

*Ожигов Д.И.,  
студент магистратуры  
2 курс, факультет «Информатики»  
Самарский национальный исследовательский университет  
имени академика С.П. Королева  
Россия, г. Самара*

*Научный руководитель: Сопченко Е.В.  
Доцент, кандидат технических наук  
Кафедра программных систем  
Самарский национальный исследовательский университет  
имени академика С.П. Королева  
Россия, г. Самара*

## **РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ МОНИТОРИНГА И АНАЛИЗА ЖУРНАЛОВ СИСТЕМЫ В РЕАЛЬНОМ ВРЕМЕНИ**

***Аннотация:** Статья посвящена описанию разработанной автоматизированной системы мониторинга и анализа журналов системы в реальном времени. Данная система предназначена для сбора, обработки и анализа данных из журналов тестируемой системы в реальном времени по обученным шаблонам, а также для представления пользователю извлеченной информации в различном виде.*

***Ключевые слова:** данные, информация, журнал, мониторинг, сбор, обработка, анализ.*

## DEVELOPMENT OF AN AUTOMATED SYSTEM FOR MONITORING AND ANALYZING SYSTEM LOGS IN REAL TIME

***Annotation:** The article describes the developed automated system for monitoring and analyzing system logs in real time. This system is designed for collecting, processing and analyzing data from the logs of the tested system in real time using trained templates, as well as for presenting the extracted information to the user in various forms.*

***Key words:** data, information, log, monitoring, collection, processing, analysis.*

В процессе разработки любая программа постоянно подвергается изменениям. Модификации, как правило, могут привести к появлению различных дефектов и сбоев в программе. Поэтому особую актуальность приобретают вопросы тестирования как деятельности, повышающей качество программных продуктов [1]. Соответственно есть необходимость в разработке специальных программных инструментов, которые позволят ускорить и повысить эффективность тестирования того или иного продукта. Многие крупные разрабатываемые системы имеют в своем составе подсистему ведения журналов, в которых содержится максимально полная информация о работоспособности системы. Отслеживание журналов в автоматическом режиме при помощи программных средств позволяет автоматизировать процесс проверки систем. Автоматизированный анализ журналов позволяет отслеживать, что все функции продукта стабильны в процессе его работы и различные скрытые сбои отсутствуют. Наравне с обнаружением ошибок также остается важным выяснение причин их возникновения. Данный анализ позволяет существенно увеличить эффективность и скорость устранения дефектов разработчиками за счёт автоматизации анализа журналов.

Главными свойствами журналов являются: тип журнала (один файл на протяжении всего анализа или файлов несколько, и они индексируются), частота записи и способ записи (мгновенный или постепенный).

Все данные в журналах представлены в виде строк. Полезная информация может находиться как в одной строке, так и в нескольких в строго идущем порядке строках. Для этого был введен термин «шаблон», который может быть простым (одна строка) или связанным (несколько строк). При этом также учитывается, что в каждой строке может быть несколько полезной информации, которую необходимо уметь извлекать.

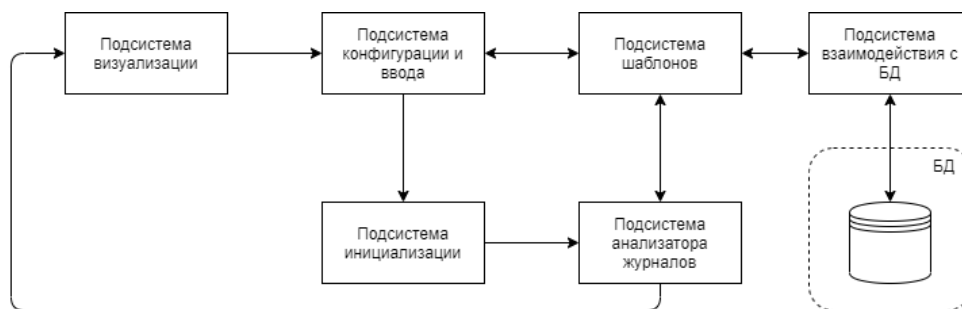
В качестве полезной информации может выступать как ключевое слово для идентификации нужной строки журнала, так и границы интервала, в которых она находится. Границей интервала может быть слово, индекс, или связка с предыдущей или следующей полезной информацией в этой же строке.

Разработанная система представлена в виде настольного кроссплатформенного приложения, способного обеспечить возможность обучения различных типов журналов с учетом всех их свойств и особенностей с целью извлечения из них полезной информации в реальном времени, которую можно представить с помощью различных видов отображения (график, диаграмма, таблица и текст). Для полезной информации с целочисленным типом данных в приложении представлен функционал отображения статистических данных по собранной информации и осуществления анализа с целью выявления характера ее зависимости от времени. Также в разработанном приложении реализован функционал сохранения в архив накопленной статистики с возможностью дальнейшего ее сравнения с другими сохраненными статистиками по этой же информации и определения ее зависимости от пользовательского параметра.

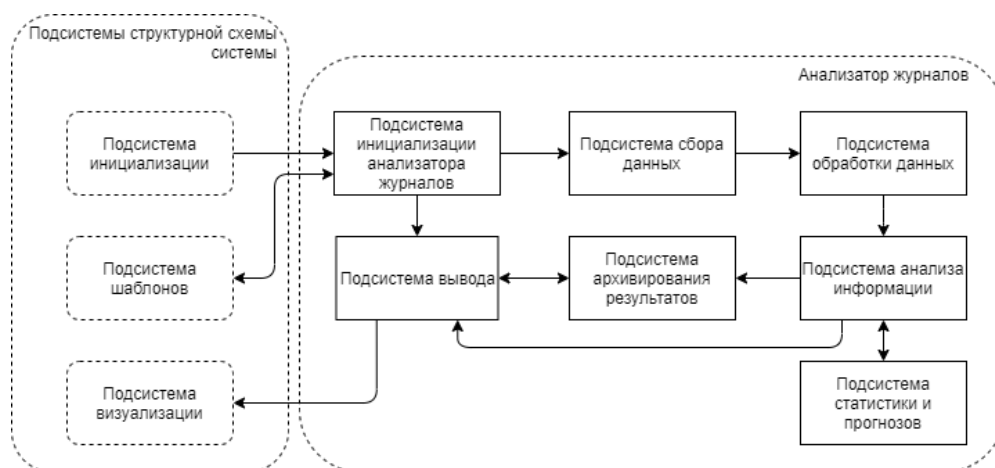
На рисунке 1 представлена общая структурная схема системы. Система состоит из подсистем конфигурации и ввода, шаблонов, анализатора журналов и других вспомогательных подсистем, реализующие дополнительные функции. Приложение содержит базу данных, за связь с которой отвечает подсистема взаимодействия с БД.

На рисунке 2 представлена структурная схема подсистемы анализатора журналов. Данная подсистема состоит из подсистем сбора данных, их обработки,

анализа информации, статистики и прогнозов, архивирования и других вспомогательных подсистем.

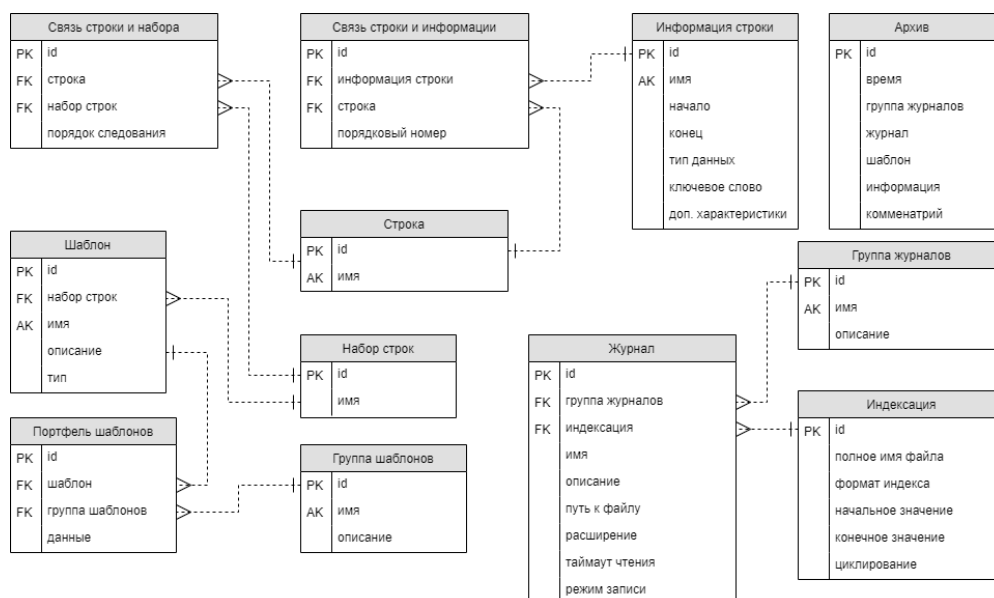


**Рисунок 1. Общая структурная схема системы**



**Рисунок 2. Структурная схема анализатора журналов**

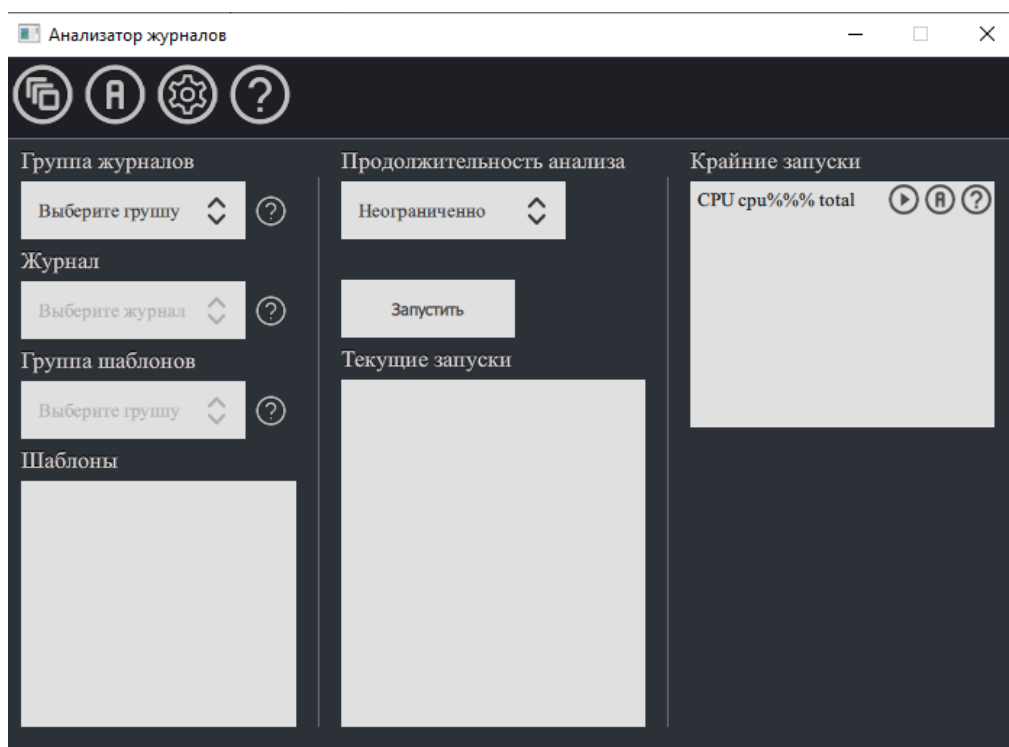
На рисунке 3 представлена логическая модель базы данных.



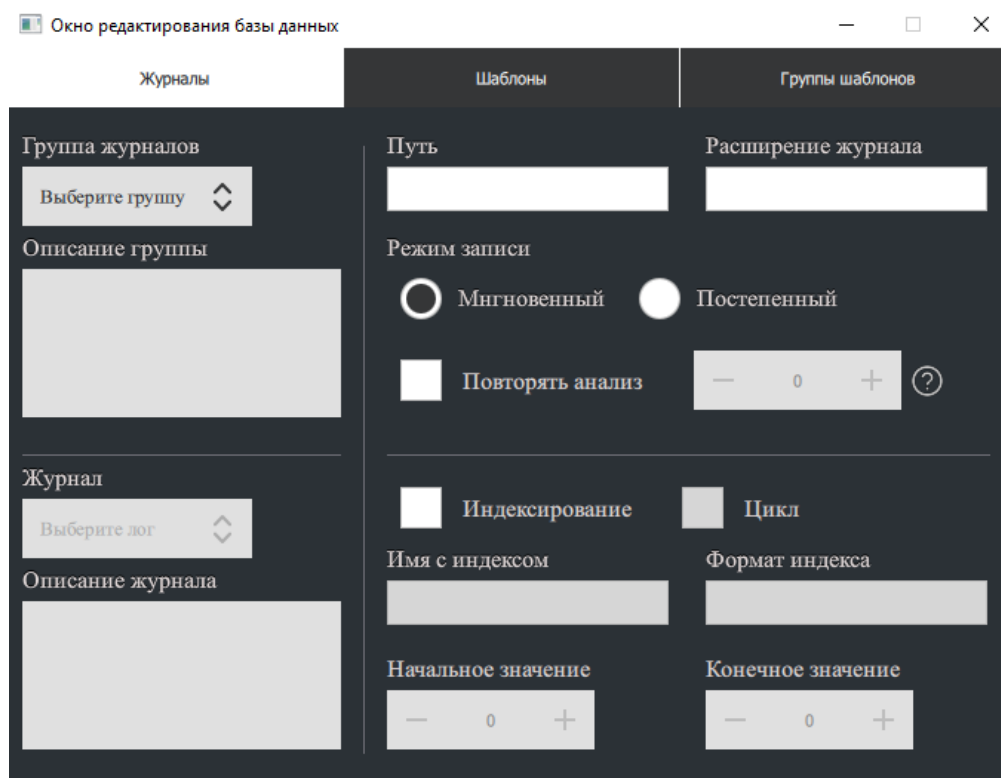
**Рисунок 3. Логическая модель базы данных системы**

В системе были реализованы главное окно (рисунок 4), окно редактирования базы данных с вкладками «Журналы», «Шаблоны» и «Группы

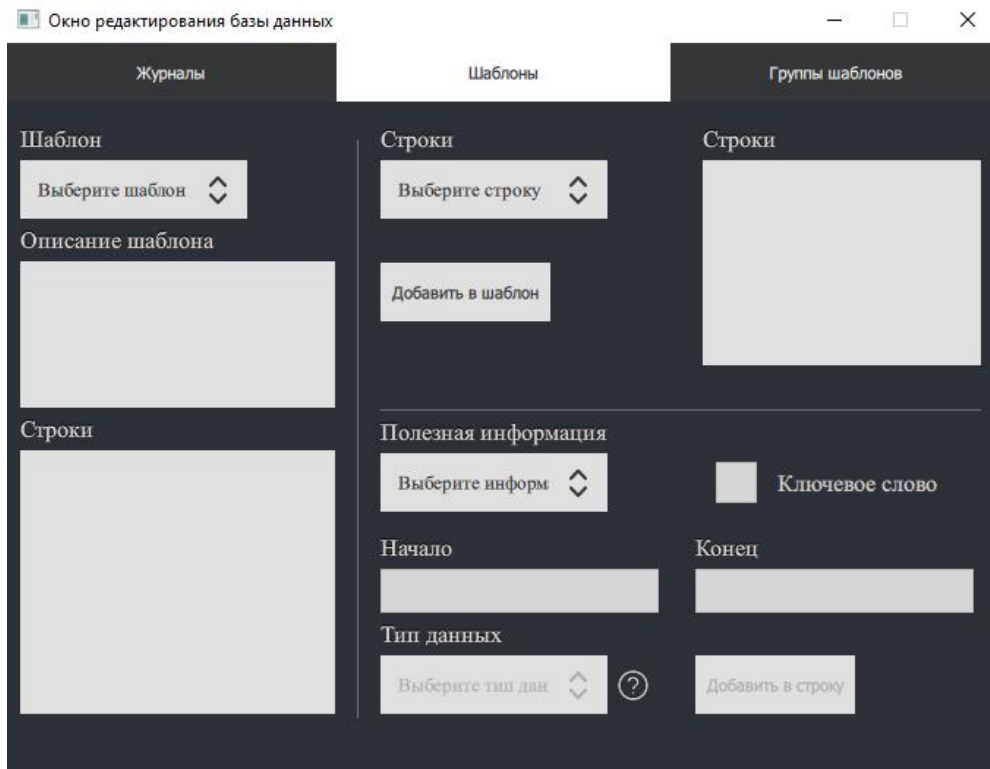
шаблонов» (рисунки 5 – 6), окно работы с архивом, окно просмотра результатов (рисунок 7), окно общих настроек и справки.



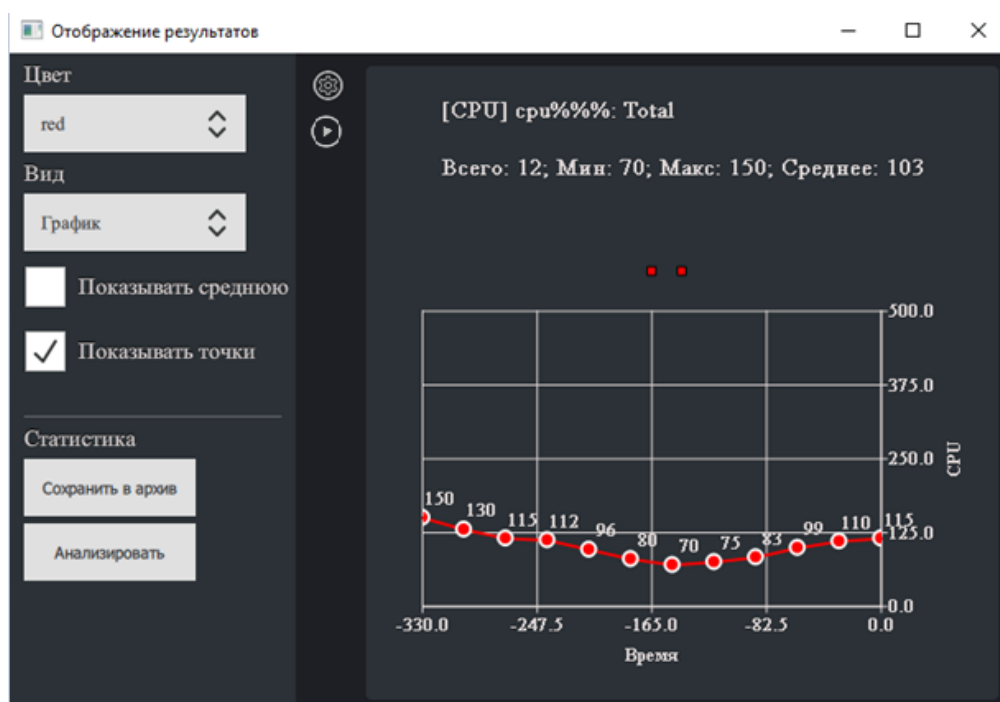
**Рисунок 4. Главное окно системы**



**Рисунок 5. Вкладка «Журналы» окна редактирования базы данных**



**Рисунок 6. Вкладка «Шаблоны» окна редактирования базы данных**



**Рисунок 7. Окно отображения результатов в режиме просмотра статистики**

Также было реализовано четыре вида представления результатов на форме просмотра результатов (график, таблица, диаграмма и текст).

С помощью разработанной системы можно осуществлять нагрузочное тестирование и оценивать характер зависимости. В качестве примера оценим характер зависимости нагрузки на ЦП модуля распознавания лиц некой тестируемой системы от количества одновременных лиц в кадре.

Известно, что основным функционалом, максимально влияющим на нагрузку данного модуля на ЦП, является детекция (обнаружение) лиц на кадрах, поступающих с камеры. Соответственно, ставится гипотеза о том, что с увеличением количества одновременных лиц в кадре нагрузка на ЦП возрастает. В случае если гипотеза подтверждается, то необходимо определить характер зависимости нагрузки на ЦП модуля распознавания лиц от количества лиц, а также осуществить регрессионный анализ и найти наиболее подходящую регрессионную модель, которая будет описывать уравнение зависимости. Все это позволяет осуществить разработанная система.

Для проведения нагрузочного тестирования и подтверждения (или опровержения) гипотезы необходимо с помощью разработанной системы осуществить сбор статистики параметра CPU из соответствующего журнала (который ведется в тестируемой системе) при различных условиях, сохранить каждую статистику в архив, далее открыть их в режиме сравнения, оценить характер зависимости и построить регрессионную модель.

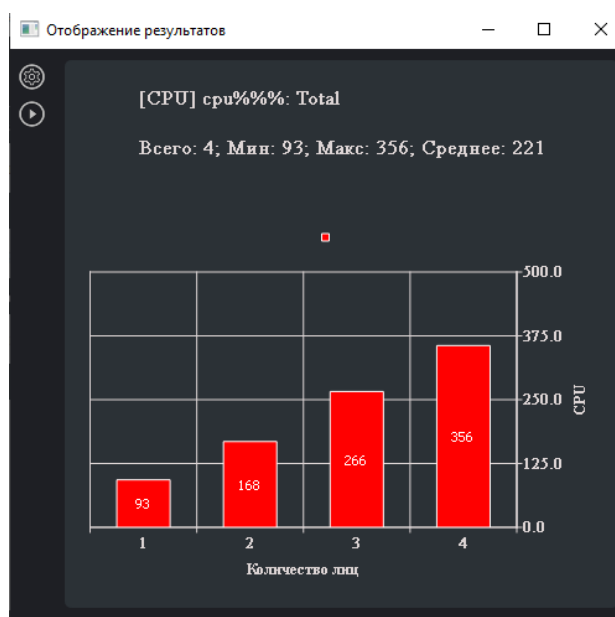
Всего будет четыре статистики по параметру CPU для четырех разных условий: одно лицо в кадре, два лица в кадре, три лица в кадре и четыре лица в кадре. Перед запуском анализатора журналов необходимо подготовить нужное начальное условие на тестируемой системе. Также необходимо в разработанной системе обучить нужный шаблон для параметра CPU.

Для того, что сравнить ранее собранные статистики, необходимо перейти в меню архива, выбрать нужные статистики для сравнения и кликнуть по кнопке открытия результатов (рисунок 8).

Дата сохранения	Группа журналов	Журнал	Шаблон	Параметр	Комментарий
2020-05-08 10:46:32	CPU	cpu%%%	cpu	CPU	[1] лица
2020-05-08 11:38:33	CPU	cpu%%%	cpu	CPU	[2] лица
2020-05-08 12:42:15	CPU	cpu%%%	cpu	CPU	[3] лица
2020-05-08 13:26:53	CPU	cpu%%%	cpu	CPU	[4] лица
2020-05-08 13:27:02	CPU	cpu%%%	cpu	CPU	4 лица
2020-05-08 16:12:43	MEMORY	mem%%%	mem	Working set	Мониторинг памяти

**Рисунок 8. Открытие нужных статистик их архива в режиме сравнения**

При открытии статистик в режиме сравнения по каждой из них берется среднее значение и выводится на вид представления. На рисунке 9 представлена диаграмма сравнения статистик.



**Рисунок 9. Отображение сравнения статистик в виде диаграммы**

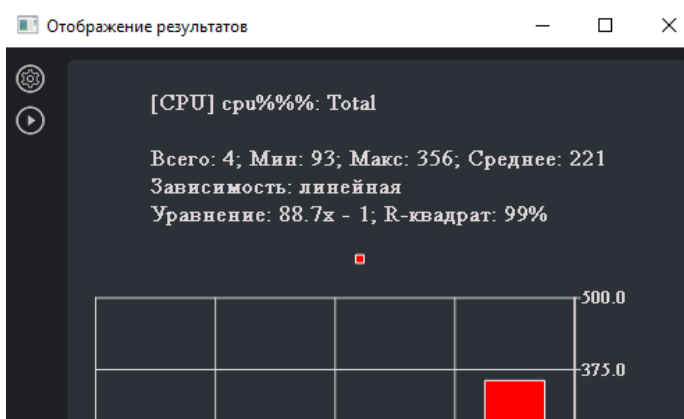
Для определения зависимости параметра CPU от количества лиц необходимо в боковом выдвигаемом меню в секции «Статистика» кликнуть по кнопке «Анализировать».

После этого программа начинает осуществлять регрессионный анализ по данной статистике, используя линейную и полиномиальные модели со степенями два и три. Затем для каждой построенной модели вычисляется



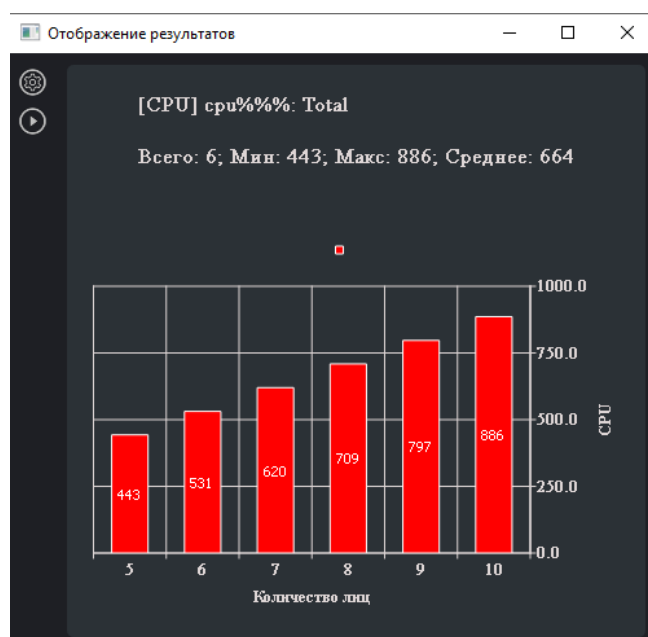
коэффициент детерминации. После чего определяется самая качественная модель, которая и будет выведена на экран.

На рисунке 10 изображен результат нахождения зависимости для данной статистики. Видно, что зависимость параметра CPU от количества одновременных лиц в кадре линейна, следовательно, гипотеза о наличии зависимости между данными двумя параметрами подтверждается. Также была подобрана линейная регрессионная модель с коэффициентом детерминации 99%.



**Рисунок 10. Результат регрессионного анализа**

После нахождения регрессионной модели можно составить прогноз при других количествах лиц в кадре (рисунок 11).



**Рисунок 11. Прогноз параметра CPU по построенной регрессионной модели**

### **Использованные источники:**

1. Автоматизированное тестирование ПО [Электронный ресурс]. URL: [https://revolution.allbest.ru/programming/00635261\\_0.html](https://revolution.allbest.ru/programming/00635261_0.html) (дата обращения: 20.05.2019).