



**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

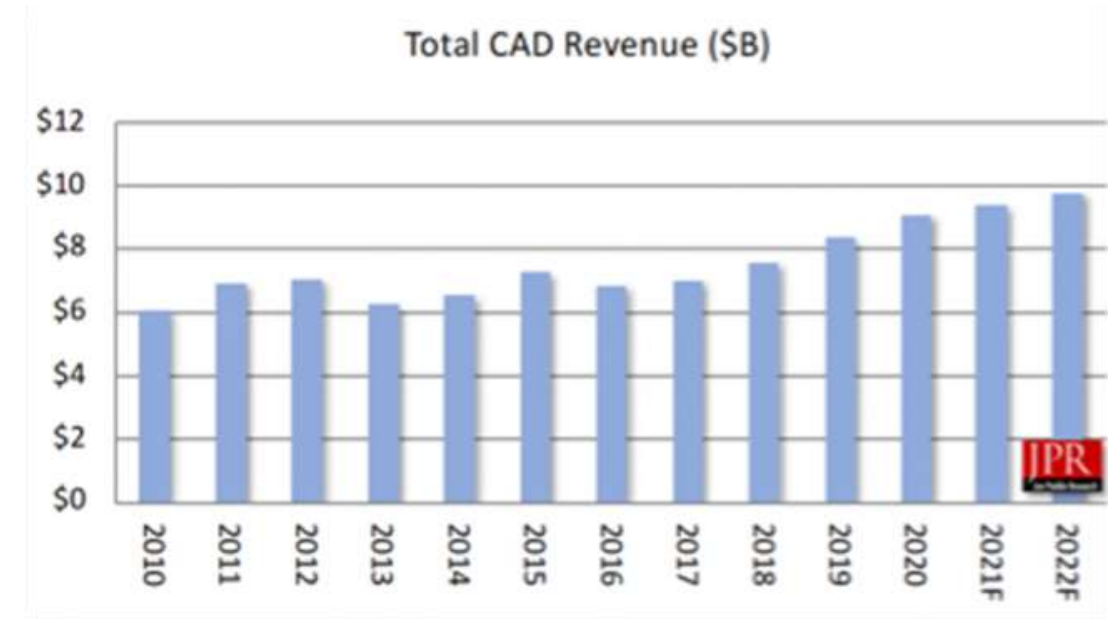
Руководитель: Никульшина Татьяна Александровна

2023 г.

# Актуальность

Показатели САПР по итогам 2021 года:

- Объем мирового рынка \$9,4 млрд
- 7,31 млн пользователей



Решаемые задачи:

- 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.

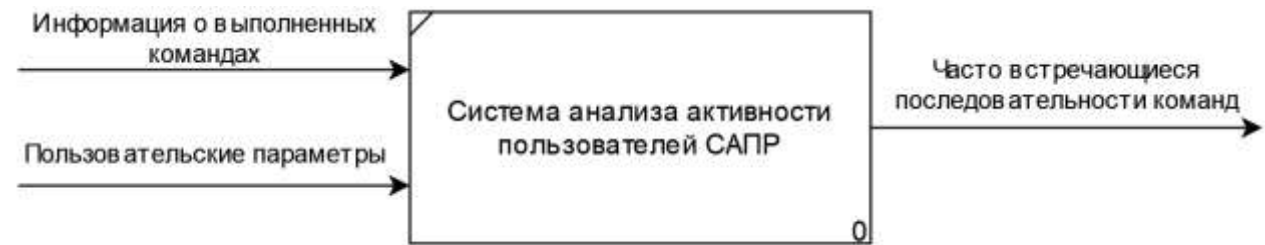
# Сравнение рассмотренных методов

| 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 временной разрыв между командами (в секундах) от 0 до INT\_MAX (2147483647)



# Уровень поддержки

Сессия 1:

| Время | Команда |
|-------|---------|
| 7     | 1       |
| 8     | 2       |
| 9     | 3       |

Сессия 2:

| Время | Команда |
|-------|---------|
| 7     | 3       |
| 8     | 1       |
| 9     | 2       |

Сессия 3:

| Время | Команда |
|-------|---------|
| 7     | 1       |
| 8     | 3       |
| 9     | 2       |

Сессия 4:

| Время | Команда |
|-------|---------|
| 7     | 2       |
| 8     | 3       |
| 9     | 1       |

При минимальном и максимальном временном разрыве между командами от 0 до 2 секунд, три первые сессии содержат последовательность  $\langle 1, 2 \rangle$ , поэтому ее поддержка  $= \frac{3}{4} = 0.75$

| Последовательность     | Поддержка |
|------------------------|-----------|
| $\langle 1, 2 \rangle$ | 0.75      |

# Коэффициент зависимости

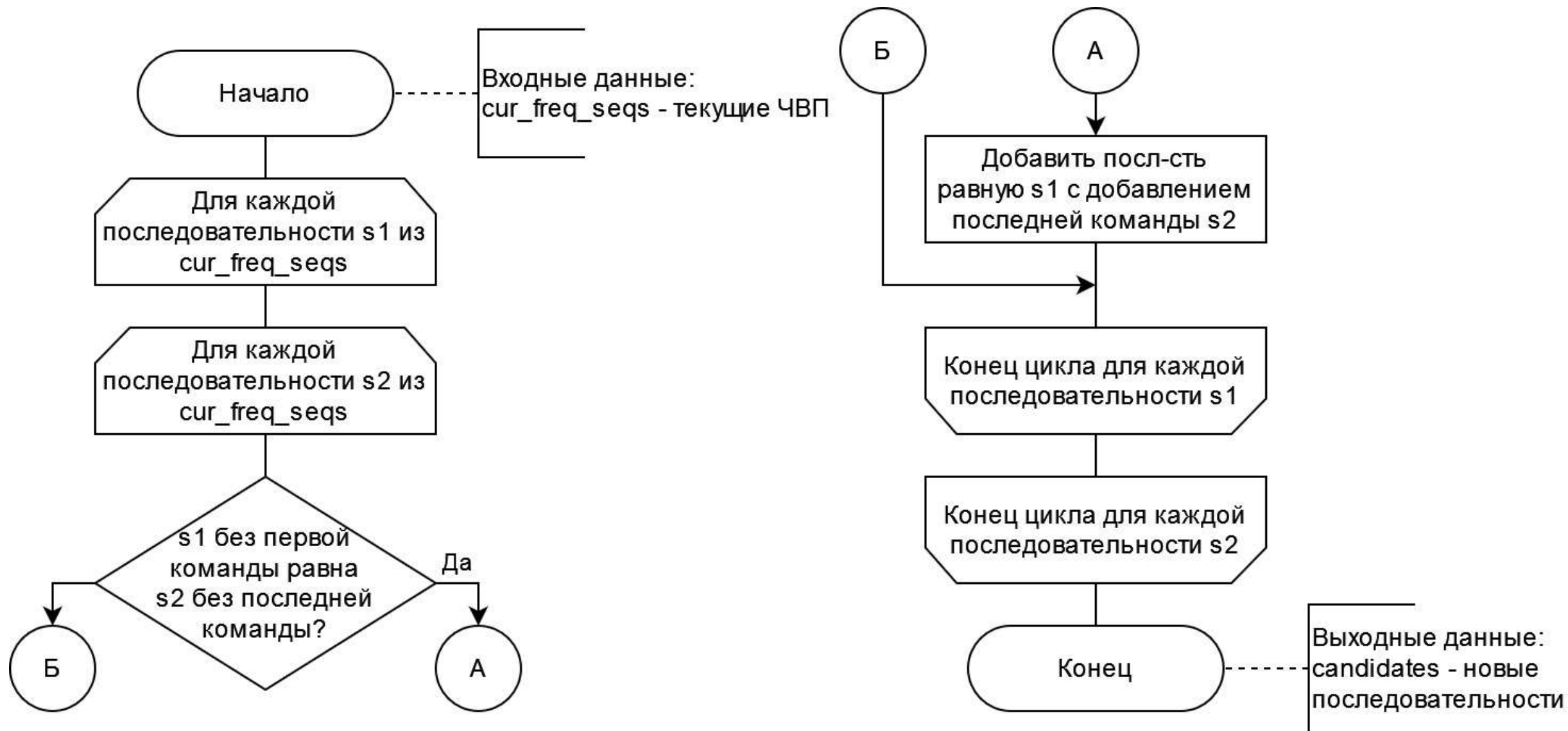
Пусть поддержка последовательности  $\langle 1, 2, 3 \rangle$  равна 0.5, а последовательностей  $\langle 1 \rangle$ ,  $\langle 2 \rangle$  и  $\langle 3 \rangle$  равны 0.6, 0.8 и 1 соответственно

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

Если коэффициент  $\leq 1$ , то команды в последовательности независимы

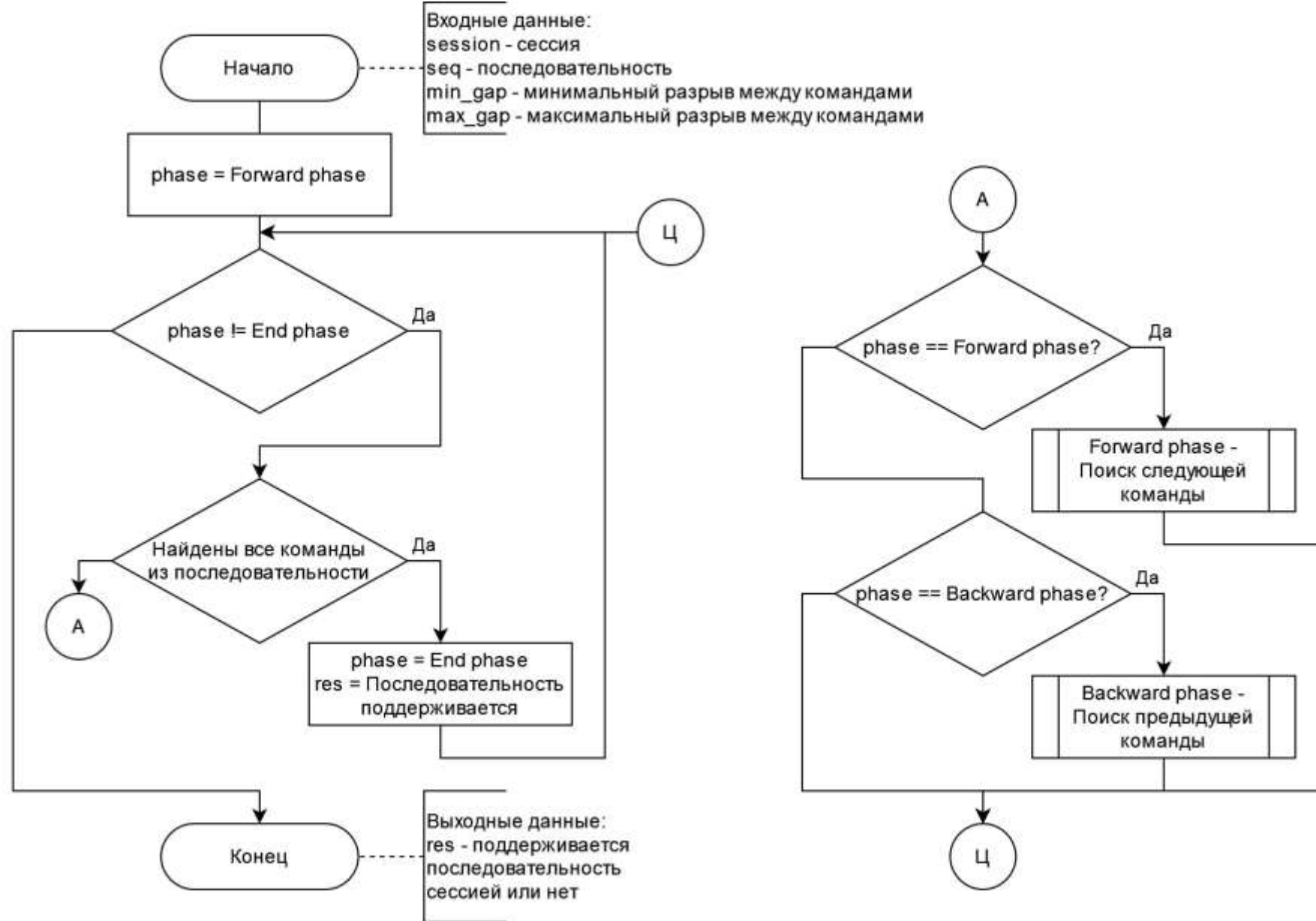
Если коэффициент  $> 1$ , то зависимость есть

# Генерация последовательностей

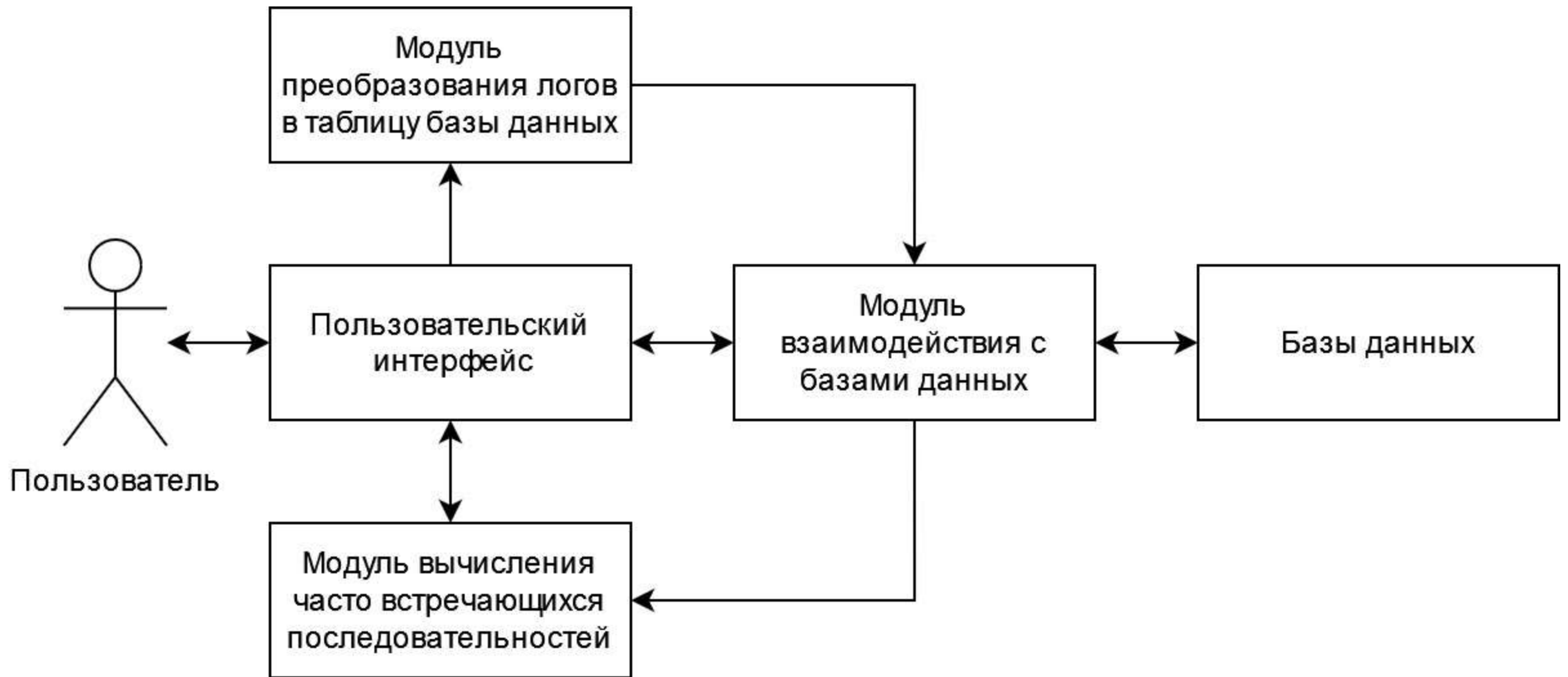




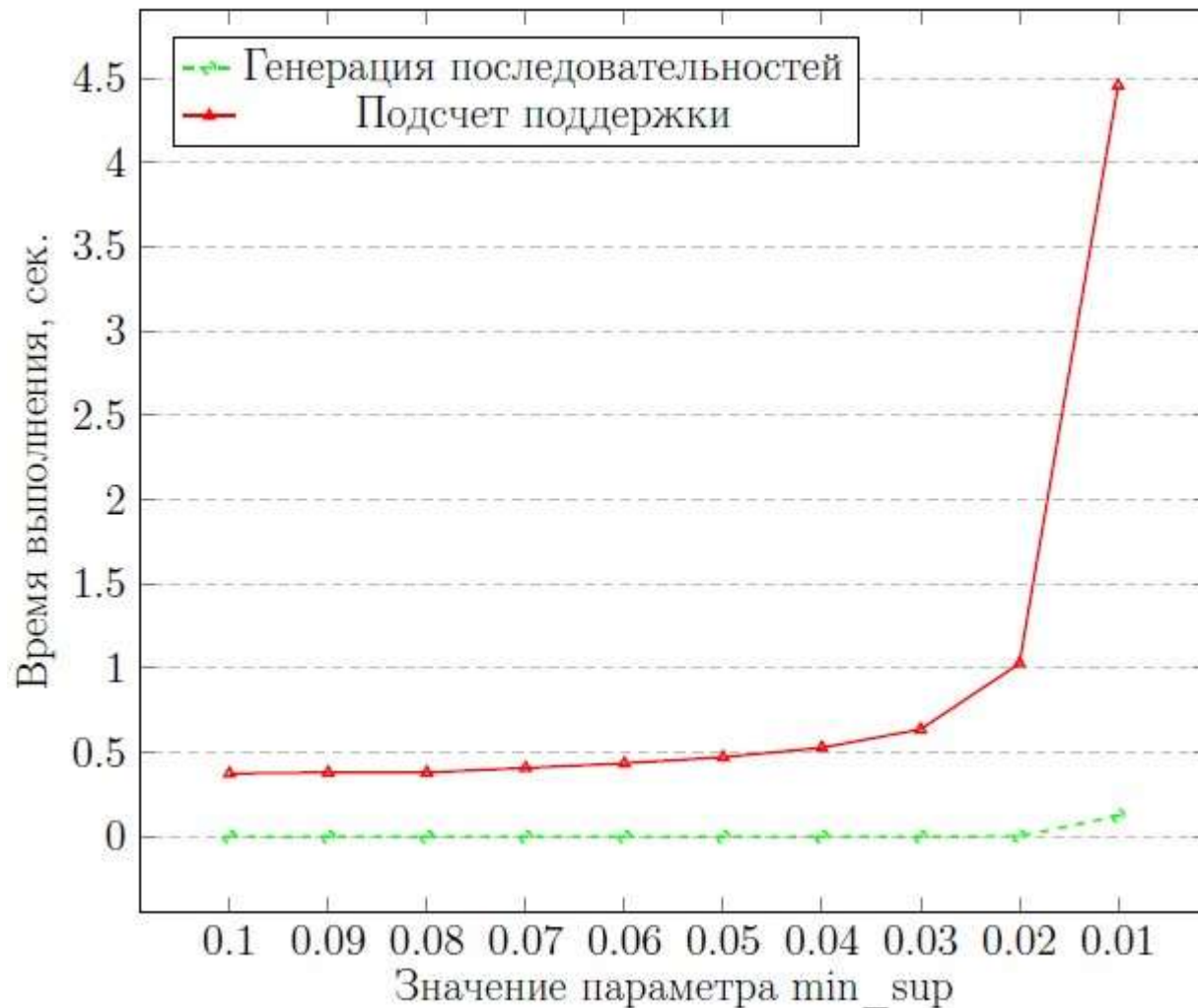
# Проверка поддержки последовательности сессией



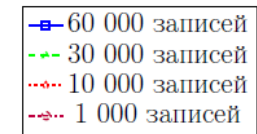
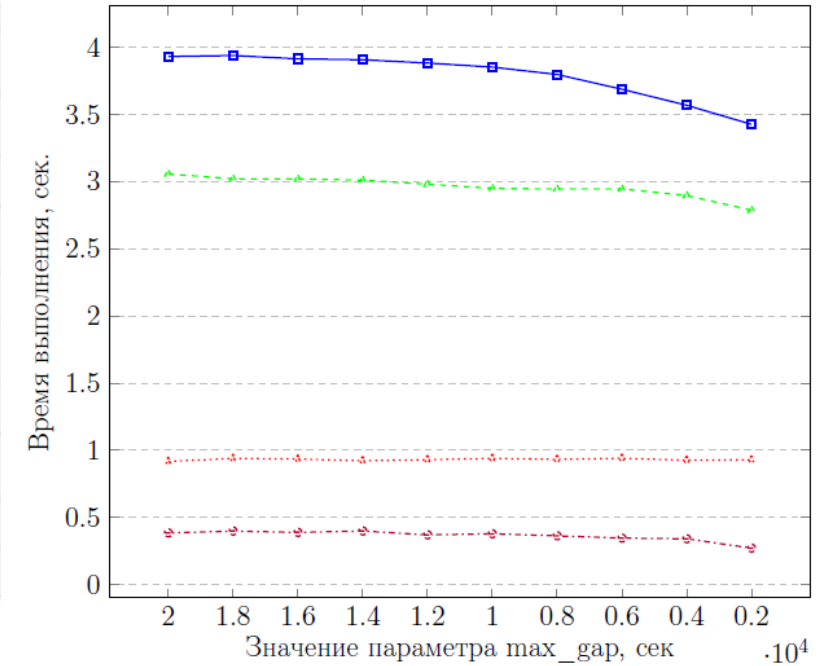
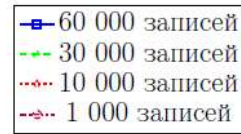
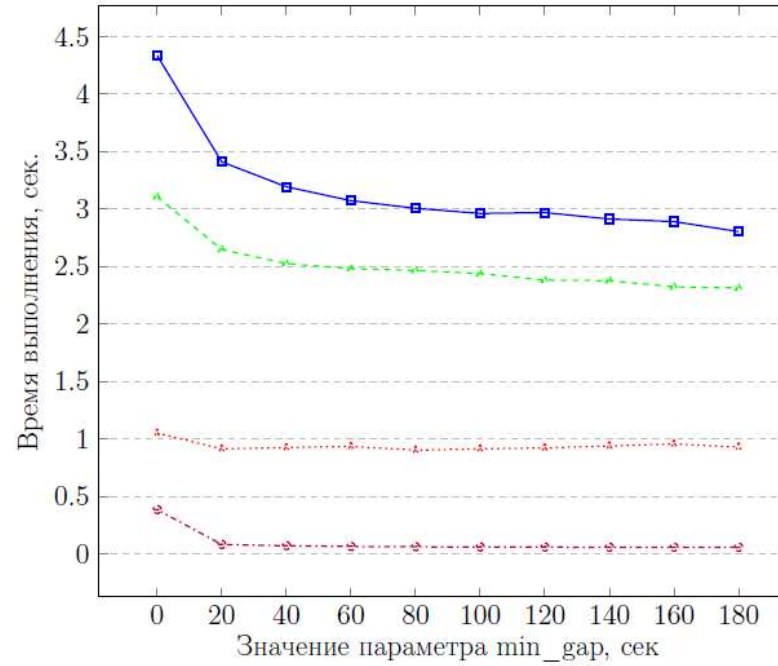
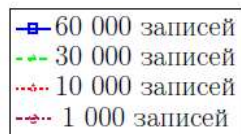
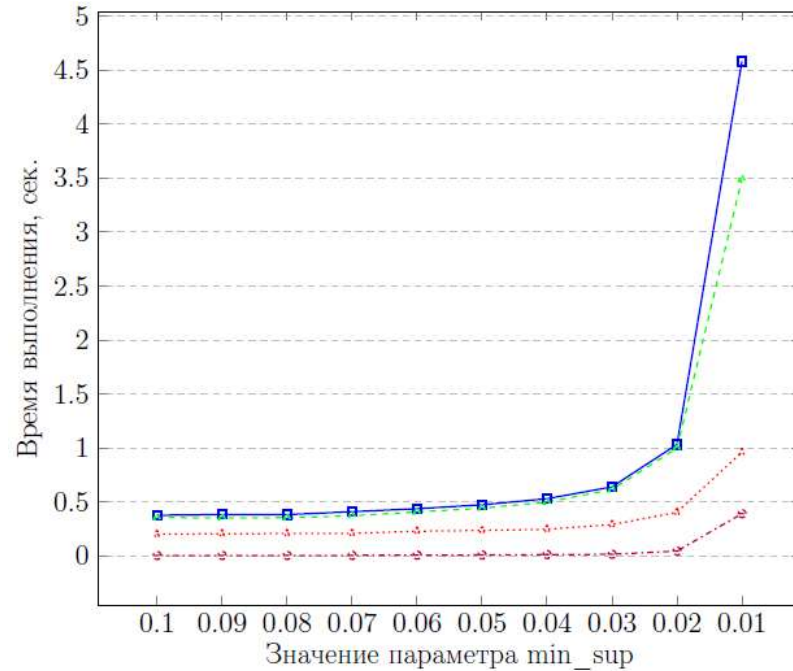
# Структура программного обеспечения



# Сравнительный анализ времени выполнения этапов метода



# Сравнительный анализ времени выполнения метода в зависимости от параметров



# 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
- Оценка времени, необходимого для выполнения последовательности