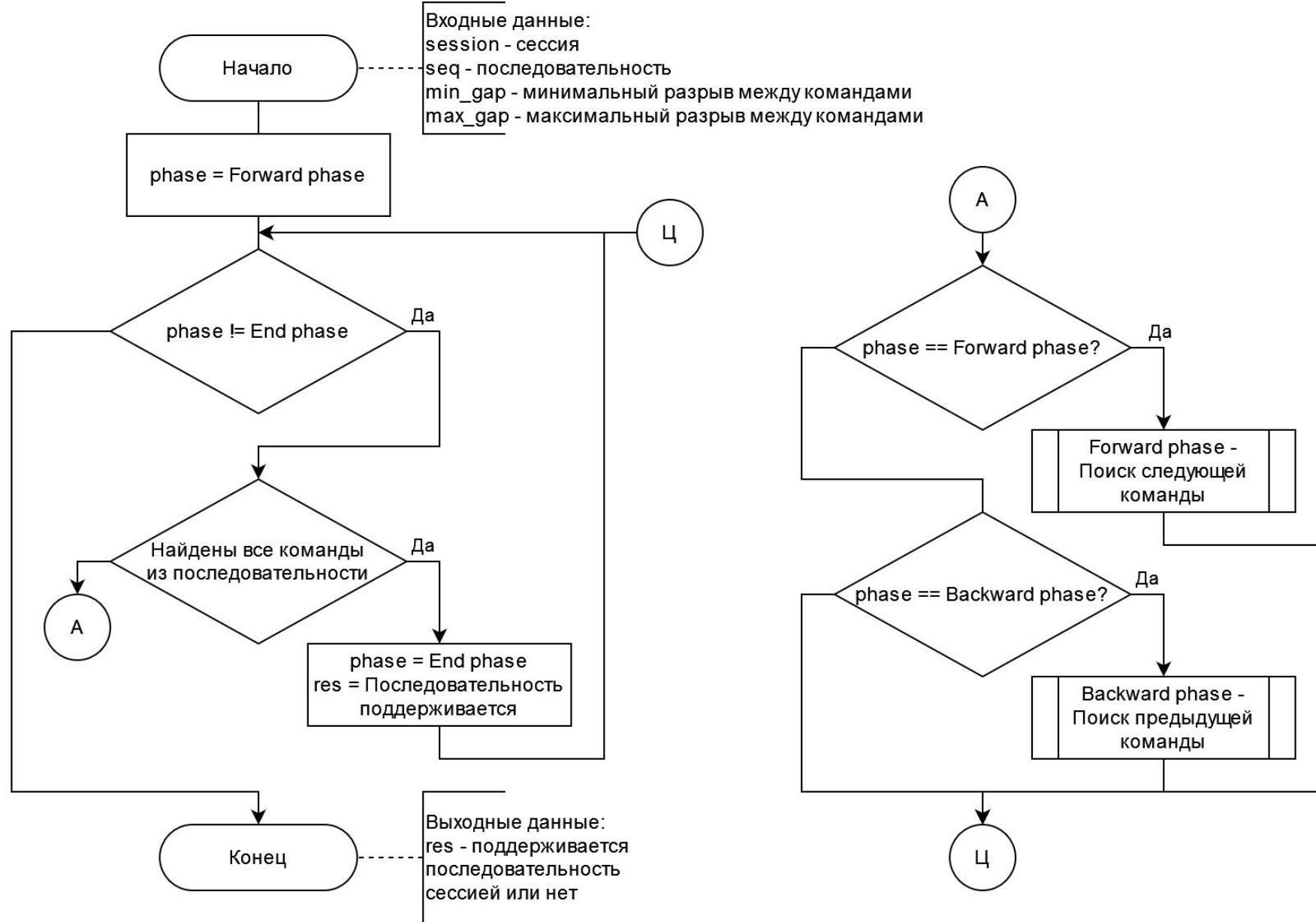
Если коэффициент ≤ 1 , то команды в последовательности независимы

Если коэффициент > 1 , то зависимость есть

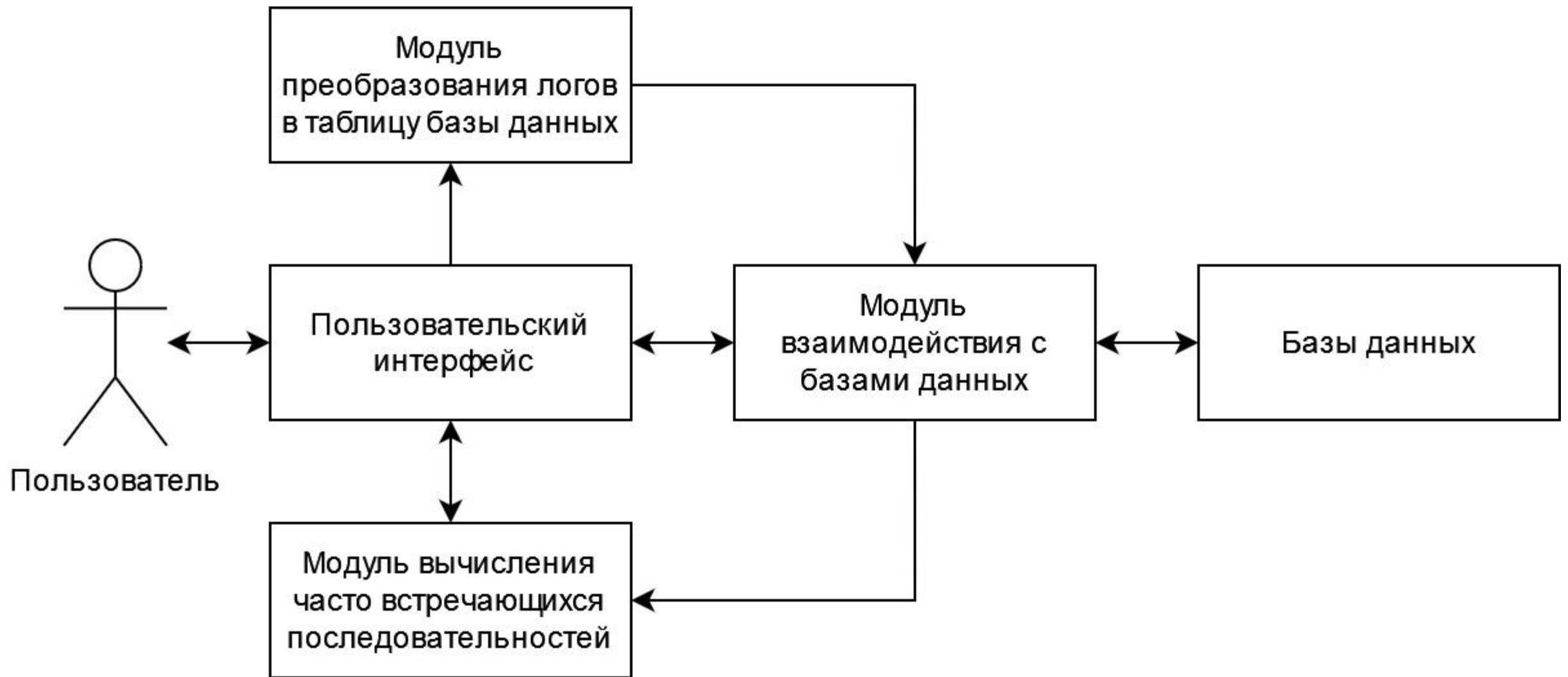
Генерация кандидатов



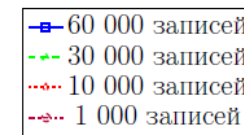
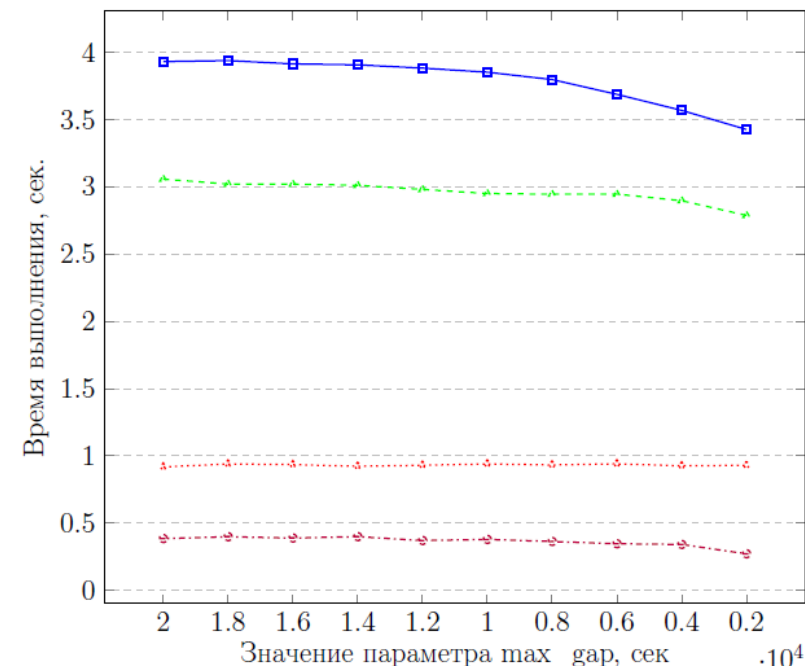
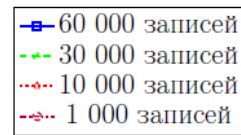
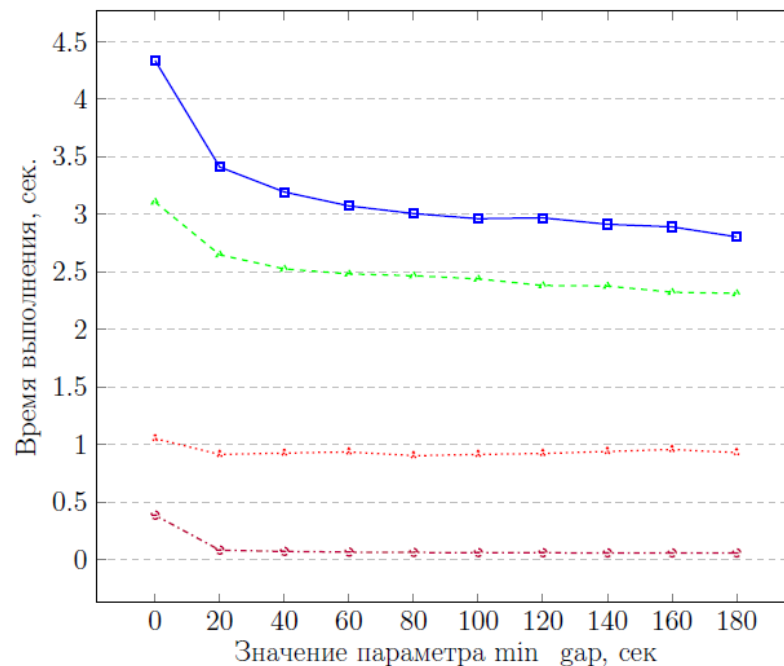
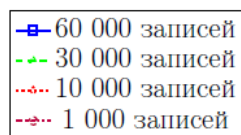
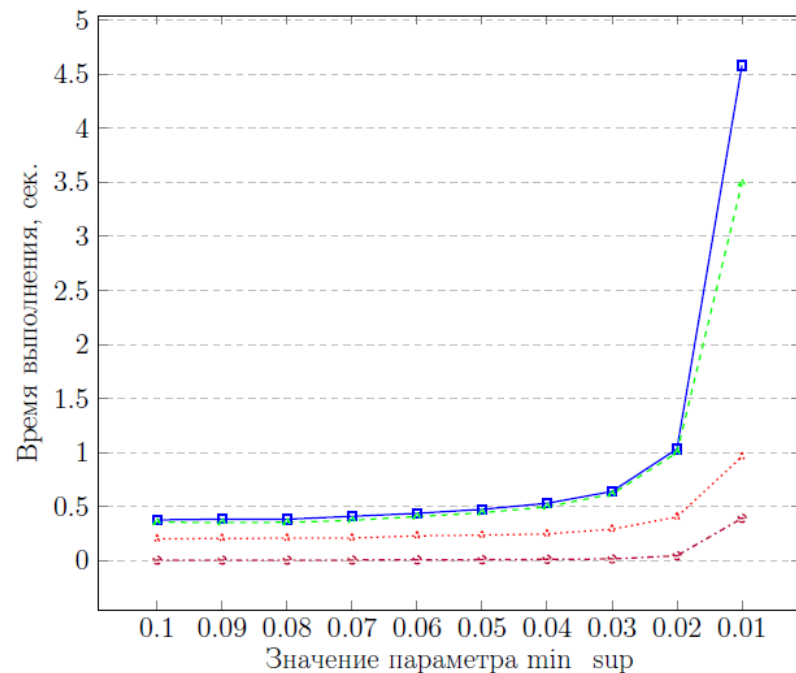
Проверка поддержки кандидата сессией



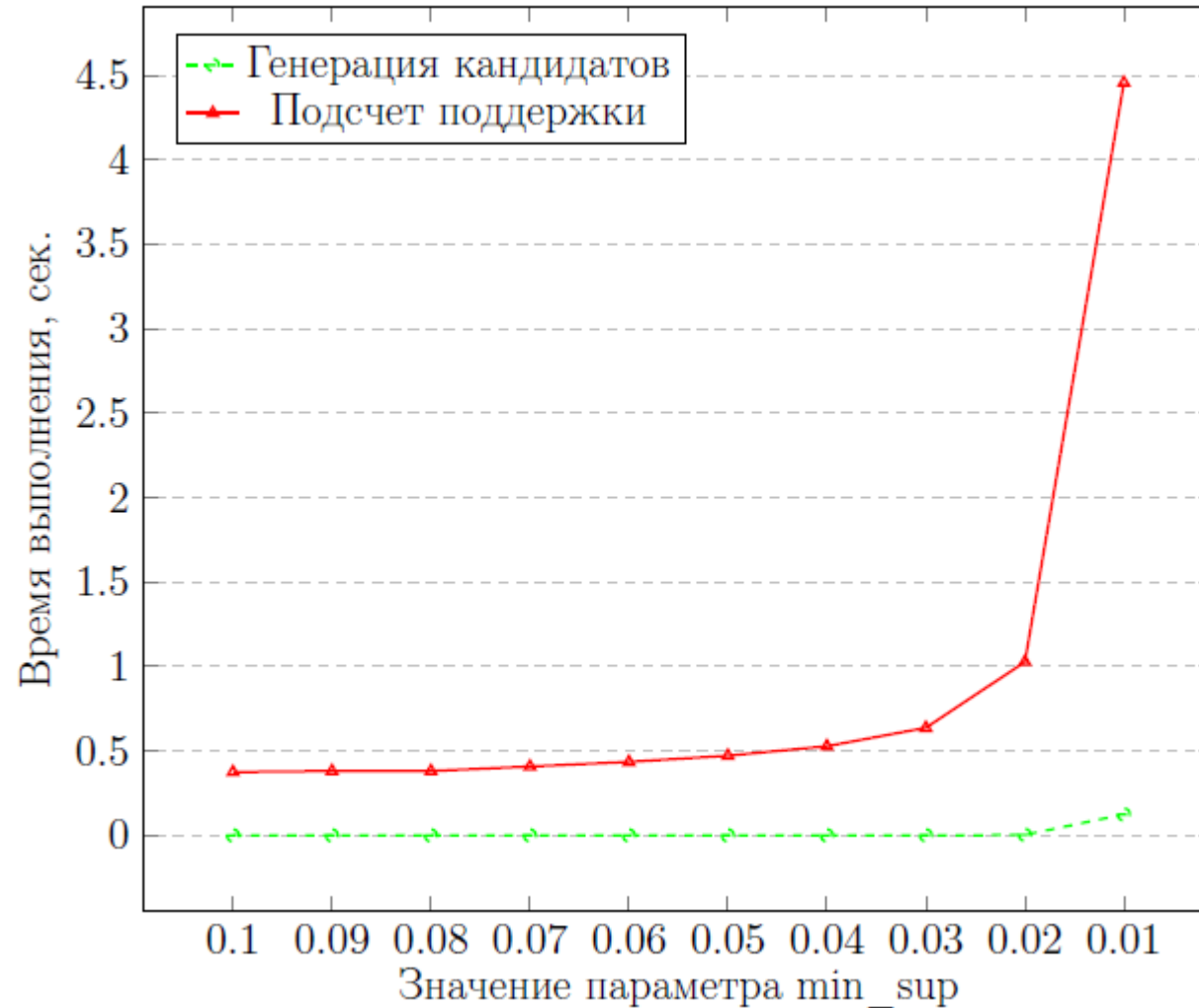
Структура программного обеспечения



Сравнительный анализ времени выполнения метода в зависимости от параметров



Сравнительный анализ времени выполнения этапов метода



Заключение

По итогу проделанной работы была достигнута **цель** - разработан и программно реализован метод анализа активности пользователей САПР с использованием поиска последовательных шаблонов.

Также были решены все поставленные **задачи**, а именно:

- 1) рассмотрены существующие решения в области анализа активности пользователей, выбраны для них критерии оценки и проведено сравнение;
- 2) формализована задача в виде IDEF0-диаграммы;
- 3) разработан метод анализа активности пользователей САПР с использованием поиска последовательных шаблонов;
- 4) разработано программное обеспечение, реализующее описанный метод;
- 5) исследованы характеристики разработанного метода.

Перспективы дальнейшего развития

- Возможность предсказания следующей команды
- Подсчет процента содержания найденных последовательностей в сессиях или других характеристик для получения дополнительной информации
- Оценка времени, необходимого для выполнения последовательности