



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)

---

ФАКУЛЬТЕТ \_\_\_\_\_ «Информатика и системы управления»  
КАФЕДРА \_\_\_\_\_ «Программное обеспечение ЭВМ и информационные технологии»

---

**РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА  
К ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЕ  
НА ТЕМУ:**

«Метод систематического распознавания усталости на  
автоматизированном рабочем месте»

Студент группы ИУ7-83Б

\_\_\_\_\_  
(Подпись, дата)      Д.В. Якуба  
(И.О. Фамилия)

Руководитель

\_\_\_\_\_  
(Подпись, дата)      Ю.В. Строганов  
(И.О. Фамилия)

Нормоконтролер

\_\_\_\_\_  
(Подпись, дата)      (И.О. Фамилия)

Москва — 2022г.

# РЕФЕРАТ

Расчетно-пояснительная записка 44 с., 5 рис., 1 табл., 26 ист., 0 прил.

Объектом разработки является метод систематического распознавания усталости на автоматизированном рабочем месте.

Цель работы – разработать и реализовать метод распознавания усталости оператора автоматизированного рабочего места по данным, приходящим с устройств взаимодействия пользователя с системой.

Для достижения поставленной цели необходимо:

- провести анализ существующих методов определения усталости;
- провести анализ действий и характеристик, позволяющих определить усталость пользователя АРМ;
- определить методы снятия выделенных действий и характеристик;
- разработать метод распознавания усталости оператора;
- реализовать разработанный метод.

# СОДЕРЖАНИЕ

<b>ВВЕДЕНИЕ</b>	<b>4</b>
<b>1 Аналитический раздел</b>	<b>5</b>
1.1 Термины предметной области . . . . .	5
1.1.1 Усталость . . . . .	5
1.1.2 Хроническая усталость . . . . .	5
1.1.3 Стресс . . . . .	6
1.1.4 Стадии общего адаптационного синдрома . . . . .	6
1.1.5 Профессиональный стресс . . . . .	7
1.2 Формализация цели метода систематического распознавания усталости на автоматизированном рабочем месте . . . . .	9
1.3 Методы определения усталости оператора с использованием устройств автоматизированного рабочего места . . . . .	9
1.3.1 Метод анализа клавиатурного почерка . . . . .	10
1.3.2 Метод анализа скорости печати и количества ошибок . . . . .	12
1.3.3 Метод анализа траекторий перемещения курсора мыши . . . . .	12
1.3.4 Метод анализа использования координатного устройства мышь с использованием модели искусственной нейронной сети . . . . .	14
1.3.5 Метод анализа внешнего состояния пользователя . . . . .	15
1.3.6 Метод анализа речевых характеристик пользователя . . . . .	15
1.3.7 Метод анализа виброакустических шумов при наборе текста или использовании мыши . . . . .	17
1.4 Биофизические факторы, позволяющие определить усталость . . . . .	17
1.4.1 Частота пульса и возраст сосудистой системы . . . . .	17
1.4.2 Индекс Баевского . . . . .	18
1.5 Формализация требований к методу систематического распознавания усталости на автоматизированном рабочем месте . . . . .	19
<b>2 Конструкторский раздел</b>	<b>20</b>
2.1 Включаемые в метод характеристики, получаемые с использованием периферийных устройств . . . . .	20
2.2 Метод систематического распознавания усталости на автоматизированном рабочем месте . . . . .	20
2.3 Формат и метод сбора данных, предоставляемых оператором автоматизированного рабочего места . . . . .	21
2.4 Диаграмма вариантов использования . . . . .	22
2.5 IDEF0-диаграмма . . . . .	23
2.6 Диаграмма “сущность-связь” в нотации Чена . . . . .	24
2.7 Используемые методы исследования корреляции . . . . .	24

2.7.1	Парная корреляция . . . . .	24
2.7.2	Множественная корреляция . . . . .	25
2.8	Используемые методы дисперсионного анализа . . . . .	26
2.8.1	Однофакторный дисперсионный анализ . . . . .	26
2.8.2	Двухфакторный дисперсионный анализ с однократными наблюдениями . . . . .	28
2.9	Алгоритмы кластеризации . . . . .	29
2.9.1	Алгоритмы иерархической кластеризации . . . . .	30
2.9.2	Алгоритмы квадратичной ошибки . . . . .	30
2.9.3	Нечеткие алгоритмы . . . . .	31
<b>3</b>	<b>Технологическая часть</b>	<b>33</b>
3.1	Средства реализации программного обеспечения . . . . .	33
3.2	Выбор СУБД . . . . .	33
3.2.1	Базы данных временных рядов . . . . .	33
3.2.2	Реляционные базы данных . . . . .	34
3.3	Выбор алгоритма кластеризации . . . . .	35
3.4	Данные для кластеризации . . . . .	35
3.5	Сведения о модулях . . . . .	35
3.5.1	Модуль логирования действий оператора . . . . .	35
3.5.2	Модуль обработки данных . . . . .	38
3.5.3	Модуль анализа данных . . . . .	38
3.6	Развертывание системы . . . . .	38
	<b>ЗАКЛЮЧЕНИЕ</b>	<b>39</b>
	<b>СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ</b>	<b>41</b>

## ВВЕДЕНИЕ

Согласно исследованиям [1], на момент 2019 года 28% сотрудников постоянно или достаточно часто чувствовали себя угнетёнными под давлением рабочих обязанностей, а 48% страдали синдромом эмоционального выгорания время от времени.

На сегодняшний день проблема эмоционального выгорания касается не только самих работников, но и компаний, которые при утрате контроля над ситуацией вынуждены увольнять сотрудников под предлогом неисполнения ими обязанностей. [2]

Причиной синдрома может быть и физическое, и эмоциональное истощение вследствие увеличения нагрузки на работе, а также количества возлагаемых обязанностей на сотрудника. [3]

Синдром хронической усталости также является одним из факторов, понижающих работоспособность сотрудников. Несмотря на то, что этиология данного заболевания до конца не раскрыта, одним из возможных факторов появления данного синдрома приписывают высокой нагрузке как умственной, так и физической. [4]

Усталость негативно влияет на производительность труда, а также на психологическое и физическое состояние человека. В условиях современной цифровизации медицины становится реальным учитывать индивидуальные особенности организма, управление его работоспособностью и проведение профилактики проявления вышеописанных синдромов, приводящих к неутешительным последствиям.

# 1. Аналитический раздел

В данном разделе...

## 1.1 Термины предметной области

### 1.1.1 Усталость

*Усталость* — ощущение физической усталости и отсутствия энергии, которое нарушает повседневную физическую и социальную жизнь, не связанное с умственным переутомлением, депрессией, сонливостью, нарушением двигательных функций. Выделяют два феномена: физическую усталость, которая заключается в снижении способности поддержания физической активности, и психическую, которая заключается в уменьшении способности выполнения умственных задач. [5]

Диагностирование усталости у пациента затруднительно в силу отсутствия общепринятой симптоматики и определения синдрома. На момент 2021 года объективных критериев усталости не выделено. Также отсутствуют инструментальные методы оценки усталости. [5]

### 1.1.2 Хроническая усталость

*Синдром хронической усталости (СХУ)* — это инвалидизирующее клиническое состояние, которое характеризуется стойкой усталостью после физических нагрузок, и сопровождающееся симптомами, связанными с когнитивной, иммунологической, эндокринологической и автономной дисфункцией. [6]

Согласно исследованиям Института медицины США, от 836 000 до 2.5 миллионов американцев страдают СХУ. Данный факт приводит к финансовым затратам от 17 до 24 миллиардов долларов в год, а каждая семья — 20 000 долларов. Уровень безработицы среди болеющих составляет от 35% до 69%. [7]

Наличие синдрома у пациента определяется наличием необъяснимой стойкой хронической усталости. К симптомам заболевания относят [6]:

- нарушение памяти или концентрации внимания;
- фарингит;
- болезненные при пальпации шейные или подмышечные лимфоузлы;
- болезненность или скованность мышц;
- болезненность суставов;
- вновь возникающая головная боль или изменение ее характеристик;

- сон, не приносящий ощущения восстановления;
- усугубление усталости, продолжающееся более 24 часов.

Более трех одновременно диагностируемых симптомов, сохраняющихся в течение более 5 месяцев, определяют наличие СХУ. [6]

### **1.1.3 Стресс**

*Стресс* — это реакция человеческого организма на неблагоприятные воздействия внешней среды, которая носит психофизиологический характер. К подобным неблагоприятным воздействиям можно определить: конфликтные ситуации, потеря близких людей, насилие, несправедливость, ограничение потребностей. [8]

Стрессовая реакция заключается в активации парасимпатического и симпатического отдела вегетативной нервной системы, в следствие чего стимулируется одна из трех психофизиологических осей стресса. [8]

Активация симпатической нервной системы заключается в общем возбуждении внутреннего органа, осуществляющего реагирование на стресс, а активация парасимпатического характера проявляется в виде чрезмерного возбуждения внутреннего органа, сменяющегося на торможение, замедление или нормализацию. Проявление симпатической реакции, в случае разрешения стрессовой ситуации, может смениться парасимпатической заторможенностью, возвращающей органы к нормальному функционированию. [8]

При продолжении течения стрессовой реакции задействуется нейроэндокринный процесс, который активирует все ресурсы организма: выделение в кровь адреналина и норадреналина, выброс гормонов. Данная реакция приводит к большому возбуждению, увеличению сердечного выброса, подъему артериального давления, снижению кровотока к почкам, повышению уровня жирных кислот в плазме и увеличению уровня холестерина. [8]

### **1.1.4 Стадии общего адаптационного синдрома**

*Общий адаптационный синдром* — это сочетание стереотипных реакций, возникающих в организме в ответ на действие стрессоров и обеспечивающих ему устойчивость не только к стрессорному агенту, но и по отношению к другим болезнетворным факторам. [9]

Общий адаптационный синдром протекает в три стадии: тревоги, устойчивости и истощения.

Стадия тревоги представляет собой мобилизацию ресурсов организма.

Она имеет продолжительность от 6 до 48 часов и включает в себя стадии шока и противошока. Стадия противошока характеризуется мобилизацией основных функциональных систем организма: нервной, симпато-адреналовой, эндокринной и адreno-кортиктропной. [9]

При длительном воздействии стрессора на организм наступает стадия устойчивости (адаптации). На данной стадии организм становится более устойчивым к действию раздражителя и другим патогенным факторам, повышается образование и секреция глюкокортикоидов. Поддерживается состояние гомеостаза в присутствии стрессора. [9]

При длительном воздействии стрессора адаптивные механизмы, участвующие в поддержании резистентности, истощаются, и наступает стадия истощения организма. Данная стадия не является обязательной и стрессовая ситуация может быть окончена до ее наступления. Протекание реакции заключается в активизации эндокринной системы и обеднении коры надпочечников. [9]

В стрессовой ситуации важно выявить наступление стадии истощения организма для его дальнейшего плодотворного функционирования. Подобные действия способны позволить корректировать нагрузку на человека, а также эффективно управлять его активностью.

### **1.1.5 Профессиональный стресс**

Профессиональный стресс возникает при воздействии на работника эмоционально-отрицательных или экстремальных факторов окружающей среды при выполнении профессиональной деятельности. Подобные стрессы влияют на здоровье человека и производительность труда, в особенности на сотрудников организаций, которые работают в условиях систематических рабочих и информационных перегрузок. [10]

Среди причин возникновения профессионального стресса различают следующие организационные факторы в трудовой деятельности человека [10]:

- режим трудовой деятельности (большая рабочая нагрузка, функциональная перегрузка, сроки выполнения);
- нерациональная организация труда и рабочего места (нечеткие ограничения полномочий и обязанностей, неоднозначные требования к выполняемой работе);
- неудовлетворительные условия труда персонала (монотонная рабо-



та, отсутствие информационных и материальных ресурсов);

- отсутствие эффективной системы мотивации и стимулирования (недостаточное вознаграждение за труд, риск штрафных санкций противоречие интересов работников и их функциональных обязанностей);
- неэффективный стиль управления,
- отсутствие или недостаток поддержки со стороны начальства;
- неудовлетворительная социально-психологическая атмосфера (неблагоприятный социально-психологический климат в коллективе, отсутствие или недостаток поддержки со стороны коллег, нарушение внутригрупповых норм поведения);
- лишение перспектив карьерного роста, бесперспективность работы.

К последствиям профессионального стресса относят [10]:

- низкую эффективность работы;
- апатичность;
- консерватизм;
- пессимистичность;
- избегание коммуникации на работе;
- физическое истощение, болезненность.

Профилактика профессиональных стрессов может включать в себя [10]:

- рациональное распределение рабочей нагрузки;
- четкое описание должностных обязанностей, полномочий;
- создание комфортной социально-психологической среды в трудовом коллективе и организации, устранение конфликтов;
- предоставление возможности карьерного роста сотрудников в рамках организации;
- предоставление информации о системе оценки труда и системе мотивации трудовой деятельности, поощрение лучших сотрудников.

Последствия профессиональных стрессов носят как физиологический, так и психологический характер. Предупреждение наступления стадии истощения с использованием информации, поступающей от пользователя, позволит сохранить здоровье и работоспособность трудящихся, а также решить проблему повышения эффективности работы и заболеваемости на высоконагруженных трудовых местах.

## **1.2 Формализация цели метода систематического распознавания усталости на автоматизированном рабочем месте**

В стрессовой ситуации важно выявить наступление стадии истощения организма для его дальнейшего плодотворного функционирования. Подобные действия способны позволить корректировать нагрузку на человека, а также эффективно управлять его активностью.

Для предупреждения наступления стадии истощения требуется определить момент наступления стадии тревоги или стадии тревоги и затем устойчивости. Данные наблюдения позволят определить, продолжается ли воздействие стрессора на организм. В случае возвращения характеристик пользователя к стандартным, построенным на наблюдении вне периода стресса, дальнейших действий не потребуется. Предупреждение о наступлении стадии истощения предполагается в случае, когда характеристики пользователя, находящегося в установленной стадии устойчивости, устремляются к нестандартным значениям, что говорит об активизации эндокринной системы, угнетающей функции внутренних органов.

Согласно исследованиям, проведенными Федеральной службой государственной статистики [11], 71.2% от общей численности населения в городской местности используют персональные компьютеры дома, в то время как 33.9% — на рабочих местах.

Подобная статистика предполагает возможность использования в качестве данных о состоянии пользователя для предупреждения состояния усталости характеристик, получаемых от внешних устройств, как на рабочих местах, так и в домашних условиях.

## **1.3 Методы определения усталости оператора с использованием устройств автоматизированного рабочего места**

*Внешние устройства (периферийные устройства)* — устройства, предназначенные для внешней машинной обработки информации (в отличие от преобразований информации, осуществляемых центральным процессором) [12]. По роду выполняемых операций данные устройства подразделяются на группы [12]:

- устройства подготовки данных (занесения информации на промежуточные носители данных);
- устройства ввода (считывание информации и её преобразование в

кодovou последовательность электрических сигналов, подлежащих передаче в центральный процессор;

- устройства вывода (регистрация результатов обработки информации или их отображения);
- устройства хранения больших объемов информации;
- устройства передачи информации на большие расстояния.

К внешним устройствам, которые являются устройствами ввода, то есть органами управления персональным компьютером, относят:

- клавиатуру;
- мышь;
- графический планшет;
- веб-камеру;
- микрофон;
- игровой манипулятор.

Из указанных выше устройств в рассмотрение не войдут:

- графический планшет;
- игровой манипулятор.

Данное решение связано с тем, что использование подобного рода устройств указывает на особый род работы. В требованиях к разрабатываемой системе указывается возможность ее использования для операторов автоматизированных рабочих мест в наиболее распространенных конфигурациях (например, для офиса), поэтому использование в решении поставленной задачи информации, получаемой с графических планшетов и игровых манипуляторов, не является целесообразным.

Согласно проведенным исследованиям [13] признаки голоса, клавиатурного почерка и характера работы исследуемого или контролируемого субъекта с компьютерной мышью содержат информацию о психофизиологических состояниях оператора: нормальное, усталость, опьянение, возбужденное, ослабленное (сонное).

### **1.3.1 Метод анализа клавиатурного почерка**

*Клавиатурный почерк* — это подвид поведенческой подгруппы аутентификации по неотчуждаемым признакам. [14]

Клавиатурный почерк определяется по времени между нажатиями клавиш. При снятии биометрического шаблона клавиатурного почерка измеряют

время нажатия двух, трех или четырех последовательных клавиш, сохраняют его и на основе полученных значений строят математические модели для сравнения шаблонов нескольких пользователей. [15]

Для системы входными данными будут два биометрических шаблона — эталонного и кандидата. Результат работы системы — рейтинг доверия к биометрическому шаблону кандидата, который является критерием схожести двух переданных шаблонов.

Различают два вида распознавания клавиатурного почерка: распознавание на статическом тексте (пароль или известная кодовая фраза), распознавание при вводе псевдослучайного текста. [14]

Математическая задача распознавания на фиксированном тексте может быть формализована. Для ее выполнения потребуется [14]:

- собрать информацию о времени между нажатиями соседних клавиш в тексте;
- сформировать из полученных данных вектор фиксированной размерности;
- по кластерной модели или другой модели сравнения двух векторов сравнить сформированный вектор и вектор-эталон для этого же текста от этого же пользователя.

Для задачи распознавания клавиатурного почерка при вводе псевдослучайного текста не существует надежных моделей формирования пользовательского шаблона и вычисления рейтинга. Время работы подобных систем не позволяет в реальном времени оценить ситуацию и выдает результат через десятки минут, а память, занимаемая векторами, лишь продолжает расти. [14]

Одним из численных показателей, которые определяют качество биометрической системы, является *ошибка первого рода (FRR, количество ложноотрицательных)* — это вероятность ложного отказа в доступе. Данная ошибка имеет место при возникновении повреждения рук пользователя или ненормального психофизического состояния человека (усталость, алкогольное или наркотическое опьянение, приступ гнева).[14]

*Ошибка второго рода (FAR, количество ложноположительных)* — это вероятность ложного допуска. Данная ошибка имеет место при ситуации, когда заданные возможные отклонения от допустимых значений при распознавании пользователя были заданы неверно, либо же когда нарушитель сумел скопиро-

вать метрики поведения пользователя и обойти систему контроля. [14]

Для проектируемой системы важнейшую роль играет ошибка первого рода. Данная ошибка способна позволить распознать состояние усталости оператора в случае единоличного пользования системой.

### **1.3.2 Метод анализа скорости печати и количества ошибок**

Изменение психофизического состояния пользователя приводит к увеличению времени совершения простейших действий. Контроль времени выполнения данных действий способен помочь определить ухудшение состояния оператора автоматизированного рабочего места и своевременно сообщить пользователю о необходимости соблюдения гигиены труда. К контролю действий оператора относят [16]:

- скорость печати (ввода) — количество символов, введенных пользователем, разделенное на время, за которое данные символы были введены — уменьшение количества символов, введенного в единицу времени, может символизировать ухудшение состояния пользователя;
- динамика печати (ввода) — время между нажатиями клавиш и временем их удержания — увеличение интервала между нажатиями может символизировать ухудшение состояния пользователя;
- частота возникновения ошибок при печати (вводе) — увеличение количества совершаемых ошибок может символизировать ухудшение состояния пользователя.

### **1.3.3 Метод анализа траекторий перемещения курсора мыши**

Компьютерная мышь используется для взаимодействия с оконным интерфейсом операционной системы и программ.

Особенности работы с мышью можно оценить, анализируя траектории передвижения курсора мыши по экрану между элементами интерфейса и среднее время выполнения передвижения. [13]

Оценка среднего времени перемещения курсора мыши между элементами интерфейса может быть выполнена с использованием адаптированным для данной задачи законом Фиттса [17]:

$$T = b \cdot \log_2 \left( \frac{D}{W} + 1 \right), \quad (1)$$

где  $b$  — величина, зависящая от типичной скорости движения курсора

мышь (отношение средней скорости движения мыши по экрану, осуществляемого субъектом, к установленному в операционной системе коэффициенту чувствительности мыши);

$D$  — дистанция перемещения курсора между элементами интерфейса ( в пикселях);

$W$  — ширина элемента интерфейса, к которому направляется курсор (в пикселях).

Адаптированный закон Фиттса связывает время, затраченное на движение к цели, с точностью движения и расстоянием до цели. Причем время перемещения курсора не совпадает с оценкой, выраженной в формуле (1), и отличается на некоторую величину, которую используют в качестве идентифицирующего признака. [18]

В качестве признаков могут быть использованы амплитуды первых десяти низкочастотных гармоник функции скорости перемещения курсора мыши по экрану:

$$V_{xy}(t) = \sqrt{\left((x(t_{i+1}) - x(t_i))^2 + (y(t_{i+1}) - y(t_i))^2\right)^2}, \quad (2)$$

где  $x, y$  — координаты курсора;

$t_i$  —  $i$ -ый момент времени регистрации координат курсора (регистрация координат курсора зависит от производительности компьютера).

Разложение функции (2) производится с помощью быстрого преобразования Фурье, тем самым достигается нормирование участков пути курсора по времени [18]. Каждый участок приводится к длительности в 0.5 секунд, амплитуды нормируются по энергии функции  $V_{xy}(t)$ , вычисляемой в соответствии с формулой:

$$E_s = \int_{-\infty}^{\infty} A^2(\omega) dt, \quad (3)$$

где  $A(\omega)$  — амплитуды гармоники с частотой  $\omega$  функции  $V_{xy}(t)$ .

Операция нормирования приводит траектории перемещений курсора между элементами интерфейса к единому масштабу. Аналогичные операции осуществляются по отношению к функциям координат курсора  $x(t)$  и  $y(t)$  с выпол-

нением предварительного перевода данных функций в систему координат, где точкой начала координат является центр элемента интерфейса, к которому двигался курсор, причем ось абсцисс будет расположена в направлении к центру данного элемента. Данные действия обусловлены необходимостью избавиться от наклона линий, связывающих элементы интерфейса, относительно исходной координатной плоскости. [18]

#### **1.3.4 Метод анализа использования координатного устройства мышь с использованием модели искусственной нейронной сети**

В качестве главных факторов биометрической информации в данном методе выбираются параметры применения клавиатуры и мыши. [19]

Для сбора и хранения особенностей ввода управляющих последовательностей применяется логирование клавиатурного буфера машины. При любом использовании клавиатуры сохраняются задержки между нажатиями, продолжительности нажима отдельных клавиш, а также продолжительность набора устойчивого сочетания. В качестве сохраняемой информации об использовании мыши и иных координатных устройств фиксируются изменения положений, скорости и времени передвижения курсора. [19]

Анализ сохраненных данных проводится с использованием специальной искусственной нейронной сети. [19]

Основной задачей нейронной сети в данном методе является разбиение поступающих данных на два кластера: случаи санкционированного доступа и несанкционированного доступа. При этом исследуемый показатель не является статическим, в связи с чем нейронная сеть работает в динамическом режиме, переобучаясь по мере изменения поведенческих особенностей пользователя. [19]

После обучения нейронной сети, с ее помощью исследуется поток данных о работе текущего пользователя. Из-за возможности изменения поведения ввода человека, с некоторой долей вероятности модель может не может однозначно определить правомерность доступа. В подобных ситуациях требуется принять решение о том, следует ли переобучить модель в связи с изменением анализируемых параметров или запретить доступ к системе. Для решения данной проблемы используется метод нечеткого вывода Мамдани. [19]

### **1.3.5 Метод анализа внешнего состояния пользователя**

Основные исследования в области использования видео-изображений для определения опасных состояний усталости проводятся для реализации систем распознавания усталости водителя.

Признаки состояний ослабленного внимания и усталости у водителя характеризуется следующими наблюдаемыми параметрами [20]:

- поворот головы влево/вправо по отношению к туловищу;
- наклон головы вперед относительно туловища;
- продолжительность моргания век;
- частота моргания век;
- степень открытости рта человека (признаки зевоты).

Работа системы строится на использовании технологий машинного обучения и предварительно обученных на наборах данных моделей. Детектирование опасных состояний в поведении повышает точность и полноту определения ослабленного внимания и усталости — накопление, анализ и обработка информации, поступающей в облачный сервис и формирующей историю вождения, приводят к созданию новых моделей поведения того или иного пользователя, вычислять характерные параметры, оказывающие влияние на эффективность работы схем распознавания опасных состояний. Таким образом высчитываются не только пороговые параметры для каждого водителя индивидуально, но и поддерживается надежность обновляемых параметров для новых пользователей, чьи параметры пока не известны. [20]

Применение рассматриваемого метода, в комплексе с уже рассмотренными, для определения состояния усталости оператора автоматизированного рабочего места потребует больших вычислительных ресурсов серверной части программного комплекса. Кроме того, со стороны клиента потребуются непрерывное соединение с сетью Интернет, увеличение ресурса оперативной памяти и современная веб-камера с приемлемым качеством фото- и видео-съемки.

### **1.3.6 Метод анализа речевых характеристик пользователя**

Микрофон позволяет регистрировать аудиопоток, исходящий от пользователя и его окружения.

Исследования показали, что признаки голоса наилучшим образом позволяют распознавать усталость или расслабленное (сонное) состояние диктора. [13]



Симптоматика усталости заключается в проявлениях слабости, раздражительности, расстройствах сна и вегетативных нарушениях. [21]

Степень раздражительности может быть оценена по шкале Рэймонда Новако. Тест Новако содержит в себе 25 вопросов, ответы на которые позволяют определить показатель раздражительности. При интерпретации результатов показатель в 0–45 баллов определяет низкую степень раздражительности, 46–55 баллов — раздражительность ниже средней, 56–75 баллов — среднюю степень раздражительности, 76–85 баллов — раздражительность выше средней и 86–100 баллов — высокую степень раздражительности. [22]

Проявления слабости, расстройства сна и вегетативные нарушения могут быть установлены с использованием клинических методов.

Наиболее полную информацию о внутреннем психоэмоциональном состоянии человека может дать анализ его связной речи [23]:

- расстановка логических ударений;
- скорость произнесения слов;
- конструкция фразы;
- неуверенный или неверный подбор слов;
- обрывание фраз на полуслове;
- изменение слов;
- появление слов-паразитов;
- исчезновение пауз.

Разработки в сфере анализа напряжения голоса тестируются в судебной психологии для выявления лжи и обмана посредством обнаружения микротремора. Однако целесообразность и надежность использования данного метода до сих пор является предметом дискуссий в силу того, что успех в выявлении лжи зависит от опыта эксперта. Кроме того, было доказано, что невиновные люди, чья невиновность оспаривается, проявляют не меньше стресса, чем виновные, увеличивая риск появления ложных срабатываний. Данные проблемы в практике привели к тому, что на сегодняшний день данные разработки рассматриваются как многообещающий инструмент обнаружения стресса, однако проблема межиндивидуальных различий пока еще остается нерешенной. [24]

### **1.3.7 Метод анализа виброакустических шумов при наборе текста или использовании мыши**

Данный метод предполагает использование виброакустических датчиков, установленных непосредственно на рабочем столе пользователя, данные с которых вводятся в звуковую карту. [25]

Для идентификации оператора из записей виброакустического сигнала при наборе произвольного текста удаляются паузы между нажатиями и отпусканиями клавиш клавиатуры. Затем очищенный исследуемый сигнал разбивается на перекрывающиеся интервалы, длины которых определяются идентификацией оператора с минимальной погрешностью. Основное требование для векторов идентификации — устойчивость для всех пользователей в системе распознавания. Таким образом формируется база виброакустических сигналов, возникающих при наборе каждым из пользователей системы. [25]

Для идентификации оператора с использованием манипулятора мышь в качестве векторов признаков используются относительные длительности периодов активности пользователя, которые вычисляются на участках виброакустического сигнала, сопоставленных с записями в журнале событий мыши. [26]

Для каждого оператора обучается нейронная сеть для его идентификации и затем используется при работе системы. [25]

В качестве оптимизации формирования обучающей выборки предлагается использование алгоритма выделения фрагментов, которые содержат виброакустический сигнал нажатия или отпускания клавиш на основе ошибки линейного предсказания и последующего удаления единичных максимумов. [25]

## **1.4 Биофизические факторы, позволяющие определить усталость**

### **1.4.1 Частота пульса и возраст сосудистой системы**

Согласно исследованиям [27], при кратковременном стрессовом воздействии было выявлено три показателя, которые подвергаются изменениям: частота пульса, возраст сосудистой системы и индекс стресса.

Частота пульса, в среднем, до стрессового воздействия составляла  $77.42 \pm 4.12$  ударов в минуту, после —  $99.67 \pm 5.54$  ударов в минуту с уровнем достоверности  $p = 0.004$ . [27]

*Возраст сосудистой системы* — это параметр, определяющий биологический возраст индивида, то есть изношенность его организма. [28]

Данный параметр, в среднем, до стрессового воздействия был равен  $29 \pm$

1.9 лет, после —  $34.5 \pm 1.69$  лет с уровнем достоверности  $p = 0.041$ . [27]

#### 1.4.2 Индекс Баевского

*Индекс стресса (индекс Баевского)* характеризует вариабельность сердечного ритма и состояние центров регуляции сердечно-сосудистой системы. Норма данного индекса для человека находится в диапазоне от 50 до 150 единиц. Увеличение данного показателя до значений от 150 до 500 может говорить о наличии физических нагрузок или хронической усталости. Увеличение индекса до значений в диапазоне от 500 до 900 говорит о наличии существенно психологического и эмоционального стресса, либо о психофизиологическом переутомлении или стенокардии. Превышение индексом значения 900 единиц свидетельствует о нарушении регуляторных механизмов или предынфарктном состоянии. [29]

Индекс Баевского может быть рассчитан по следующей формуле [30]:

$$I = \frac{AM_o}{2M_o\Delta X}, \quad (4)$$

где  $M_o$  — это наиболее часто встречающееся в динамическом ряде значение кардиоинтервала (мода);

$AM_o$  — это число кардиоинтервалов, соответствующих значению моды (амплитуда моды), в процентах от объема выборки;

$\Delta X$  — вариационный размах среднего значения продолжительности сердечного цикла, показатель деятельности контура автономной регуляции ритма сердца, который целиком связан с дыхательными колебаниями тонуса блуждающих нервов, в секундах.

$M_o\Delta X$  — отражает степень вариативности значений кардиоинтервалов в исследуемом динамическом ряду.

В исследовании данный параметр, в среднем, до стрессового воздействия был равен  $75.27 \pm 11.84$  единиц, после —  $334.55$  единиц с уровнем достоверности  $p = 0.001$ .

Частота сердечных сокращений является включением для индекса стресса, из чего следует факт того, что в системе отсутствует потребность в анализе данной характеристики отдельно. Возраст сосудистой системы — параметр, определяемый сложной медицинской техникой. В силу изложенных фактов, в дальнейшем в систему в качестве характеристики для определения стресса вой-

дет лишь индекс Баевского.

В качестве устройства для снятия индекса, основанного на вариабельности сердечного ритма, могут быть предложены смарт-часы с поддержкой данной функции. Модуль программного обеспечения, сохраняющий данные характеристики, должен записывать ее значения в базу данных с использованием интерфейса, предоставляемым выбранной моделью.

### **1.5 Формализация требований к методу систематического распознавания усталости на автоматизированном рабочем месте**

На текущий момент не каждое автоматизированное рабочее место может включать в себя все рассмотренные внешние устройства взаимодействия пользователя с системой. Данное обстоятельство указывает на требование возможности работы системы в условиях отсутствия тех или иных периферийных устройств.

К требованиям также отнесена потребность в сохранении ресурсов персонального компьютера автоматизированного рабочего места в силу отсутствия стандартизации конфигурации технической составляющей компьютера.

Важно отметить, что в системе требуется разграничение личностей пользователей в силу индивидуальности характеристик каждого.

### **Вывод**

## **2. Конструкторский раздел**

В данном разделе представлены диаграмма вариантов использования, диаграмма "сущность-связь" базы данных для хранения характеристик пользователя и IDEF0-диаграмма прикладной задачи определения усталости оператора АРМ.

### **2.1 Включаемые в метод характеристики, получаемые с использованием периферийных устройств**

При проведении анализа метода распознавания усталости с использованием веб-камеры было указано, что применение данного метода может потребовать больших вычислительных ресурсов серверной части программного комплекса, а также постоянное соединение с сетью Интернет со стороны клиента. Данные факторы указывают на то, что метод не отвечает формализованным требованиям к прототипу метода систематического распознавания усталости на автоматизированном рабочем месте.

При проведении анализа метода распознавания усталости и стресса с использованием микрофона стало известно, что данный метод позволяет определить лишь проявления слабости, которые не могут однозначно указывать на наступление стадии истощения, адаптации или тревоги. Таким образом, метод не может быть включен в прототип метода систематического распознавания усталости на автоматизированном рабочем месте.

Для определения усталости будут использоваться устройства клавиатура и мышь. Для решения задачи будет построена нечеткая модель.

### **2.2 Метод систематического распознавания усталости на автоматизированном рабочем месте**

Метод систематического распознавания усталости на автоматизированном рабочем месте включает в себя:

- хранение и анализ данных, получаемых от клавиатуры — для определения усталости используется нечеткая модель;
- хранение и анализ данных, получаемых от мыши — для определения усталости используется нечеткая модель.

Построение нечеткой модели происходит на этапе синхронизации поступающих данных о действиях пользователя с данными о скорости его реакции в отдельные моменты времени. В дальнейшем данная модель используется до

ее актуализации. Система актуализирует модель в случае, если будет определено, что имеют место ложные срабатывания и “эталонные” паттерны поведения субъекта изменились.

### **2.3 Формат и метод сбора данных, предоставляемых оператором автоматизированного рабочего места**

В качестве данных, характеризующих действия оператора с клавиатурой и мышью, выступают нажатия клавиш клавиатуры и мыши, а также передвижения курсора мыши по экрану.

Каждое нажатие на клавишу клавиатуры характеризуется двумя полями:

- наименование нажатой клавиши (в случае литеры — литера, в случае специальных клавиш — полное наименование);
- временная метка, включающая в себя дату совершенного действия, и время в формате ЧЧ:ММ:СС.ССС, где Ч — часы, М — минуты, С — секунды.

Каждое нажатие на клавишу мыши характеризуется четырьмя полями:

- координата X, в которой было совершено действие;
- координата Y, в которой было совершено действие;
- номер нажатой кнопки мыши;
- временная метка, включающая в себя дату совершенного действия, и время в формате ЧЧ:ММ:СС.ССС.

Передвижения курсора мыши характеризуется тремя полями:

- координата X нового положения, в котором находится курсор;
- координата Y нового положения, в котором находится курсор;
- временная метка, включающая в себя дату совершенного действия, и время в формате ЧЧ:ММ:СС.ССС.

При непрерывном передвижения курсора мыши по экрану его положения фиксируются каждые  $\approx 0.010$  секунд.

В качестве данных, формирующих нечеткую модель, используется тест на реакцию. Данный тест включает в себя появление некоторого элемента на экране и регистрацию времени, за которое оператор среагировал на его появление и нажал на него. При следующем появлении элемент не изменяет своей формы или положения, время появления элемента определяется случайным образом в интервале от двух до десяти секунд. Объекту предоставляется порядка десяти попыток, в результате чего среднее время реакции принимается за те-

кущий показатель времени реакции пользователя. Тест проводится каждые 40 минут работы пользователя.

Может быть использован как локальный, так и удаленный метод сбора данных. Отличие данных методов заключается только в расположении хранилища данных. Первично данные должны записываться в логирующие файлы, которые в дальнейшем могут быть направлены в базу данных и удалены из локального хранилища.

Метод включает в себя логирование прерываний, поступающих с клавиатуры и мыши, включающих необходимую информацию о действиях пользователя, а также результатов тестов на реакцию.

## 2.4 Диаграмма вариантов использования

На рисунке 2.1 предоставлена диаграмма вариантов использования.

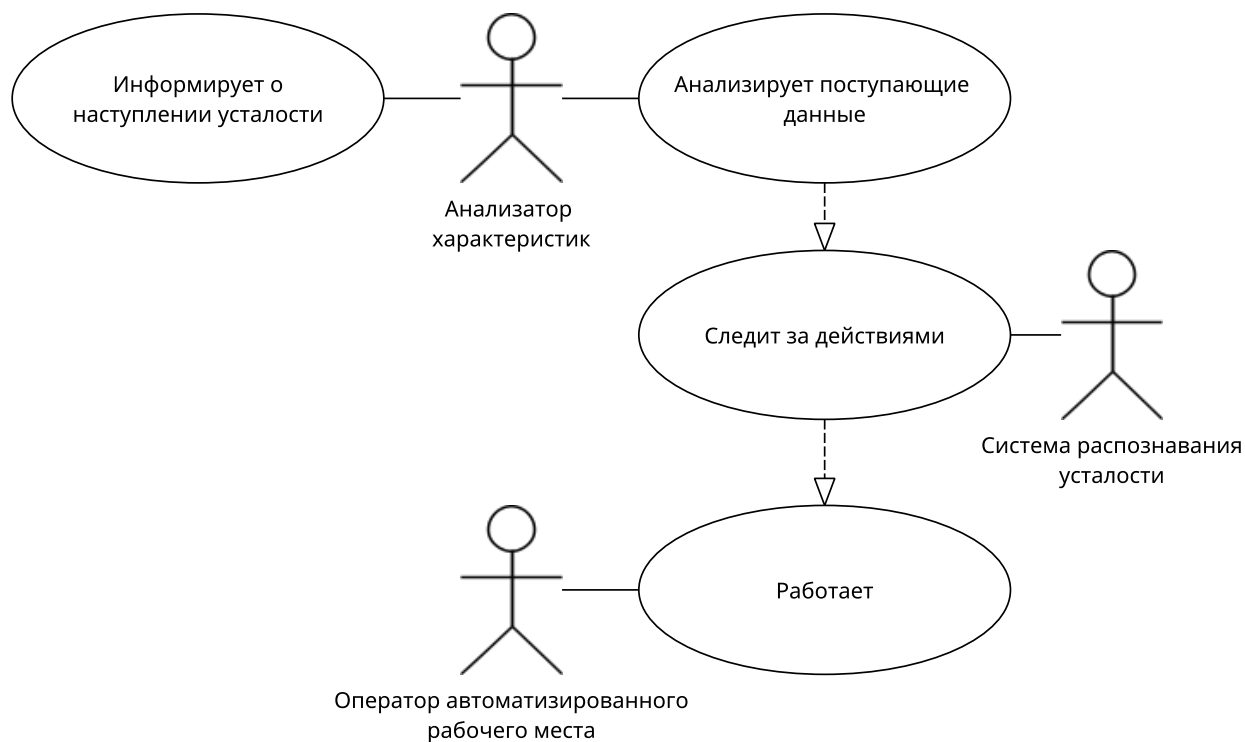


Рис. 2.1: Диаграмма вариантов использования.

В системе определены 3 роли: система распознавания усталости, анализатор характеристик и оператор автоматизированного рабочего места.

Роль системы распознавания усталости заключается в сборе и систематизации поступающей информации от оператора. Данная часть системы не анализирует информацию и определяется в качестве медиатора для базы данных и программного обеспечения снятия данных.

Анализатор поступающих характеристик формирует некоторую нечеткую модель по первичным данным, принимаемых за эталонных, а также отвечает за актуализацию участвующей в распознавании усталости модели. Также данная часть системы отвечает за информирование о наступлении усталости.

## 2.5 IDEF0-диаграмма

На рисунках 2.2–2.3 предоставлена диаграмма IDEF0 задачи определения усталости оператора АРМ.

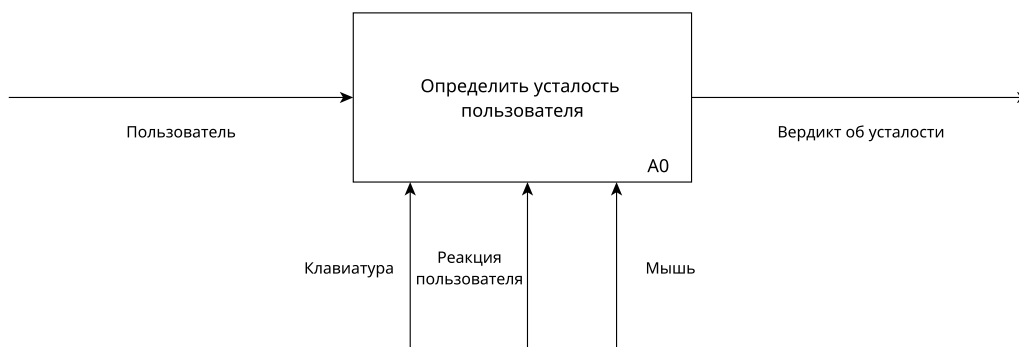


Рис. 2.2: IDEF0-диаграмма уровня A0.



Рис. 2.3: IDEF0-диаграмма уровня A1-A3.



## 2.6 Диаграмма “сущность-связь” в нотации Чена

На рисунке 2.4 предоставлена диаграмма “сущность-связь” в нотации Чена.

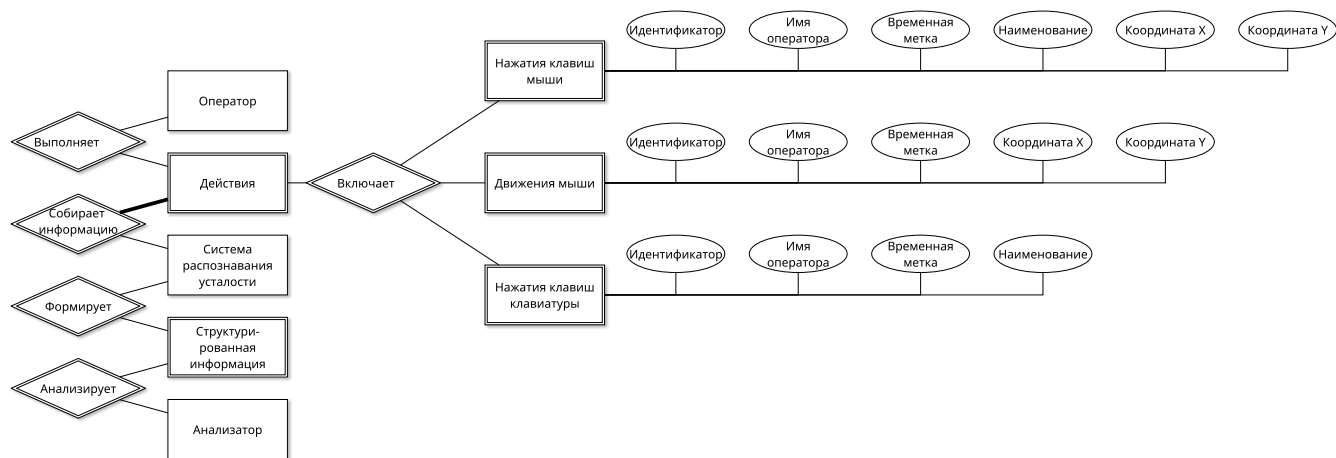


Рис. 2.4: Диаграмма “сущность-связь” в нотации Чена.

## 2.7 Используемые методы исследования корреляции

Корреляционные методы позволяют решать задачи определения изменения зависимой переменной под влиянием одного или набора факторов, установкой тесноты связи результативного признака с отдельным фактором, включенным в анализ. Также данные методы предоставляют возможность оценки общего объема вариации зависимой переменной и определения роли каждого фактора, а также проведения статистической оценки выборочных показателей корреляционной связи. [31]

Характеристикой корреляционной зависимости является статистическая величина, называемая коэффициентом корреляции. [32]

### 2.7.1 Парная корреляция

Для парной корреляции (фактор и отклик имеют нормальные распределения), коэффициент корреляции вычисляется по формуле [32]:

$$\rho = \frac{M \{[x - M(X)][y - M(Y)]\}}{\sqrt{M[x - M(X)]^2} M[y - M(Y)]^2}} = \frac{K_{XY}}{\sqrt{D(X) D(Y)}}, \quad (5)$$

где  $K_{XY}$  — корреляционный момент, представляющий собой математическое ожидание произведения отклонений значений  $x$  и  $y$  случайных величин  $X$  и  $Y$  от их математических ожиданий  $M(X)$

и  $M(Y)$ ;

$D(X)$  — дисперсия случайной величины  $X$ ;

$D(Y)$  — дисперсия случайной величины  $Y$ .

На практике случайные величины представляются ограниченным числом значений, поэтому вместо значения коэффициента корреляции  $\rho$  используется его оценка  $r$ , которая рассчитывается с использованием выборочных характеристик отклика и фактора [32]:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)\bar{S}_X\bar{S}_Y}, \quad (6)$$

где  $\bar{x}$  и  $\bar{y}$  — средние выборочные значения фактора и отклика;

$\bar{S}_X$  и  $\bar{S}_Y$  — выборочные стандартные отклонения отклика и фактора;

$n$  — число наблюдений.

Следует отметить, что значение  $r$  лежит на интервале от  $-1$  до  $+1$ . В случае, когда  $r = \pm 1$  зависимость между фактором и откликом является функциональной. Также положительное значение коэффициента корреляции указывает на возрастание отклика с увеличением фактора, когда отрицательное значение свидетельствует об убывании  $Y$  при возрастании  $X$ . Равенство коэффициента парной корреляции нулю указывает, что взаимосвязь не является линейной. [32]

### 2.7.2 Множественная корреляция

Множественная корреляция определяет обусловленность некоторого признака одновременным действием нескольких других признаков. Взаимодействия отклика с каждым из факторов и факторов между собой отображают в виде матрицы корреляции. [32]

Таблица 1: Матрица корреляции

	$Y$	$X_1$	$\dots$	$X_j$	$\dots$	$X_m$
$Y$	1	$r_{Y,X_1}$	$\dots$	$r_{Y,X_j}$	$\dots$	$r_{Y,X_m}$
$X_1$	$r_{Y,X_1}$	1	$\dots$	$r_{X_1,X_j}$	$\dots$	$r_{X_1,X_m}$
$\dots$	$\dots$	$\dots$	1	$\dots$	$\dots$	$\dots$
$X_j$	$r_{Y,X_j}$	$r_{X_1,X_j}$	$\dots$	1	$\dots$	$r_{X_m,X_m}$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	1	$\dots$
$X_m$	$r_{Y,X_m}$	$r_{X_1,X_m}$	$\dots$	$r_{X_j,X_m}$	$\dots$	1

Коэффициент множественной корреляции определяют из предположе-

ния, что отклик связан с факторами линейной зависимостью, причём коэффициент вычисляется по следующей формуле:

$$R = \sqrt{1 - \frac{\Delta_{YX}}{\Delta_{XX}}}, \quad (7)$$

где  $\Delta_{YX}$  — определитель матрицы корреляции;

$\Delta_{XX}$  — определитель матрицы, получаемой из матрицы корреляции вычеркиванием первой строки и первого столбца.

Значимость множественного коэффициента корреляции проверяется с использованием критерия Фишера:

$$F_\rho = \frac{R^2}{1 - R^2} \frac{n - m - 2}{m} > F[\alpha; m; n - m - 2], \quad (8)$$

где  $F_\rho$  и  $F[\alpha; m; n - m - 2]$  — рассчитанное и табличное числа Фишера.

При выполнении условия (8) коэффициент множественной корреляции можно считать значимым с доверительной вероятностью  $p = 1 - \alpha$ .

## 2.8 Используемые методы дисперсионного анализа

Дисперсионный анализ позволяет определить, влияет ли некоторый фактор или группа факторов на изучаемую величину, какой из них имеет наибольшее влияние, а также зависит ли влияние факторов от их взаимодействия друг с другом. [33]

### 2.8.1 Однофакторный дисперсионный анализ

Некоторый фактор  $A$  имеет  $m$  уровней и число  $y_{ij}$  получено в результате  $j$ -го опыта, проведенного на его  $i$ -м уровне. Числа  $y_{ij}$  — это наблюдения, а  $n_i$  — число наблюдений, полученных на  $i$ -м уровне. Наблюдения представляются в виде [33]:

$$y_{ij} = \mu_i + \varepsilon_{ij}, \quad (9)$$

где  $\mu_i$  — математическое ожидание  $y$  на  $i$ -ом уровне;

$\varepsilon_{ij}$  — случайная ошибка.

Если фактор не влияет на переменную  $y$ , то рассеяние ее значений вызывается лишь случайными ошибками, а математические ожидания на всех уровнях одинаковы. [33]

Обозначим  $\alpha_i = \mu_i - \mu$ ,  $\mu = \frac{1}{n} \sum_{i=1}^m n_i \mu_i$ . Значение  $\alpha$  называется эффек-

том фактора  $A$  на уровне  $i$ . Таким образом уравнение (10) примет вид [33]:

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad (10)$$

Следующие рассуждения строятся на предположении о том, что случайные ошибки имеют нулевое математическое ожидание, постоянную дисперсию и подчиняются нормальному распределению.

Определяются следующие величины:

$$y_{ig} = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}, \quad (11)$$

$$e_{ij} = y_{ij} - y_{ig}, \quad (12)$$

$$y_{gg} = \frac{1}{n} \sum_{i,j} y_{ij}, \quad (13)$$

$$\hat{\alpha}_i = y_{ig} - y_{gg}, \quad (14)$$

где  $y_{ig}$  — средние значения по столбцам;  
 $e_{ij}$  — отклонения от среднего в каждом столбце;  
 $y_{gg}$  — общее среднее,  $n = \sum_{i=1}^m n_i$ ;  
 $\hat{\alpha}_i$  — отклонения средних по столбцам от общего среднего.

Можно доказать, что:

$$M(e_{ij}) = 0, M(y_{gg}) = \mu, M(\hat{\alpha}_i) = \mu_i - \mu = \alpha_i, \mu = \frac{1}{n} \sum_{i=1}^m n_i \mu_i, \quad (15)$$

Соотношения (15) означают, что случайные величины  $y_{ig}$  и  $\hat{\alpha}_i$  являются несмещенными оценками параметров  $\mu_i$  и  $\alpha_i$ .

Справедливо соотношение:

$$S = S + S, \quad (16)$$

Причем:

$$S = \sum_{i,j} (y_{ij} - y_{gg})^2, \quad (17)$$

где  $S$  — Полная сумма квадратов.

$$S = \sum_{i=1}^m n_i (y_{ij} - y_{gg})^2 = \sum_{i=1}^m n_i \hat{\alpha}_i^2, \quad (18)$$

где  $S$  — межгрупповая сумма квадратов, характеризующая рассеяние средних по столбцам относительно общего среднего.

$$S = \sum_{i,j} (y_{ij} - y_{ig})^2, \quad (19)$$

где  $S$  — внутригрупповая сумма квадратов, характеризующая рассеяние значений  $y_{ij}$  относительно  $y_{ig}$ , то есть рассеяние внутри групп.

Если гипотеза верна и  $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$ , то величины  $\hat{\alpha}_i$  должны быть достаточно близки к нулю. Тогда вклад  $S$  и  $S$  по сравнению с  $S$  должен быть мал. Поэтому малое значение  $S$  является доводом в пользу гипотезы, а большое значение  $S$  является доводом против гипотезы. Точный метод проверки гипотезы основан на F-критерии. [33]

### 2.8.2 Двухфакторный дисперсионный анализ с однократными наблюдениями

В данном методе рассматривается возможное влияние нескольких факторов на некоторую переменную  $y$ .

При изучении факторов  $A$  и  $B$  на переменную  $y$  можно представить модель в виде:

$$y_{ij} = \mu_{ij} + \varepsilon_{ij}, i = 1, \dots, I; j = 1, \dots, J, \quad (20)$$

где  $\mu_{ij}$  — некоторые константы;

$\varepsilon_{ij}$  — случайные ошибки;

$I, J$  — число уровней факторов  $A$  и  $B$  соответственно;

$y_{ij}$  — наблюдение, полученное на  $i$ -м уровне фактора  $A$  и на  $j$ -м уровне фактора  $B$ .

Случайные ошибки также должны удовлетворять требованиям, что и в однофакторном случае.

Если между факторами отсутствует взаимодействие, то влияние одного фактора на величину  $y$  не зависит от того, на каком уровне находится другой фактор. От Таким образом уравнение (20) можно записать в следующем виде:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, i = 1, \dots, I; j = 1, \dots, J, \quad (21)$$

где  $\mu$  — общее среднее;  
 $\alpha_i$  — эффект фактора  $A$  на его  $i$ -ом уровне;  
 $\beta_j$  — эффект фактора  $B$  на его  $j$ -м уровне.

Для проверки предположения о том, что фактор  $A$  или  $B$  не влияет на переменную  $y$  потребуется проверить гипотезы о равенстве нулю соответствующих эффектов:

$$H_A : \alpha_i = 0, i = 1, \dots, I, H_B : \beta_j = 0, j = 1, \dots, J, \quad (22)$$

Для этого потребуется вычислить величины:

$$y_{ig} = \frac{1}{J} \sum_{j=1}^J y_{ij}, \quad (23)$$

где  $y_{ig}$  — средние по строкам.

$$y_{gj} = \frac{1}{I} \sum_{i=1}^I y_{ij}, \quad (24)$$

где  $y_{gj}$  — средние по столбцам.

$$y_{gg} = \frac{1}{IJ} \sum_{i,j} y_{ij}, \quad (25)$$

где  $y_{gg}$  — общее среднее.

Далее вычисляются значения  $\hat{\alpha}_i = y_{ig} - y_{gg}$  и  $\hat{\beta}_j = y_{gj} - y_{gg}$ . Эти величины являются несмещенными оценками соответствующих параметров. Если гипотезы верны, то данные оценки не должны значительно отличаться от нуля. Алгоритм проверки основан на критерии Фишера, как и в предыдущем варианте. [33]

## 2.9 Алгоритмы кластеризации

Задачей кластеризации является объединение в группы объектов, схожих по некоторому признаку. Данная задача включает в себя разбиение множества объектов, называемых кластерами. [34]

Одним из главных отличий кластеризации от классификации является

факт того, что перечень групп четко не задан и определяется в процессе работы алгоритма. [34]

Этапы кластерного анализа включают в себя [34]:

- 1) Отбор выборки объектов;
- 2) Определение множества переменных, по которым будут оцениваться объекты в выборке;
- 3) Вычисление значений меры сходства между объектами;
- 4) Применение метода кластерного анализа;
- 5) Представление результатов анализа.

Выделяют две основные классификации алгоритмов кластеризации: иерархические и плоские и четкие и нечеткие. [34]

Иерархические алгоритмы строят не одно разбиение выборки на непесекающиеся кластеры, а систему вложенных разбиений. Плоские алгоритмы строят разбиение объектов на кластеры. [34]

Четкие алгоритмы каждому объекту выборки ставят в соответствие номер кластера, то есть каждый объект определенно относится к одному из выделенных кластеров. Нечеткие алгоритмы каждому объекту ставят в соответствие набор вещественных значений, показывающих степень отношения объекта к кластерам, каждый объект относится к каждому кластеру с некоторой вероятностью. [34]

### **2.9.1 Алгоритмы иерархической кластеризации**

Среди алгоритмов иерархической кластеризации выделяются два типа: нисходящие и восходящие алгоритмы. В первом типе в начале все объекты помещаются в один кластер, который затем разбивается на все более мелкие кластеры. Восходящие алгоритмы в начале работы помещают каждый объект в отдельный кластер, а затем объединяют кластеры во все более крупные, пока все объекты выборки не будут содержаться в одно кластере. [34]

### **2.9.2 Алгоритмы квадратичной ошибки**

Задачу кластеризации можно рассматривать как построение оптимального разбиения объектов на группы. При этом оптимальность может быть определена как требование минимизации среднеквадратической ошибки разбиения [34]:

$$e^2(X, L) = \sum_{j=1}^K \sum_{i=1}^{n_j} \left\| x_i^{(j)} - c_j \right\|^2, \quad (26)$$

где  $c_j$  — “центр масс” кластера  $j$  (точка со средними значениями характеристик для данного кластера),

$K$  — количество кластеров,

$n$  — количество точек.

Алгоритмы квадратичной ошибки относятся к типу плоских алгоритмов. Самым распространенным алгоритмом этой категории является метод  $k$ -средних. Этот алгоритм строит заданное число кластеров, расположенных как можно дальше друг от друга. Алгоритм описывается следующими действиями [34]:

- 1) Случайно выбрать  $k$  точек, являющихся начальными “центрами масс” кластеров;
- 2) Отнести каждый объект к кластеру с ближайшим “центром масс”;
- 3) Пересчитать “центры масс” кластеров согласно их текущему составу;
- 4) Если критерий остановки алгоритма не удовлетворен, вернуться к п. 2.

В качестве критерия остановки алгоритма выбирают минимальное изменение среднеквадратической ошибки. Также возможность останавливаться, если на шаге 2 не было объектов, переместившихся из кластера в кластер. [34]

Недостаток данного алгоритма заключается в том, что необходимо задавать количество кластеров для разбиения. [34]

### 2.9.3 Нечеткие алгоритмы

Наиболее популярным алгоритмом нечеткой кластеризации является алгоритм  $s$ -средних. Он представляет собой модификацию метода  $k$ -средних. [34]

Алгоритм может быть описан следующей последовательностью действий [34]:

- 1) Выбрать начальное нечеткое разбиение  $n$  объектов на  $k$  кластеров путем выбора матрицы принадлежности  $U$  размера  $(n, k)$ ;
- 2) Используя матрицу  $U$ , найти значение критерия нечеткой ошибки:



$$E^2(X, L) = \sum_{i=1}^N \sum_{j=1}^K U_{ik} \left\| x_i^{(k)} - c_k \right\|^2, \quad (27)$$

где  $c_k$  — “центр масс” нечеткого кластера  $k$ :  $c_k = \sum_{i=1}^N U_{ik} x_i$ ,  
 $K$  — количество кластеров,  
 $N$  — количество точек.

- 3) Перегруппировать объекты с целью уменьшения этого значения критерия нечеткой ошибки;
- 4) Возвращаться в п. 2 до тех пор, пока изменения матрицы  $U$  не станут незначительными.

Данный алгоритм может не подойти, если заранее неизвестно число кластеров, либо необходимо однозначно отнести каждый объект к одному кластеру. [34]

Вычислительная сложность для алгоритмов  $k$ -средних и  $s$ -средних одинакова и равна  $O(n \cdot k \cdot l)$ , где  $k$  — число кластеров,  $l$  — число итераций, а  $n$  — количество точек для кластеризации. [34]

### **Вывод**

Юрий Владимирович запретил мне работать на стройке

## 3. Технологическая часть

### 3.1 Средства реализации программного обеспечения

При написании программного продукта был использован язык программирования Kotlin [35].

Данный выбор обусловлен следующими факторами:

- возможность запуска программного кода на любом устройстве, поддерживающем Java Virtual Machine;
- большое количество актуализируемой справочной литературы, связанной как с языком программирования Java, так и Kotlin;
- возможность интеграции программного кода в приложения для ОС Android.

При написании программного продукта использовалась среда разработки IntelliJ IDEA. Данный выбор обусловлен тем, что Kotlin является продуктом компании JetBrains, поставляющей данную среду разработки.

### 3.2 Выбор СУБД

#### 3.2.1 Базы данных временных рядов

Базы данных временных рядов отличаются от статических баз данных тем, что содержат записи, в которых некоторые из атрибутов ассоциируются с временными метками. В качестве таких записей могут выступать данные мониторинга, биржевые данные о торгах или транзакции продаж. [36]

**InfluxDB** InfluxDB - это база данных временных рядов, предназначенная для обработки высокой нагрузки записи и запросов.

Основным назначением является хранение больших объемов данных с метками времени. Например, данные мониторинга, метрики приложений и данные датчиков IoT (Internet of Things, интернет вещей).

В традиционной реляционной базе данных данные хранятся до тех пор, пока вы не решите их удалить. Учитывая сценарии использования баз данных временных рядов, можно не хранить данные слишком долго: это или слишком дорого, или данные со временем теряют актуальность. [37]

Системы вроде InfluxDB могут удалять данные спустя определённое время, используя концепцию, называемую политикой хранения. Также имеется возможность выполнять непрерывные запросы к оперативным данным для выпол-

нения определённых операций. [37]

**OpenTSDB** OpenTSDB состоит из демона временных рядов, а также набора утилит командной строки. Взаимодействие с OpenTSDB в первую очередь достигается путем запуска одного или нескольких независимых демонов.

Демон использует базу данных с открытым исходным кодом HBase или службу Google Bigtable для хранения и получения данных временных рядов. Схема данных высоко оптимизирована для быстрого объединения аналогичных временных рядов, чтобы минимизировать пространство хранения. Пользователям никогда не требуется прямой доступ к базовому хранилищу. Можно общаться с демоном через протокол telnet, HTTP API или простой встроенный графический интерфейс.

### 3.2.2 Реляционные базы данных

Реляционная база данных — это организованный по реляционной модели набор таблиц, в которых каждая ячейка этих таблиц имеет некоторое соответствующее описание. [38]

Использование реляционной модели предполагает возможность идентификации элементов по совокупности уникальных идентификаторов: имя столбца, первичный ключ. Для построения логической связи между строками и ячейками разных таблиц используются внешние ключи. [38]

Среди подобных СУБД, которые основаны на реляционных базах данных выделяются: Oracle, PostgreSQL.

Каждая из указанных СУБД имеет некоторые отличительные особенности. Так, например, PostgreSQL поддерживает вставки кода, написанного на языке программирования Python, в тело процедуры. Однако выделить среди данных особенностей важных для данной работы не представляется возможным.

### Вывод

Базы данных временных рядов являются наиболее подходящими для решаемой задачи, так как они нацелены на хранение, извлечение и анализ большого количества статистических данных, в которых имеются временные метки.

Для организации хранения данных будет использоваться СУБД InfluxDB, так как она является одной из самых популярных, среди известных баз данных временных рядов, а также по той причине, что поддержка данной СУБД всё ещё

не прекращена на сегодняшний день.

### 3.3 Выбор алгоритма кластеризации

В качестве используемого алгоритма кластеризации был выбран метод с-средних в силу того, что число кластеров заранее известно, а также задача рассматривает установку соответствия некоторого объекта (например, значения скорости печати) набора вещественных значений, показывающих степень отношения объекта к кластерам.

### 3.4 Данные для кластеризации

В качестве данных для кластеризации используются действия оператора автоматизированного рабочего места, производимые с использованием клавиатуры и мыши. Данные действия логируются с использованием программного обеспечения, а затем направляются в базу данных для возможности миграции определенной модели поведения на другое автоматизированное место.

Действия пользователя, участвующие в построении модели, соотносятся с временем реакции, которое фиксируется каждые 20 минут, причем по времени реакции определяется, в каком состоянии в текущий момент времени находится организм оператора.

### 3.5 Сведения о модулях

Программное обеспечение состоит из модулей логирования действий оператора, обработки и анализа данных.

#### 3.5.1 Модуль логирования действий оператора

Данный модуль предназначен для записи информации о действиях оператора.

Внешние зависимости модуля:

- Java Swing [39] — библиотека легковесных компонентов для реализации оконного интерфейса приложения;
- JNativeHook [40] — библиотека, предоставляющая средства перехвата прерываний, поступающих от клавиатуры и мыши.

Модуль состоит из четырех пакетов.

**Пакет window** Данный пакет включает в себя абстрактный класс Window, предоставляющий родительский класс с определенными свойствами для всех окон реализуемого программного обеспечения.

```

1 package window
2
3 import javax.swing.GroupLayout
4 import javax.swing.ImageIcon
5 import javax.swing.JComponent
6 import javax.swing.JFrame
7
8 open class Window: JFrame() {
9
10     init {
11         iconImage =
12             ImageIcon(this.javaClass.getResource("/bb.png")).image
13     }
14
15     fun createLayout(vararg components: JComponent) {
16         val gl = GroupLayout(contentPane)
17         contentPane.layout = gl
18
19         gl.autoCreateContainerGaps = true
20
21         val horizontalGroup =
22             gl.createParallelGroup(GroupLayout.Alignment.CENTER)
23         components.forEach { horizontalGroup.addComponent(it) }
24
25         gl.setHorizontalGroup(horizontalGroup)
26
27         val verticalGroup = gl.createSequentialGroup()
28         components.forEach { verticalGroup.addComponent(it) }
29
30         gl.setVerticalGroup(verticalGroup)
31
32         pack()
33     }
34 }

```

**Пакет bigBrother** Данный пакет включает в себя класс BigBrotherWindow, который является реализацией класса Window и определяет функционал главного экрана приложения.

**Пакет `loggers`** Данный пакет включает в себя пакеты реализаций логирующих классов: `KeyLogger` (нажатия на клавиши клавиатуры), `MouseLogger` (нажатия на клавиши мыши и движения данного устройства), `ReactionLogger` (результаты пройденных тестов на реакцию).

Также в данном пакете предоставлена реализация класса `ReactionTestWindow` предоставляющая интерфейс и логику определения реакции пользователя по нажатию на кнопку, появляющуюся в случайные моменты времени (от 2 до 10 секунд).

Каждый логирующий класс локально создает текстовый файл, в который записывает в определенном формате собранные данные. Для исключения попыток изменения файла конкурирующими потоками в каждом из них представлена реализация очереди записи, которая переносится в файл при достижении размера в сотню записей, либо по завершению работы приложения.

**Пакет `random`** Данный пакет включает в себя класс, реализованного по паттерну “Одиночка”, предоставляющий доступ к классу `Random`, инициализированного отложено. Данный класс используется преимущественно для получения данных о реакции пользователя.

Также в модуле определен файл `main.kt`, который является точкой входа в приложения.

На рисунке 3.1 предоставлена диаграмма классов модуля.

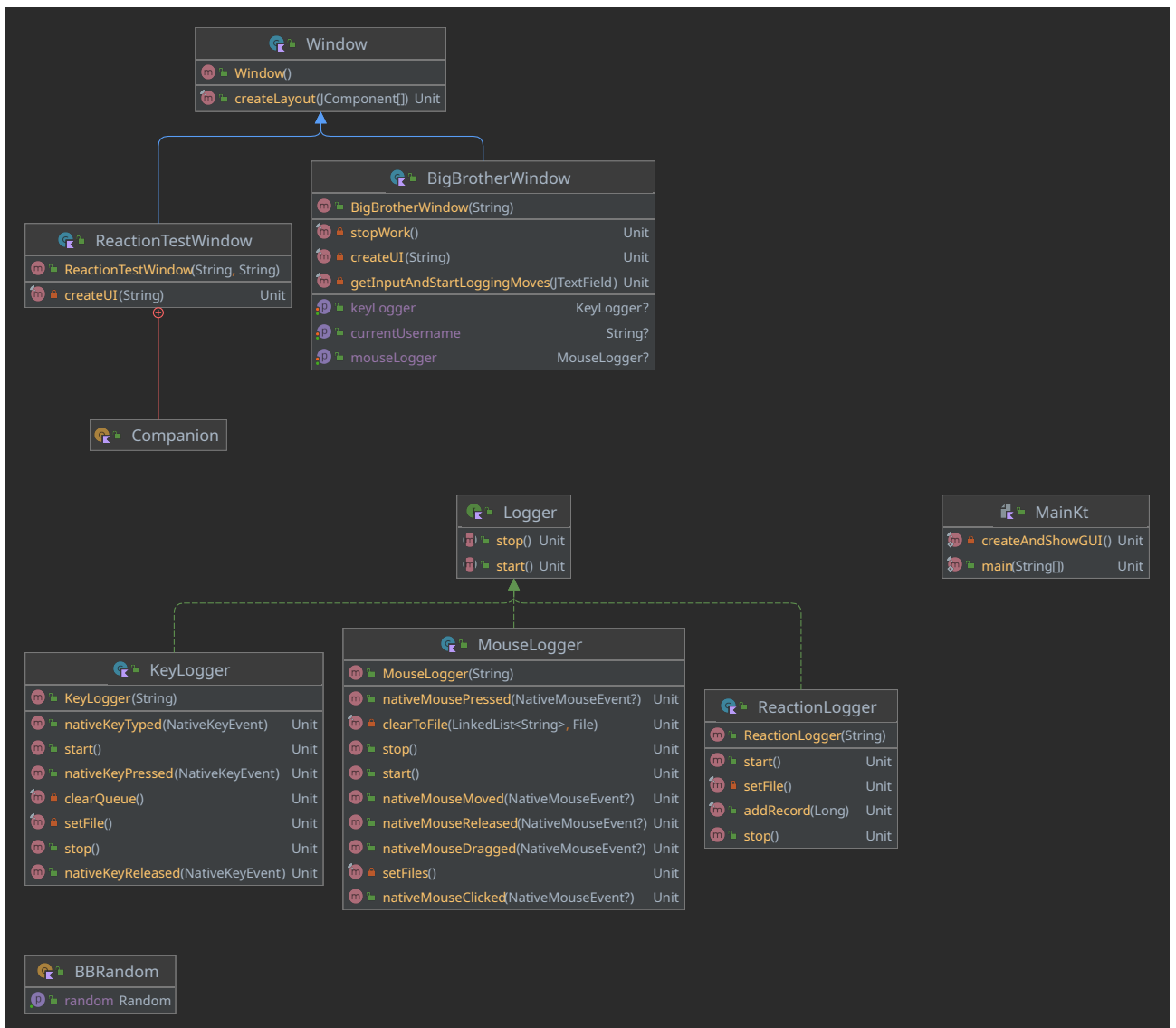


Рис. 3.1: Диаграмма классов модуля логирования.

### 3.5.2 Модуль обработки данных

Модуль обработки данных включает в себя два пакета: парсинга данных и приведения данных.

Модуль парсинга данных...

Модуль приведения данных...

### 3.5.3 Модуль анализа данных

## 3.6 Развертывание системы

## ЗАКЛЮЧЕНИЕ

Была описана проблематика стресса на рабочем месте и метод распознавания усталости пользователя на АРМ с использованием доступных технологий определения усталости, построенных на анализе отдельных характеристик и действий.

Были описаны термины предметной области: усталость, хроническая усталость, стресс. Также рассмотрены стадии общего адаптационного синдрома и понятие профессионального стресса. Проведенная аналитическая работа позволила обозначить проблему: на рабочем месте человек испытывает стресс, избыток которого может навредить человеку и работодателю.

Была формализована цель прототипирования метода СРУ на АРМ: предупреждение наступления стадии истощения посредством регистрации момента наступления стадии тревоги или стадии тревоги и затем устойчивости. Назначение данного метода заключается в выявлении стадии истощения организма для его дальнейшего плодотворного функционирования.

Проведён анализ действий и характеристик, позволяющих определить усталость пользователя АРМ. Для регистрации действий и характеристик были выбраны внешние и носимые устройства.

Рассмотрены следующие внешние устройства, позволяющие определить усталость пользователя АРМ: клавиатура, координатное устройство типа мышь, веб-камера, микрофон.

Рассмотрен метод распознавания усталости пользователя по клавиатурному почерку, который определяется по времени между нажатиями клавиш.

Рассмотрен метод распознавания усталости пользователя по разнице скорости перемещения курсора и оценки среднего времени перемещения курсора мыши между элементами интерфейса.

Рассмотрен метод распознавания усталости с использованием изображений, получаемых от веб-камеры, который строится на использовании технологий машинного обучения и предварительно обученных на наборах данных моделей.

Рассмотрен метод распознавания усталости с использованием аудио-потока, построенного на анализе связной речи человека.

Рассмотрены биофизические факторы, позволяющие определить уста-



лость пользователя: частота пульса, возраст сосудистой системы, индекс стресса (индекс Баевского). В качестве задействованной в системе характеристики был выбран индекс стресса. В качестве устройства для определения выбранной характеристики могут использоваться смарт-часы, поддерживающие подобную функцию.

Таким образом, были определены методы снятия выделенных действий и характеристик.

Формализованный прототип метода СРУ на АРМ включил в себя:

- хранение и анализ данных, получаемых от клавиатуры;
- хранение и анализ данных об индексе стресса субъекта.

# СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

## Список литературы

1. Gallup. Employee Burnout: Causes and Cures // Gallup. 2020. p. 32.
2. Moss J. Burnout Is About Your Workplace, Not Your People [Электронный ресурс]. Режим доступа: <https://hbr.org/2019/12/burnout-is-about-your-workplace-not-your-people> (дата обращения 27.03.2021).
3. Г.А. Макарова. Синдром эмоционального выгорания. М.: Просвещение, 2009. с. 432.
4. Е.А. Пигарова А.В. Плещева. Синдром хронической усталости: современные представления об этиологии // Ожирение и метаболизм. 2010. с. 13.
5. Т.В. Байдина Т.И. Колесова Б.В. Малинина. Усталость как симптом неврологических заболеваний // Пермский медицинский журнал. 2021. Т. 38, № 2. С. 37–44.
6. Н.В. Пизова А.В. Пизов. Когнитивные нарушения и синдром хронической усталости // Нервные болезни. 2021. № 3. С. 10–16.
7. Committee on the Diagnostic Criteria for Myalgic Encephalomyelitis/Chronic Fatigue Syndrome; Board on the Health of Select Populations; Institute of Medicine // The National Academies Collection: Reports funded by National Institutes of Health. 2015.
8. Долбышев А.В. Нейрофизиологические механизмы стресса // StudNet. 2020. № 7. С. 163–167.
9. Г.В. Порядина. Стресс и патология: учеб. пособие. М.: РГМУ, 2009. с. 23.
10. Деева О.С. Причины профессионального стресса и методы его профилактики // Ученые записки Тамбовского отделения РоСМУ. 2019. Т. 38, № 14. С. 140–146.

11. Информационное общество в Российской Федерации, 2020. Статистический сборник [Электронный ресурс]. Режим доступа: <https://rosstat.gov.ru/storage/mediabank/lqv3T0Rk/info-ob2020.pdf> (дата обращения 08.12.2021). 2020.
12. Большой энциклопедический политехнический словарь [Электронный ресурс]. Режим доступа: <https://rus-big-polyheh-dict.slovaronline.com/> (дата обращения 22.12.2021).
13. В.И. Васильев А.Е. Сулавко Р.В. Борисов. Распознавание психофизиологических состояний пользователей на основе скрытого мониторинга действий в компьютерных системах // ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ПРИНЯТИЕ РЕШЕНИЙ. 2017. С. 21–37.
14. Е. В. Шкляр Е. Г. Воробьев М. Ф. Савельев. Распознавание клавиатурного почерка в браузере // Известия СПбГЭТУ «ЛЭТИ». 2019. С. 58–63.
15. Axelsson S. The Base-Rate Fallacy and its Implications for the Difficulty of Intrusion Detection. 1999. 12. Р. 1–10.
16. Попов А.Ю. Макушкина Л.А. ИССЛЕДОВАНИЕ ПСИХОФИЗИЧЕСКОГО СОСТОЯНИЯ ПОЛЬЗОВАТЕЛЯ ИНФОРМАЦИОННОЙ СИСТЕМЫ // Вестник магистратуры. 2015. Т. 41, № 2-2.
17. Раскин Д. Интерфейс: новые направления в проектировании компьютерных систем. Символ-плюс, 2010. с. 272.
18. Р.В. Борисов Д.Н. Зверев А.Е. Сулавко. Оценка идентификационных возможностей особенностей работы пользователя с компьютерной мышью // Вестник СибАДИ. 2015. С. 106–113.
19. Савинова В.М. Бесхмельницкий А.А. Бибина Е.С. Идентификация пользователей корпоративной системы с помощью поведенческого анализа с использованием модели искусственной нейронной сети // ТДР. 2017. № 5.
20. И.Б. Лашков А.М. Кашевник. Определение опасных состояний водителя на основе мобильных видеоизмерений его лицевых характеристик // ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ. 2019. С. 84–96.

21. М.А. Лебедев С.Ю. Палатов Г.В. Ковров. Усталость и ее проявления // Медицинское обозрение. 2014. С. 282–288.
22. Д. Бернс. Терапия настроения: Клинически доказанный способ победить депрессию без таблеток. М.: Альпина Паблишер, 2019. с. 271.
23. В.В. Киселев. Автоматическое определение эмоций по речи // Технологии и практика обучения. 2012. С. 85–89.
24. Voice Stress Analysis: A New Framework for Voice and Effort in Human Performance / Martine Van Puyvelde, Xavier Neyt, Francis McGlone [и др.] // Frontiers in Psychology. 2018. Т. 9, № 14.
25. Федеров В.М. Реблев Д.П. Панченко Е.М. Идентификация пользователя по виброакустическим шумам // Известия ЮФУ. Технические науки. 2013. Т. 149, № 12.
26. Рублев Д.П. Федоров В.М. Идентификация пользователя по динамическим характеристикам работы с манипулятором “мышь” с использованием нейронных сетей // Известия ЮФУ. Технические науки. 2017. Т. 190, № 5.
27. Кропачев И.Г. Архипова Е.И. Нора С.А. Показатели сердечно-сосудистой системы и гемодинамики в условиях стрессового воздействия как фактор развития острых респираторных инфекций // Вестник НовГУ. 2020. Т. 119, № 3.
28. А.В. Капустин. Изменение показателей биологического возраста при патологии сосудов // Бюллетень Северного Государственного медицинского Университета. 2018. № 1.
29. Э.Э. Ибрагимова. Мониторинг уровня стресса обучающихся как подход профилактики нарушения регуляторных механизмов // Ученые записки Крымского федерального университета имени В. И. Вернадского. Социология. Педагогика. Психология. 2019. № 2.
30. Р.М. Баевский. Прогнозирование состояний на грани нормы и патологии. М.: Медицина, 1979. с. 298.

31. Стукова Ю.Е. Иваницкий Д.К. Савченко М.И. ПРИМЕНЕНИЕ МЕТОДА КОРРЕЛЯЦИОННОГО АНАЛИЗА В ОЦЕНКЕ ЛИКВИДНОСТИ И ПЛАТЕЖЕСПОСОБНОСТИ ОРГАНИЗАЦИЙ // ЕГИ. 2021. Т. 36, № 4.
32. М.И. Румянцев. Корелляционный анализ: методические указания. Магнитогорск: Магнитогорский государственный технический университет имени Г. И. Носова, 2017. с. 18.
33. Абжандадзе З.Л. Федорова Е.В. Юдовин М.Э. Элементы дисперсионного анализа: методические указания. СПб., 2009. с. 22.
34. Ершов К.С. Романова Т.Н. Анализ и классификация алгоритмов кластеризации // Новые информационные технологии в автоматизированных системах. 2016. № 19.
35. Kotlin language specification [Электронный ресурс]. Режим доступа: <https://kotlinlang.org/spec/introduction.html> (дата обращения 09.10.2020).
36. Шабельников А.Н. Шабельников В.А. Поиск аномалий в технических базах данных временных рядов // Известия ЮФУ. Технические науки. 2008. № 4.
37. TProger : Знакомство с InfluxDB и базами данных временных рядов [Электронный ресурс]. Режим доступа: <https://tproger.ru/translations/influxdb-guide/> (дата обращения 01.04.2020).
38. Н.Р. Булахов. Основы реляционных баз данных // Вестник науки и образования. 2019. Т. 76, № 22-2.
39. Документация Oracle - Java Swing [Электронный ресурс]. Режим доступа: <https://docs.oracle.com/javase/7/docs/api/javax/swing/package-summary.html> (дата обращения 16.05.2022).
40. Официальный репозиторий GitHub проекта JNativeHook [Электронный ресурс]. Режим доступа: <https://github.com/kwhat/jnativehook> (дата обращения 16.05.2022).