

Санкт-Петербургский государственный университет

Кафедра системного программирования  
Программная инженерия

Горбатюк Сергей Петрович

# Сбор и анализ биометрических маркеров с помощью смартфона

Курсовая работа

Научный руководитель:  
старший преподаватель Кириленко Я. А.

Консультант:  
кандидат психологических наук, доцент кафедры общей психологии СПбГУ Волков Д. Н.

Санкт-Петербург  
2019

# Оглавление

<b>Введение</b>	<b>3</b>
<b>1. Цель работы</b>	<b>5</b>
<b>2. Обзор</b>	<b>6</b>
2.1. Система в целом . . . . .	6
2.2. Измерение ЧСС с помощью камеры смартфона . . . . .	7
<b>3. Реализация клиентской части</b>	<b>8</b>
3.1. Архитектура . . . . .	8
3.2. Аспекты реализации и стек технологий . . . . .	9
<b>4. Алгоритмы вычисления ЧСС</b>	<b>10</b>
4.1. Измерение в состоянии покоя . . . . .	10
4.2. Измерение в состоянии релаксации . . . . .	11
<b>5. Эксперименты и анализ выявленных проблем</b>	<b>13</b>
5.1. Обобщенные численные результаты . . . . .	13
5.1.1. Первая серия тестов . . . . .	13
5.1.2. Вторая серия тестов . . . . .	14
5.2. Точность и полнота алгоритма . . . . .	14
5.3. Технические проблемы . . . . .	15
<b>Заключение</b>	<b>17</b>
<b>Список литературы</b>	<b>18</b>

# Введение

В современном мире становится все более актуальным вопрос мониторинга состояния здоровья человека и принятия мер для его поддержания и укрепления. Эта тема — основа для многочисленных исследований научных групп и лабораторий как частных компаний [7], так и подразделений университетов, институтов физической культуры и спорта, ученых в области психологии, и других (например, [12], [11]). Исследования невозможны без экспериментов и данных, которые специалисты собирают преимущественно в лабораторных условиях с привлечением групп испытуемых — спортсменов или обычных людей разных возрастов. Часть исследования может состоять в оценке одного или нескольких биометрических маркеров человека в заданных условиях, например, частоты сердечных сокращений, способности удерживать равновесие, объема выдыхаемого воздуха, и так далее. Лабораторное оборудование позволяет делать измерения с высокой точностью, выполнять соответствующие расчеты, и получать метрики, на которых в дальнейшем ученые выстраивают и проверяют гипотезы, однако, оно, обычно, сложное и имеет размеры, исключаяющие его транспортировку. Последние аспекты означают, что возможности сбора данных в ряде случаев ограничены наличием лаборатории и ее возможностями. В противоположность этому, смартфон имеется уже практически у каждого человека, и сенсоры, которыми оснащены все устройства, находят на первый взгляд не очевидное, но вполне осмысленное применение в перечисленных выше задачах. При этом точность этих сенсоров достаточная для того, чтобы иметь возможность собирать требуемые данные, и соответствующие алгоритмы могут быть разработаны для их анализа. В то же время массовость распространения смартфонов влечет возможность собирать данные в гораздо больших объемах и в более короткие сроки. В связи с этим возникает идея построения централизованной, и в должной степени анонимизированной системы для сбора биометрических маркеров с помощью смартфона. Собранные данные будут храниться на сервере с возможностью их обработки различными

алгоритмами для выработки способов анализа и метрик, необходимых специалистам для исследований.

# 1. Цель работы

Цель этой работы — разработка полноценной системы для сбора биометрических показателей, которая будет представлять серверную и клиентскую части, и позволять встраивать модули для измерения конкретного показателя. Также в рамках работы планируется создать модуль для измерения частоты сердечных сокращений (ЧСС) с помощью камеры смартфона. Отметим, что основное отличие от существующих работ, часть из которых будет рассмотрена ниже, состоит в том, что планируется создать именно целостную систему для сбора данных, призванную оказать весомую помощь в научных исследованиях, направленных на изучение аспектов, связанных со здоровьем человека, и его физических особенностей.

## **Задачи:**

1. Разработка и реализация архитектуры приложения (клиентской части)
2. Разработка решения для серверной части (выполнено коллегой по команде и описано в отдельной работе)
3. Изучение имеющихся и выработка собственного алгоритма определения ЧСС с учетом специфики исследовательских задач
4. Реализация модуля на основе выработанного алгоритма
5. Исследование средств тестирования ПО и их применение в курсовой работе

## 2. Обзор

### 2.1. Система в целом

Системы сбора показателей с портативных устройств, вероятнее всего, существуют внутри компаний, производящих соответствующие устройства, однако, они не находятся в открытом доступе, или не используются массово. В то же время в магазинах приложений для разных платформ существует огромное множество приложений для измерения ЧСС, использующих другие сенсоры мобильных устройств, а также приложения, предоставляющие программы оздоровления и тренировок, основанные, по заверениям авторов, на научных данных и исследованиях (например [1], [10]). Все существующие приложения решают задачу с достаточной точностью, обладают множеством функций, однако не отвечают некоторым специальным требованиям, возникающим в процессе научных исследований, так как ориентированы сразу на широкую аудиторию пользователей. Поясним, что здесь имеется в виду под вышеуказанными требованиями. Для научных изысканий в области оценки физического состояния и здоровья человека проводятся обычные испытания, но ученых зачастую интересуют специфичные детали и закономерности в полученных данных. Например, для метода оценки аэробной выносливости требуется анализ кривой снижения пульса. При этом требуется знать ЧСС в момент сразу после выполнения специального упражнения, и ровно через 60 секунд после этого. Существующие на данный момент алгоритмы не решают конкретно эту задачу, так как в большинстве случаев делают априорное предположение о том, что ЧСС слабо меняется в процессе измерения, что в поставленной задаче допустить невозможно. Таких задач, то есть довольно специфичных, хотя и поддающихся обобщению тестов, может быть много, и поэтому основной задачей данной работы является создание системы, способной вмещать в себя разные компоненты, реализующие требуемую логику. Аналогов системы, подлежащей разработке, обнаружено не было.

## 2.2. Измерение ЧСС с помощью камеры смартфона

Существует немалое число работ, направленных на решение задачи определения частоты сердечных сокращений с помощью камеры смартфона. В [14] дается подробный разбор методов анализа ряда интенсивности красного канала изображений, получаемых с камеры. Под анализом здесь понимается алгоритм определения частоты сигнала, то есть частоты сердечных сокращений. Два основных метода выделения доминирующей частоты сигнала — подсчет пиков и преобразование Фурье. Первый метод обладает меньшей устойчивостью, так как измерение с помощью камеры может отражаться шумом на данных временного ряда, что сильно повлияет на результаты подсчета пиков. Для решения этой проблемы применяются различные эвристики, например, восстановление потенциально пропущенных пиков. Метод, использующий преобразование Фурье более устойчив к шумам, но асимптотически работает дольше, и для приемлемой точности (малого шага) требует сравнительно большое число данных. В работе [2] описывается комплексный подход, сочетающий описанные выше стратегии, для снижения недостатков обоих методов, и содержащий полезные технические идеи, как например перевод сигнала с камеры не в RGB, а в HSV формат, позволяющий сохранить больше информации, а также обрезание картинки с камеры для обработки таким образом, чтобы поймать баланс между скоростью и производительностью на конкретном устройстве. В одной из ранних работ [8] для подсчета числа сердечных сокращений использовался алгоритм поиска пиков, дополненный эвристикой вырезания из временного ряда релевантного окна, содержащего заранее заданное число «ударов». Несмотря на указанные авторами проблемы (например, световые помехи и шумы при измерении), результаты алгоритма обладают низкой относительной ошибкой.

## 3. Реализация клиентской части

### 3.1. Архитектура

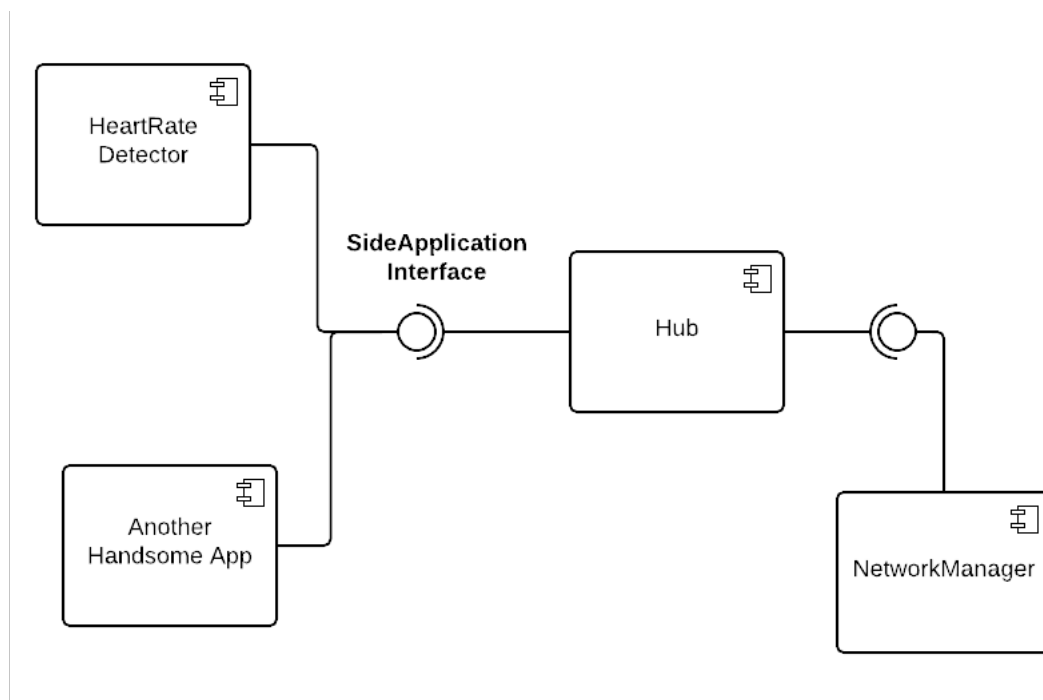


Рис. 1: Обобщенная архитектура

Ниже приводится описание компонентов, указанных на схеме (Рис. 1)

- **Hub** – главный управляющий класс. Он осуществляет взаимодействие между сторонними приложениями (SideApplication), сервером, базой данных на устройстве (при ее наличии), и логгирование. Чтобы исключить конфликты, архитектурой требуется, чтобы существовал единственный объект этого класса, что и реализуется с помощью паттерна Singleton. Взаимодействие с вызывающими его приложениями происходит с помощью обратных вызовов, а также он контролирует, чтобы в один момент было запущено лишь одно стороннее приложение, вновь, чтобы исключить возможные конфликты доступа и запросов к данным или серверу.
- **NetworkManager** – класс для взаимодействия с сервером. Осуществляет асинхронные запросы на сервер, передавая требуемые



данные. Запрос осуществляется к API сервера по конкретному адресу, определяемому с помощью имени приложения, которое делает запрос. Принятые данные передаются вместе с тегом обратно в Hub, который осуществляет их передачу приложению, сделавшему запрос.

- **SideApplication** – интерфейс, описывающий стороннее приложение, включаемое в систему. Hub не вызывает напрямую с помощью объектов Intent классы приложений, реализующие интерфейс Activity, потому что возможны сценарии, когда до запуска непосредственно интерфейса, приложению нужно, например, показать флеш-заставку, при этом обратившись в память, или к серверу. Для этого в методе start() приложению дается возможность сделать любую предварительную работу. При вызове метода stop() приложение обязано завершить работу, это - служебный метод.

## 3.2. Аспекты реализации и стек технологий

Основная часть клиентской части, а именно управляющие модули, написана на языке Kotlin [6], разработанный компанией JetBrains [5] с использованием библиотеки Volley [9] для выполнения асинхронных запросов.

В то же время модуль для измерения ЧСС написан на Java [4], и включает в себя лишь компонент для преобразования IUUV формата в RGB для того, чтобы отправлять на сервер сырые данные интенсивности красного канала снимков с камеры. Остальная часть алгоритма находится на серверной части для того, чтобы ее удобнее было обновлять, а также для сохранения кода в тайне на случай возможного создания патента и коммерциализации системы.

## 4. Алгоритмы вычисления ЧСС

### 4.1. Измерение в состоянии покоя

Состояние покоя характеризуется тем, что субъект не выполнял предварительно никакое физическое упражнение для поднятия ЧСС, то есть, измерение проводится в так называемых нормальных условиях. В течение некоторого времени (порядка 20 секунд) происходит считывание данных с камеры, при включенном светодиоде, во время того, как пользователь приложил к ней указательный палец. Полученные данные из формата YUV420SP (стандартный формат кадров для камеры смартфона на ОС Android) переводятся в RGB представление, после чего для каждого измерения берется значение красного канала. Таким образом, имея последовательность измерений значений этого цвета на кадрах, можем интерпретировать ее как временной ряд для интенсивности приливов крови в капиллярах пальца. Дальнейшая задача заключается в том, чтобы определить по нему ЧСС, зная частоту съема кадров с камеры (fps). Для определения доминирующей частоты сигнала выбран метод с применением быстрого преобразования Фурье (БПФ), алгоритма, позволяющего выполнять дискретное преобразование Фурье за время  $O(n \log n)$  вместо  $O(n^2)$ :

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} = \sum_{n=0}^{N-1} x_n \left( \cos \frac{2\pi kn}{N} - i \sin \frac{2\pi kn}{N} \right)$$

Метод, использующий преобразование Фурье, выбран как наиболее точный и устойчивый к помехам. Для улучшения точности используется также метод пересекающихся отрезков, заключающийся в том, что данные разбиваются на некоторое число отрезков, каждый из которых (кроме первого) имеет начальную точку ровно посередине предыдущего. После этого вычисление доминирующей частоты производится на каждом из отрезков, и на всем промежутке измерений. Имея все результаты, можно реализовывать разные стратегии вычисления ответа. В самом строгом случае можно потребовать совпадения всех значений,

и тогда выдавать таковое в качестве ответа, а в противном случае сообщать пользователю о неточном измерении. Стоит заметить, что ситуация совпадения значений на всех интервалах достаточно редка, поэтому такой способ будет слишком часто сообщать об ошибке измерений. Другой подход — сообщать в качестве ответа самый частый из вариантов ответов на промежутках. В этой ситуации, за исключением отбраковывания заведомо невозможных для человека результатов, мы всегда будем получать какой-то ответ, однако рискуем потерять точность, отбрасывая остальные данные. В приложении реализован алгоритм, объединяющий преимущества двух крайних подходов, то есть следящий за тем, чтобы большинство значений совпадало (иначе сообщаящий о неточном измерении), но в некоторых случаях усредняющий значения на промежутках для сохранения информации обо всем измерении.

## 4.2. Измерение в состоянии релаксации

Релаксация в этом контексте понимается как восстановление обычной функциональности организма после выполнения специальных физических упражнений. В этом режиме измерение происходит похожим образом, но длится дольше (примерно 60 секунд). По полученному большому временному ряду методом скользящего окна вычислим несколько десятков значений ЧСС. Конкретнее, зададим ширину интервала, и будем вычислять на нем пульс с помощью БПФ, но без применения метода пересекающихся отрезков, так как в этом случае, находясь в предположении, что пульс постоянно снижается, отрезки скорее всего дадут разные результаты, которые будет затруднительно интерпретировать. Так как измерение ЧСС по единичному БПФ может давать ошибку (например, вдвое меньшая частота окажется доминирующей), то эвристическими методами из всего измерения удаляются значения ЧСС, сильно отклоняющиеся от среднего. После этого мы хотим наиболее точно приблизить имеющиеся данные некоторой функцией от времени, чтобы иметь возможность вычислять значение ЧСС в любой адекватный момент времени. После проведенных экспериментов был сделан

вывод, что наибольшую точность обеспечивает экспоненциальная зависимость —  $y = ae^{bt}$ , где  $t$  - время. Теперь, поиск коэффициентов  $A$ ,  $b$  — это оптимизационная задача для решения которой выбран метод наименьших квадратов. Будем искать решение для логарифмированного равенства:  $y = a + bt$ , где  $a \equiv \ln A$ . Согласно методу наименьших квадратов, будем минимизировать функцию

$$\sum_{i=1}^n y_i(\ln y_i - a - bt_i)$$

Применяя метод наименьших квадратов, получим систему:

$$a \sum_{i=1}^n y_i + b \sum_{i=1}^n x_i y_i = \sum_{i=1}^n y_i \ln y_i$$

$$a \sum_{i=1}^n x_i y_i + b \sum_{i=1}^n x_i y_i \ln y_i$$

Решая для  $a$ ,  $b$ , получаем:

$$a = \frac{\sum_{i=1}^n (x_i^2 y_i) \sum_{i=1}^n (y_i \ln y_i) - \sum_{i=1}^n (x_i y_i) \sum_{i=1}^n (x_i y_i \ln y_i)}{\sum_{i=1}^n y_i \sum_{i=1}^n (x_i^2 y_i) - (\sum_{i=1}^n (x_i y_i))^2} \quad (1)$$

$$b = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n (x_i y_i \ln y_i) - \sum_{i=1}^n (x_i y_i) \sum_{i=1}^n (y_i \ln y_i)}{\sum_{i=1}^n y_i \sum_{i=1}^n (x_i^2 y_i) - (\sum_{i=1}^n (x_i y_i))^2} \quad (2)$$

Теперь, вспоминая, что  $A \equiv e^a$ , получаем формулу для вычисления ЧСС в произвольный момент времени, в частности, при  $t=0$ ,  $t=60$ , т. е. в случаях, для которых эта задача и была сформулирована. Отметим, что вычисления для уравнений (1), и (2) могут быть сделаны за линейное от числа точек время, а так как оно обычно не превосходит сотни, вычисление можно считать очень быстрым. Также заметим, что мы решили более сильную задачу, так как анализ графика снижения частоты сердечных сокращений является центральным местом некоторых тестов для определения физического состояния человека.

## 5. Эксперименты и анализ выявленных проблем

Мои коллеги из НИИФК [13] провели ряд экспериментов, направленных на тестирование функциональности и точности алгоритма для измерения ЧСС. Было проведено две серии тестов - до и после улучшения алгоритма с помощью экспоненциальной функции в режиме релаксации. До введения последней алгоритм по сути состоял только лишь из одного измерения с применением быстрого преобразования Фурье и метода пересекающихся отрезков. Различий между режимами измерения не делалось. Для проведения экспериментов тестовая версия приложения была установлена на смартфон Sony Xperia M4 Aqua Dual (E2333), а в качестве контрольного прибора использовался профессиональный пульсометр Polar M600. Была проведена серия экспериментов в трех режимах: режим покоя, режим релаксации, и режим восстановления (измерение после минутного восстановления вслед за выполненной нагрузкой).

### 5.1. Обобщенные численные результаты

#### 5.1.1. Первая серия тестов

При измерении в режиме покоя выявлены лишь незначительные отклонения от контрольного прибора, и точность алгоритма признана удовлетворительной. При измерении в двух других случаях количество расхождений с контрольным прибором возрастало. Ошибочные ответы в некоторых случаях, вероятнее всего, были обусловлены зашумленностью сигнала, что дало ошибку при преобразовании Фурье, так как ответ отличался от действительного значения в целое число раз (в подавляющем большинстве из этих случаев - в 2 раза). Однако, были и другие ошибки, необъяснимые подобным образом. Например, при измерении "ЧСС релаксации" алгоритм показывал результат 86, а контрольный прибор – 110. Те же случаи имели место при измерении "ЧСС восстановления": алгоритм показывает 39, а контрольный прибор

– 65. Подобного рода ошибки, вероятнее всего, были связаны с шумом при измерении (например, движение пальца), и несовершенностью алгоритма.

### **5.1.2. Вторая серия тестов**

После разделения режимов, и реализации отдельного метода для измерения специфического случая релаксации точность повысилась значительно. Эти изменения не затронули алгоритм измерения в состоянии покоя, однако в двух других случаях наблюдалось значительное улучшение результатов: в нескольких измерениях не было значимых различий между результатами алгоритма и контрольного прибора, в других наблюдались расхождения в 5-7 ударов в режиме релаксации, и в 5-10 ударов в режиме восстановления. Мы будем иметь возможность улучшать алгоритм, создав систему, описываемую в данной работе, так как получим доступ к большому числу данных, и сможем тестировать различные алгоритмы очистки и обработки сигналов в неограниченном масштабе.

## **5.2. Точность и полнота алгоритма**

В процессе экспериментов появлялись случаи, когда измерение признавалось плохим, то есть алгоритм сообщал о невозможности точно определить пульс. Напомним, что это решение принимается алгоритмом на основании сравнения значений на отдельных отрезках, и на всем измерении в целом. До введения отдельного алгоритма для измерения пульса в режиме релаксации более 75% сообщений о неточном измерении приходились именно на этот режим, что легко объясняется высокой вариативностью значений ЧСС в этом случае. Это являлось одной из главных причин создания для этого случая отдельного режима измерений для релаксации. Однако сообщения о неточном измерении все еще случаются в режиме покоя. Здесь мы сталкиваемся с проблемой баланса следующего рода: можно сделать совсем нестрогие критерии для оценки точности измерения, и в этом случае мы будем сообщать

о плохом измерении крайне редко, или вовсе никогда. Однако, в этом случае чаще будут появляться неадекватные результаты и значения, обусловленные реальной зашумленностью входных данных вследствие тех или иных событий. При такой стратегии полнота охвата будет большой, но точность значительно пострадает. Обратный подход заключается в максимально строгом сравнении значений на всех интервалах, и выдаче ответа только в том случае, если они полностью совпадают, или максимально близки, то есть, находятся на расстоянии шага дискретизации. В этом случае полнота охвата сильно снизится, однако точность измерений скорее всего максимально приблизится к профессиональным приборам. В будущем возможно введение для пользователя возможности выбирать режим "строгости" измерений.

### **5.3. Технические проблемы**

Основной проблемой, обнаруженной при измерении, оказался сильный нагрев светодиода камеры смартфона. Это вызывало дискомфортные ощущения у участников эксперимента, и могло значительно сказаться на результатах эксперимента. Было сделано предположение, что в более ранних моделях смартфонов были установлены незащищенные от чрезмерного нагрева светодиодные элементы, что и являлось причиной нагрева. Чтобы проверить эту гипотезу, был проведен ряд экспериментов с разными моделями смартфонов, результаты которого представлены ниже (см. Таблицу 1). Эксперимент состоял в изучении ощущений от длительного контакта со включенным светодиодом. Время тестирования составляло 2 минуты, в то время как самый долгий период измерения для алгоритма составляет 1 минуту.

<b>Значительный нагрев</b>	<b>Незначительный нагрев</b>
Xiaomi Mi A2	Sony Xperia Z
Xiaomi Mi Max	Xiaomi Mi 8
Nokia 5	OnePlus 6
Alcatel Pop	Xiaomi Mi A1
	Samsung Galaxy S6
	Nokia 3.1 Plus
	Samsung Galaxy S5
	OnePlus 5
	IPhone X
	IPhone 6
	Samsung Galaxy J2 Prime
	IPhone 6S
	IPhone SE
	IPhone 5S
	Xiaomi Redmi 5+

Таблица 1: Сравнение смартфонов по признаку нагрева светодиода



## Заключение

Результатом работы является репозиторий с кодом [3], который хорошо документирован и покрыт тестами, которые запускаются автоматически при добавлении кода (совершении commit-push). Создана хорошая платформа для работы других специалистов, готовых писать модули-приложения для других физических проб и экспериментов. В реализованном модуле для измерения ЧСС многое можно улучшить, но это, вероятнее всего, задача для специалистов в области обработки сигналов, удобные условия для которых (а именно – большое число примеров, нам также удалось создать). Интерфейс модуля для тестирования также допускает доработки в области UI, но функциональность его полностью реализована. При тестировании, однако, выявлен ряд проблем, связанных в первую очередь с чрезмерным нагревом светодиода камеры на некоторых устройствах. Хотя небольшое исследование показало, что такие устройства составляют малую часть от всего многообразия смартфонов, все же, область применения технологии несколько сужается. Серверная компонента также реализована, и готова к развертыванию на любом сервере. Цели работы считаю в большей степени достигнутыми, и результат в целом успешным.

## Список литературы

- [1] Accurate Heart Rate Monitor. Приложение // Google Play Market. — 2017. — URL: <https://play.google.com/store/apps/details?id=com.repsi.heartrate> (дата обращения: 15.12.2018).
- [2] Arpan Pal Aishwarya Visvanathan Anirban Dutta Choudhury Aniruddha Sinha. Improved heart rate detection using smart phone. — 2014.
- [3] BiometricHub. GitLab code repository. — 2019. — URL: <https://gitlab.com/SergeyGorbatyuk171/biometrichub> (online; accessed: 13.05.2019).
- [4] Java. Programming language. — 2019. — URL: [https://www.java.com/ru/about/whatis\\_java.jsp](https://www.java.com/ru/about/whatis_java.jsp) (дата обращения: 11.05.2019).
- [5] JetBrains. Software Development Company. — 2019. — URL: <https://www.jetbrains.com/> (дата обращения: 11.05.2019).
- [6] Kotlin. Programming language by JetBrains. — 2019. — URL: <https://kotlinlang.org/> (дата обращения: 11.05.2019).
- [7] Nike Research. Сайт лаборатории // Google. — 2019. — URL: <https://about.nike.com/pages/nike-explore-team-sport-research-lab> (дата обращения: 15.12.2018).
- [8] Pelegris P. Banitsas K. Orbach T. Marias K. A Novel Method to Detect Heart Beat Rate Using a Mobile Phone. — 2010.
- [9] Volley. A library for asynchronous network requests by Google. — 2019. — URL: <https://developer.android.com/training/volley> (дата обращения: 11.05.2019).
- [10] Здоровье и фитнес трекер. Приложение // Google Play Market. — 2017. — URL: <https://play.google.com/store/apps/details?id=com.droidinfinity.healthplus> (дата обращения: 15.12.2019).

- [11] Мисникова И.В., Ковалева Ю.А. Влияние физической нагрузки на обменные процессы у пациентов с метаболическим синдромом. — Русский медицинский журнал, 2018. — URL: <https://elibrary.ru/item.asp?id=35663868>.
- [12] Нувальцева Е.П. Функциональная активность человека и взаимосвязь физической и умственной деятельности. Утомление и восстановление. Вклад молодых ученых в аграрную науку: Материалы международной научно-практической конференции. — 2018. — URL: <https://elibrary.ru/item.asp?id=36337288>.
- [13] СПбНИИФК ФГБУ. Научно-исследовательский институт физической культуры. — 2019. — URL: <http://www.spbniifk.ru/> (дата обращения: 11.05.2019).
- [14] Ташкинов Михаил Ильич. Оценка пульса с помощью камеры мобильного телефона. — 2016.