

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное
образовательное учреждение высшего образования
«Южно-Уральский государственный университет
(национальный исследовательский университет)»

Высшая школа электроники и компьютерных наук
Кафедра «Электронные вычислительные машины»

ДОПУСТИТЬ К ЗАЩИТЕ
Заведующий кафедрой ЭВМ
_____ Г.И. Радченко
« ___ » _____ 2020 г.

Разработка внешнего модуля голосового управления элементами "умного" дома

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
К ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЕ

Руководитель работы,
к.т.н., доцент каф. ЭВМ
_____ Д.В. Топольский
« ___ » _____ 2020 г.

Автор работы,
студент группы КЭ-405
_____ Г.Ю. Григорьев
« ___ » _____ 2020 г.

Нормоконтролёр,
ст. преп. каф. ЭВМ
_____ С.В. Сяськов
« ___ » _____ 2020 г.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное
образовательное учреждение высшего образования
«Южно-Уральский государственный университет
(национальный исследовательский университет)»
Высшая школа электроники и компьютерных наук
Кафедра «Электронные вычислительные машины»

УТВЕРЖДАЮ

Заведующий кафедрой ЭВМ

_____ Г.И. Радченко

«___» _____ 2020 г.

ЗАДАНИЕ

на выпускную квалификационную работу бакалавра
студенту группы КЭ-405
Григорьев Александр Юрьевич
обучающемуся по направлению
09.03.01 «Информатика и вычислительная техника»

1. **Тема работы:** «Разработка внешнего модуля голосового управления элементами "умного" дома» утверждена приказом по университету от 24 апреля 2020 г. №627.
2. **Срок сдачи студентом законченной работы:** 1 июня 2020 г.
3. **Исходные данные к работе:**
Спроектировать программно-аппаратный комплекс с возможностью:
 1. Передача аудиосигнала с модуля на сервер.
 2. Распознавания голосовых команд.
 3. Интерпретация команд в инструкции управления устройствами умного дома.

4. **Перечень подлежащих разработке вопросов:**

- анализ предметной области;
- определение требований к системе;
- проектирование программно-аппаратного комплекса;
- разработка программной и аппаратной части комплекса;
- тестирование программно-аппаратного комплекса.

5. **Дата выдачи задания:** 1 декабря 2019 г.

Руководитель работы _____ /Д.В. Топольский/

Студент _____ /А.Ю. Григорьев/

КАЛЕНДАРНЫЙ ПЛАН

Этап	Срок сдачи	Подпись руководителя
Введение и анализ предметной области	01.03.2020	
Определение требований к системе	01.04.2020	
Проектирование программно-аппаратного комплекса	15.04.2020	
Разработка программной и аппаратной части комплекса	01.05.2020	
Тестирование программно-аппаратного комплекса	15.05.2020	
Компоновка текста работы и сдача на нормоконтроль	24.05.2020	
Подготовка презентации и доклада	30.05.2020	

Руководитель работы _____ / *Д.В. Топольский* /

Студент _____ / *А.Ю. Григорьев* /

Аннотация

А.Ю. Григорьев. Разработка внешнего модуля голосового управления элементами "умного" дома. – Челябинск: ФГАОУ ВО «ЮУрГУ (НИУ)», ВШЭКН; 2020, 43 с., 13 ил., библиогр. список – 16 наим.

В рамках выпускной квалификационной работы производится детальный анализ современных систем управления устройствами умного дома. Организуется разработка программно-аппаратного комплекса распознавания голосовых команд и интерпретации их в инструкции для устройств умного дома. Производится выборка средств разработки программного обеспечения и аппаратных решений. Созданы прототип аппаратного комплекса и программное обеспечение. Производится анализ результатов работы системы, в домене специально разработанных задач. Созданы прототип аппаратного комплекса и программное обеспечение. Производится анализ результатов работы системы, в домене специально разработанных задач.

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	8
1 АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ	9
1.1 ОБЗОР АНАЛОГОВ	9
1.2 АНАЛИЗ ОСНОВНЫХ ТЕХНОЛОГИЧЕСКИХ РЕШЕНИЙ.....	12
1.2.1 Микроконтроллер.....	12
1.2.2 Микрофон.....	13
1.2.3 Хаб умного дома	14
1.2.4 Технология распознавания речи.....	15
1.2.3 Технологии для программирования микроконтроллера.....	16
1.3 ВЫВОД	17
2 ОПРЕДЕЛЕНИЕ ТРЕБОВАНИЙ К СИСТЕМЕ	18
2.1 ФУНКЦИОНАЛЬНЫЕ ТРЕБОВАНИЯ.....	18
2.3 ТЕХНОЛОГИЧЕСКИЕ ТРЕБОВАНИЯ	20
2.4 ТРЕБОВАНИЯ К ДОКУМЕНТАЦИИ.....	20
2.5 ВЫВОД	21
3 ПРОЕКТИРОВАНИЕ ПРОГРАММНО-АППАРАТНОГО КОМПЛЕКСА	22
3.1 АРХИТЕКТУРА ПРЕДЛАГАЕМОГО РЕШЕНИЯ.....	22
3.2 ОПИСАНИЕ ДАННЫХ.....	22
3.3 ВЫВОД	23
4 РАЗРАБОТКА ПРОГРАММНОЙ И АППАРАТНОЙ ЧАСТИ КОМПЛЕКСА.....	24
4.1 СБОРКА ПРОТОТИПА.....	24
4.1 РЕАЛИЗАЦИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ	28
4.1.1 Программное обеспечение модуля.....	28
4.1.1 Сервер.....	31

4.2 ВЫВОД.....	31
5 ТЕСТИРОВАНИЕ ПРОГРАММНО-АППАРАТНОГО КОМПЛЕКСА	32
5.1 МЕТОДОЛОГИЯ ТЕСТИРОВАНИЯ.....	32
5.2 ПРОВЕДЕНИЕ ПРОЦЕДУРЫ ТЕСТИРОВАНИЯ	32
5.3 ВЫВОД.....	34
ЗАКЛЮЧЕНИЕ.....	35
БИБЛИОГРАФИЧЕСКИЙ СПИСОК.....	36
ПРИЛОЖЕНИЕ А ИСХОДНЫЙ КОД ARDUINO СКРИПТА	38
ПРИЛОЖЕНИЕ Б ИСХОДНЫЙ КОД СЕРВЕРНОГО ПРИЛОЖЕНИЯ	42
ПРИЛОЖЕНИЕ В СХЕМА СБОРКИ МОДУЛЯ	44

ВВЕДЕНИЕ

Голосовые ассистенты в значительной мере упростили взаимодействие с системами умного дома. Данные системы в основном распространены на мобильных устройствах и «умных» колонках. В течение последнего десятилетия многие технологические компании стремятся к развитию голосового управления в качестве интерфейса управления различными системами.

Со временем возможность использования голосового управления становится наиболее ожидаемой функцией. Более 50% пользователей голосовых ассистентов используют их дома. Как для развлечения, так и для управления умным домом.

Цель исследования: Создать программно-аппаратный комплекс для интеграции в системы умного дома, способное распознавать голосовые команды и интерпретировать их в инструкции для устройств умного дома

Задачи необходимые для достижения цели:

- проанализировать рынок на предмет подобных устройств, выявить их качества;
- проанализировать стандарты, протоколы и устройства умного дома, выявить способы взаимодействия с ними;
- подобрать микроконтроллер и платформу, которые оптимально подходит для требований к таким системам;
- собрать рабочий прототип устройства, разработать необходимое для интеграции ПО.

1 АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ

1.1 ОБЗОР АНАЛОГОВ

Сейчас существуют 3 решения от крупных корпораций. Но не все из них удовлетворяют по разному ряду причин.

1.1.1. Apple Homepod

«Умная» колонка, разработанная компанией Apple, снабженная голосовым управлением, и использующая в своей работе облачный персональный голосовой помощник Siri работающий по принципу вопросно-ответной системы [1].

Достоинства:

- наличие программной платформы умного дома Apple HomeKit;
- семи-микрофонный массив для распознавания голоса.

Недостатки:

- не продаётся в России;
- не имеет интерфейса, переведенного на русский язык;
- высокая стоимость, \$349.

1.1.2. Google Home

«Умная» колонка, разработанная в Google, также снабженная персональным голосовым помощником Google Assistant [2].

Достоинства:

- наличие программной платформы умного дома Nest;

Недостатки:

- не продаётся в России;
- не имеет интерфейса, переведенного на русский язык;
- высокая стоимость \$139.

1.1.3. Яндекс.Станция

«Умная» колонка, разработанная компанией Яндекс, которая использует в своей работе голосовой помощник Яндекс.Алиса работающий по принципу вопросно-ответной системы. [3]

Достоинства:

- наличие программной платформы умного дома на базе помощника Алиса;
- Семи-микрофонный массив для распознавания голоса;
- Видео-интерфейс HDMI.

Недостатки:

- Высокая стоимость Р10 990.

1.1.4. Amazon Echo

«Умная» колонка, разработки корпорации Amazon.com, снабженная голосовым ассистентом Amazon Alexa [4].

Достоинства:

- Наличие программной платформы умного дома «Smart Home Skill»;
- Семи-микрофонный массив для распознавания голоса;
- Относительно дешевая стоимость \$99.

Недостатки:

- Не имеет интерфейса, переведенного на русский язык.

Результаты сравнения в таблице 1.1.

Таблица 1.1 – Преимущества и недостатки аналогов:

Критерий	Apple Homepod	Google Home	Яндекс.Станция	Amazon Echo
Наличие программной платформы умного дома	+	+	+	+
Видео-интерфейс HDMI	-	-	+	-
Продвинутый микрофонный массив	+	-	+	+
Интерфейс на русский языке	-	-	+	-
Стоимость	\$349	\$139	₽10 990	\$99

1.2 АНАЛИЗ ОСНОВНЫХ ТЕХНОЛОГИЧЕСКИХ РЕШЕНИЙ

1.2.1 Микроконтроллер

Для распознавания речи необходимо решить проблему передачи аудиопотока на сервер и принятия инструкций.

Для передачи большого количества трафика различного вида в домашних условиях, наиболее распространённой технологией беспроводного доступа на данный момент является стандарт беспроводных локальных сетей IEEE 802.11

В качестве микроконтроллера для отправки записанного звука на сервер используется микроконтроллер ESP32. Микроконтроллер поставляется с интегрированным WiFi-модулем для отправки данных. Благодаря этому не придётся использовать аналоги и подключать к ним внешние модули WiFi.

На данный момент популярным и дешёвым решением будет сборка Wi-Fi ретранслятора на базе модуля ESP32 (рисунок 1.2.1).

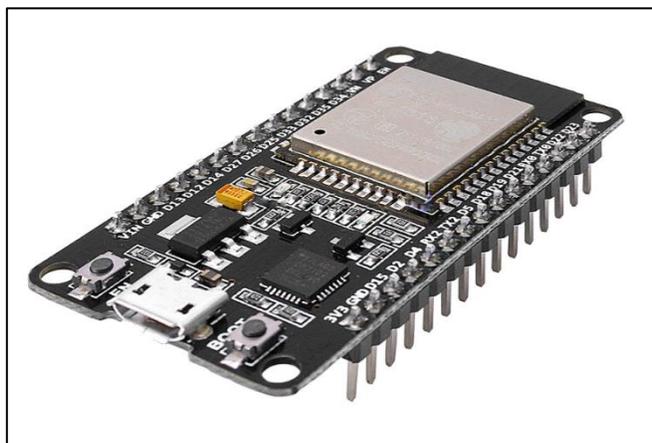


Рисунок 1.2.1 – Микроконтроллер “ESP32”

Особенности:

- поддержка беспроводного стандарта 802.11 b/g/n;
- поддержка 2 режима работы Wi-Fi Direct (P2P), soft-AP;
- интегрирован стек протокол TCP/IP;

- поддержка подключения нескольких TCP Client.

1.2.2 Микрофон

Для записи звука нужен микрофон, способный записывать звук на достаточно большом расстоянии в бытовых условиях. Для решения этой задачи был выбран микрофон МАХ9814 с автоматической регулировкой усиления (рисунок 1.2.2).

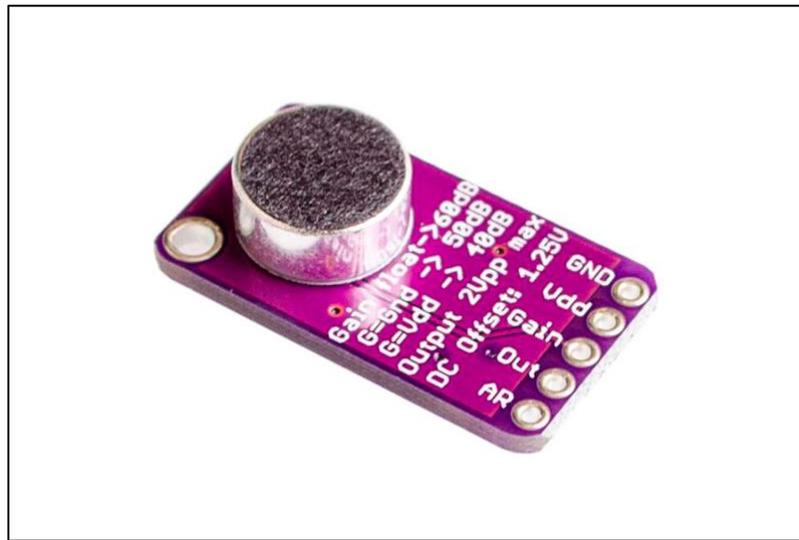


Рисунок 1.2.2 – Микрофон “МАХ9814”

Особенности:

- автоматический контроль усиления (AGC);
- три настройки усиления (40 дБ, 50 дБ, 60 дБ);
- программируемое время атаки;
- частотный диапазон 20 - 20000 кГц;
- коэффициент гармоник 0.04%;
- входная плотность шума 30 нВ.

1.2.3 Хаб умного дома

В данном проекте обработка звука с модуля должна производиться на сервере, который, также, должен исполнять роль хаба устройств умного дома. Хаб нужен для контроля элементов системы умного дома, а также распознавания речи и интерпретации их в команды для управления устройствами умного дома.

В качестве технологии сервера (хаба) умного дома выбран OpenHUB, преимуществами которого являются [5]:

- python-библиотека для взаимодействия;
- возможность интегрирования с более чем 200 технологий и систем;
- можно запустить на: операционных системах для ПК: Linux, macOS, Windows; одноплатных компьютерах: Raspberry Pi, PINE64; систем виртуализации: Docker, Synology.

На сервер операционной системой выбран Linux дистрибутив Ubuntu. Данный дистрибутив, в отличие, от операционной системы Windows, бесплатный.

Популярные языки программирования на Linux дистрибутивах:

- Shell;
- C/C++;
- Java;
- Python.

Для управления элементами системы умного дома с помощью фреймворка OpenHUB выбираем язык Python [6].

1.2.4 Технология распознавания речи

Для распознавания речи должна использоваться технология, которую можно развернуть на операционной системе Linux и которая распространяется по свободной лицензии.

Основным показателем качества для систем автоматического распознавания речи является точность распознавания, которая, как правило, определяется как процент неправильно распознанных слов (WER – Word Error Rate). Этот показатель на данный момент чаще всего используется как показатель точности работы системы распознавания речи, когда сравниваются разные системы распознавания речи между собой. Формула для вычисления WER:

$$WER = \frac{S + D + I}{T}$$

где, S — количество “подмен” D — количество пропусков I — количество лишних слов N — количество слов в выборке [7].

В качестве технологии распознавания речи был выбран фреймворк Kaldi, после сравнения характеристик WER с другими технологиями.

Сравнение метрик Kaldi и наиболее популярных технологий удалённого распознавания речи: Yandex Speech Kit и Google Speech API, представлено на таблице 1.2.1.

Таблица 1.2.1 – Сравнительный анализ WER Kaldi и популярных технологий удалённого распознавания речи.

	Yandex Speech Kit	Kaldi ASR(kaldi-ru-0.6)	Google Speech API
WER	0,3	0,36	0,31

Результаты WER были получены после распознавания речевого корпуса VoxForge, после распознавания каждой системой 3677 звуковых файлов [8].

Сравнение метрик Kaldi и наиболее популярных технологий распознавания речи с открытым исходным кодом: CMU Sphinx, HTK, iAtros, Julius, Kaldi и RWTH ASR.

Таблица 1.2.2 – Сравнительный анализ WER Kaldi и популярных технологий распознавания речи с открытым исходным кодом.

	CMU Sphinx	HTK	iAtros	Julius	Kaldi	RWTH ASR
WER	0,214	0,198	0,16	0,231	0,065	0,155

Результаты WER были получены после распознавания речевого корпуса Wall Street Journal 1 [9].

Так же код Kaldi находится под лицензией Apache 2.0 которая является одной из наименее ограничительных лицензий [10].

Фреймворк Kaldi был подобран по метрике Word Error Rate и параметрам производительности. Он незначительно уступает популярных технологий удалённого распознавания речи и выигрывает по сравнению с аналогичными популярными технологиями распознавания речи с открытым исходным кодом.

1.2.3 Технологии для программирования микроконтроллера

Для написания программы микроконтроллера был выбран язык Arduino Language. Причиной выбора данной технологии является наличие в языке библиотек для работы с Wi-Fi сетями, в частности WiFi.h и Autoconnect.h [11].

Также разработчики данного языка предоставляют компилятор и программатор для микроконтроллеров ESP32.

1.3 ВЫВОД

Технологии распознавания речи на данный момент достаточно развиты и имеют широкий спектр аналогов. Следовательно, создание программно-аппаратного комплекса для распознавания голосовых команд является актуальной задачей. Для ее решения выбраны следующие технологические решения: для получения звукового сигнала и отправки его по Wi-Fi – микроконтроллер ESP32 и микрофон MAX9814, для распознавания речи фреймворк Kaldi. Для взаимодействия с системами умного дома выбран фреймворк OpenHUB.

Для разработки ПО выбраны языки программирования Python и Arduino Language. Подобраны бесплатные средства разработки Visual Studio Code и Arduino IDE.

2 ОПРЕДЕЛЕНИЕ ТРЕБОВАНИЙ К СИСТЕМЕ

2.1 ФУНКЦИОНАЛЬНЫЕ ТРЕБОВАНИЯ

На рис. 2.1 показана схема взаимодействия объектов системы. Модуль голосового управления отправляет аудио на компьютер. На компьютере аудио обрабатывается и интерпретируется в команды управления устройствами умного дома. Команды отправляются на устройство для выполнения. На модуль отправляется информация о результате выполнения команды.

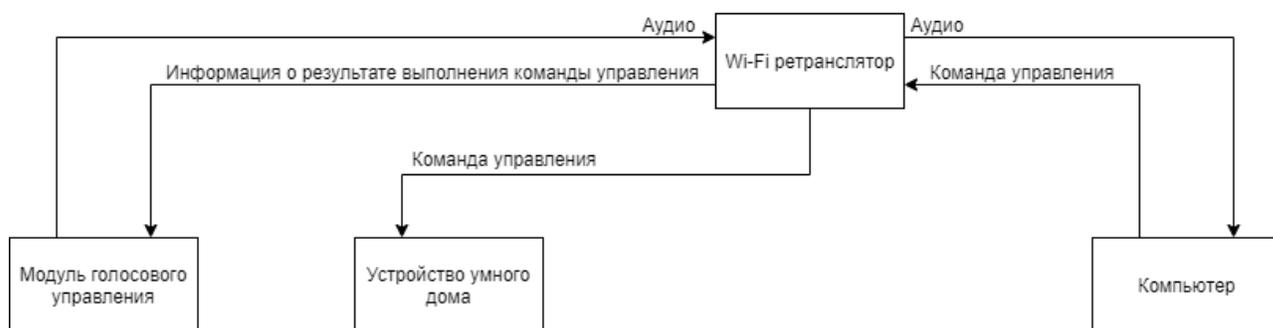


Рисунок 2.1 – Схема взаимодействия объектов программно-аппаратного комплекса.

На рисунке 2.2 показаны основные варианты использования программно-аппаратного комплекса.

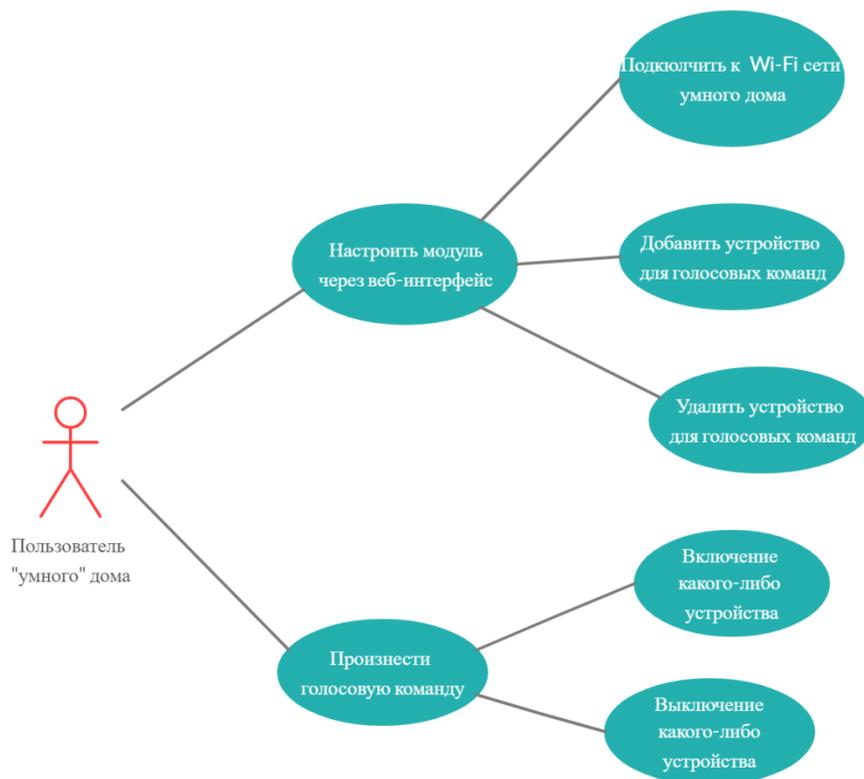


Рисунок 2.2 – Варианты взаимодействия с модулем

2.2 ТРЕБОВАНИЯ К ГОЛОСОВЫМ КОМАНДАМ

1. Активация распознавания команд должна производиться голосовой командой «Старт». Распознавание команд должно производиться в течение 5 секунд после произнесения команды «Старт».

2. Включение устройства должно производиться командой «включить «название устройства», где «название устройства» – название которое добавлено в веб-интерфейсе настройки (см. пункт 1.2.5).

3. Выключение устройства должно производиться командой «выключить «название устройства», где «название устройства» – это название которое добавлено в веб-интерфейсе настройки (см. пункт 1.2.5).

4. При невозможности выполнения команды устройство должно подавать сигнал.

5. Новая голосовая команда не может быть произнесена раньше чем через 5 секунд после произнесения прошлой голосовой команды.

2.3 ТЕХНОЛОГИЧЕСКИЕ ТРЕБОВАНИЯ

Требования к модулю:

1. Размеры устройства: не более 100(Ш)х100(Д)х30(В) мм;
2. Вес устройства: не более 200 г;
3. Микрофон с усилением более 20 дБ и автоматической регулировкой усиления;
4. WI-FI: 5 ГГц;
5. Питание от сети 220 В.

Требования к серверу:

1. Процессор: архитектура: x86, x64;
2. Частота: от 1 Гц;
3. ОЗУ: от 1 Гб;
4. ПЗУ: от 4 Гб;
5. ОС: Linux;
6. ПО: Интерпретатор языка Python.

2.4 ТРЕБОВАНИЯ К ДОКУМЕНТАЦИИ

В документации на устройство должны содержаться технические характеристики устройства, которые включают следующие требования: вес, размеры, тип источника питания, класс пыле- и влаго- защита. Документация должна предоставляться Заказчику в печатном виде на бумажных носителях, а

также в электронном виде, в виде файлов формата MS Word (DOC) или Adobe Acrobat Reader (PDF), пригодных для печати. Документация должна быть на русском языке.

2.5 ВЫВОД

В данной главе были определены ключевые требования к разрабатываемой системе.

В частности требования по взаимодействию модуля с пользователем и другими устройствами умного дома, а так же технологические требования системы.

3 ПРОЕКТИРОВАНИЕ ПРОГРАММНО-АППАРАТНОГО КОМПЛЕКСА

3.1 АРХИТЕКТУРА ПРЕДЛАГАЕМОГО РЕШЕНИЯ

На микроконтроллере ESP32 имеется 36-пиновая рейка портов ввода/вывода общего назначения (Рисунок 3.1.1).

Спроектируем схему подключений модулей Рисунок В.1 приложения В.

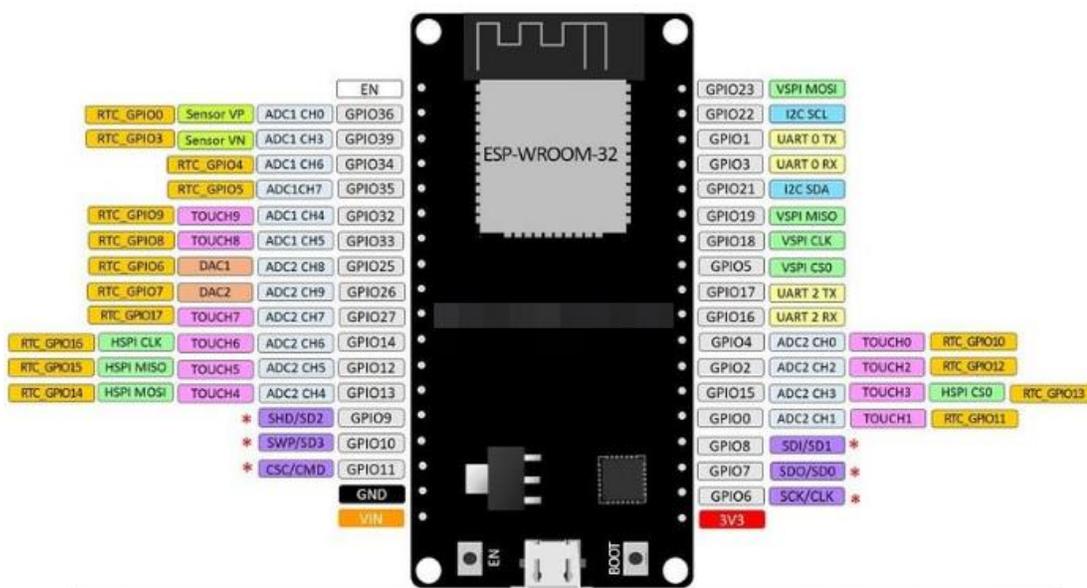


Рисунок 3.1.1 – Схема портов ESP32

3.2 ОПИСАНИЕ ДАННЫХ

Составим схему данных для нашего проекта. Структурная схема данных системы представлена на рисунке 3.3.

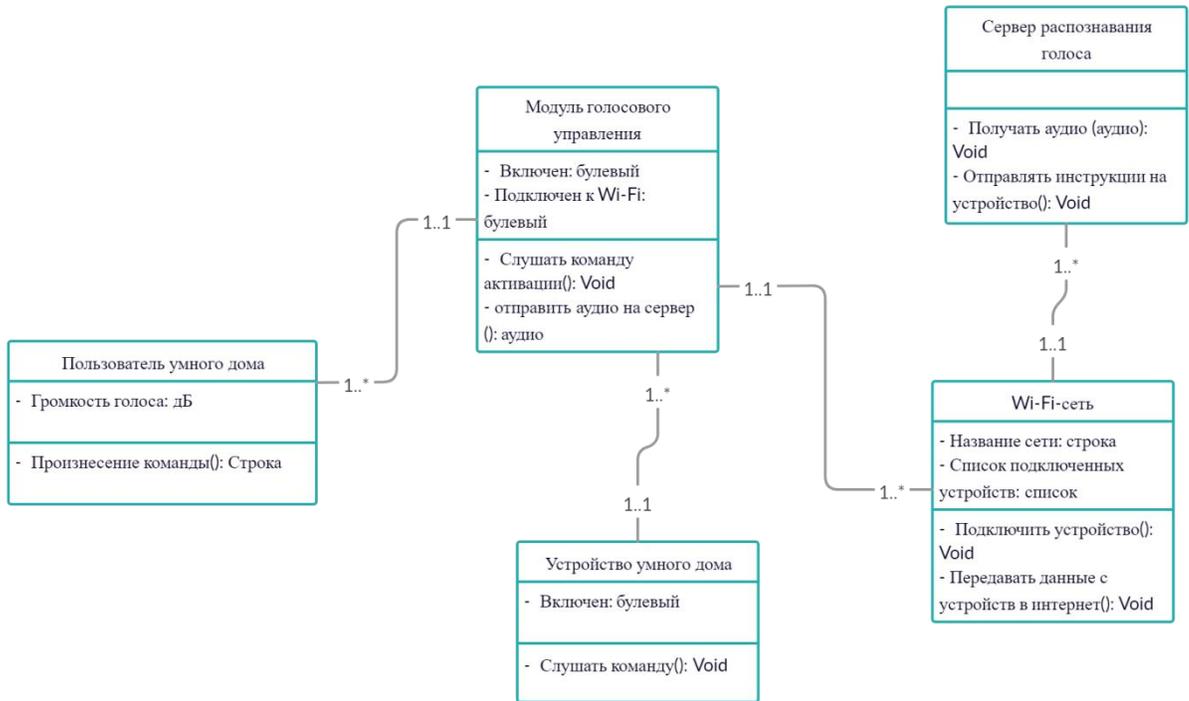


Рисунок 3.3 – Структурная схема данных системы.

3.3 ВЫВОД

По итогу данной главы, была определена схема подключения модулей в устройстве и была спроектирована структурная схема данных системы

4 РАЗРАБОТКА ПРОГРАММНОЙ И АППАРАТНОЙ ЧАСТИ КОМПЛЕКСА

4.1 СБОРКА ПРОТОТИПА

Для проверки в действии работы системы необходимо собрать прототип, в котором будут использоваться упомянутые выше платы. Заказ оборудования производился из российского магазина «ЧИП и ДИП»[12]. Не найденное на территории России оборудование было заказано с иностранного онлайн-магазинов «AliExpress»[13].

Блок питания GS06E-1P1J



Рисунок 4.1.1 – Блок питания GS06E-1P1J

Таблица 4.1.1 – Технические характеристики Блока питания GS06E-1P1J

Габариты	32 x 66 x 42,5 мм
Уровень пульсаций (размах)	50 мВ
КПД	70%
U _{вых}	5 В
I _{вых}	0...1 А
Стоимость	820 руб.
Источник	https://www.chipdip.ru/product/gs06e-1p1j

Микроконтроллер “Espressif ESP32”



Рисунок 4.1.2 – Микроконтроллер “Espressif ESP32”

Технические характеристики модуля Espressif ESP32.

Таблица 4.1.2 – Технические характеристики Микроконтроллера “Espressif ESP32”

Питание	2,2...3,6 Вольт
Процессор	Двухъядерный 32-бит Tensilica Xtensa® LX6 с FPU и MAC. 240 МГц (600 DMIPS)
Память	448 кБайт ПЗУ, 520 кБайт ОЗУ. Внешние ОЗУ/ПЗУ на SPI интерфейсе, до 4*16 МБайт. Внешняя память может быть криптографически защищена. [14]
Стоимость	930 руб.
Источник	https://www.chipdip.ru/product1/8002969391

Микрофон “МАХ9814”

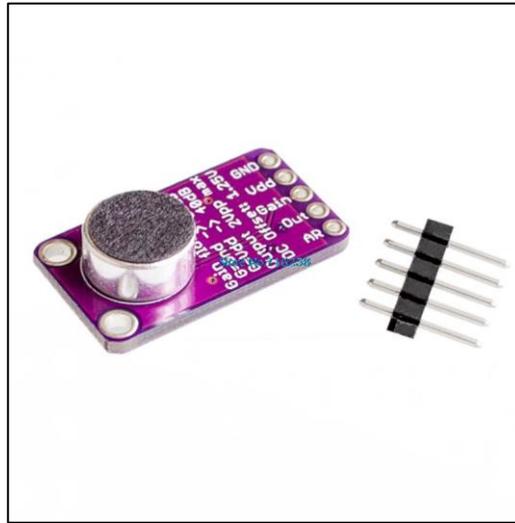


Рисунок 4.1.3 – Микрофон “МАХ9814”

Таблица 4.1.3 – Технические характеристики Микроконтроллера “Espressif ESP32”

Питание	2,7...5,5 Вольт
Частотная характеристика	20 Гц - 20 кГц [15]
Выход	2В при смещении 1,25В
Стоимость	64 руб.
Источник	https://www.aliexpress.com/snapshot/0.html?cv=815660&orderId=5001607696324465&af=857570&utm_campaign=857570&aff_platform=link-c-tool&utm_medium=cpa&sk=mr6wR6CC&aff_trace_key=5a42b2be01f0491ab3a3f6d54f1f1336-1590820755563-08797-mr6wR6CC&dp=7421e41f03fd96140e54f3aa40b329c6&terminal_id=bd5b4786364c47a193af217d2fdbb038&utm_source=admitad&utm_content=815660&aff_request_id=5a42b2be01f0491ab3a3f6d54f1f1336-1590820755563-08797-mr6wR6CC

Для проверки работоспособности системы и взаимодействия её ключевых компонентов реализован прототип модуля.

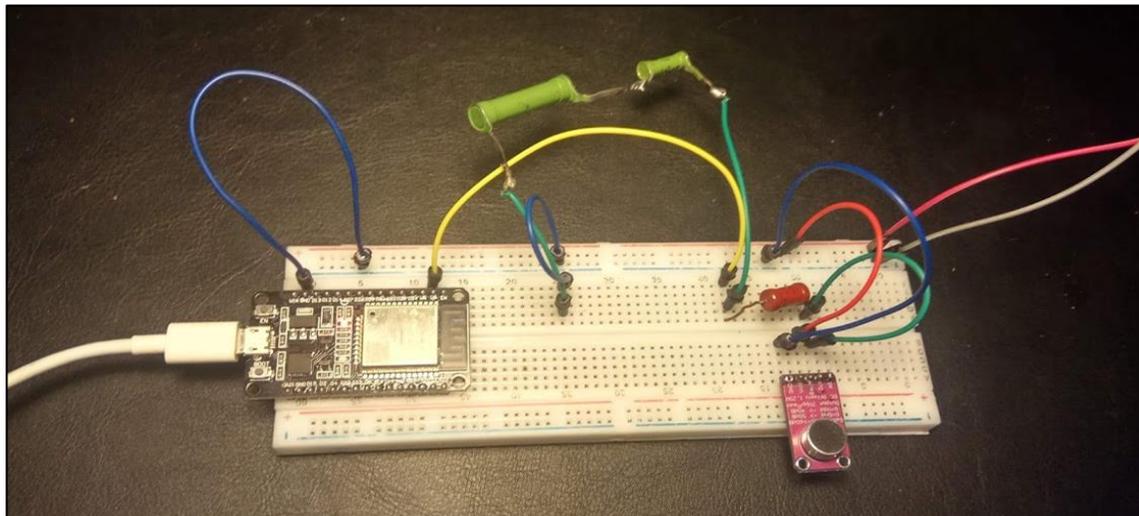


Рисунок 4.1.4 – Прототип

Также была подсчитана сумма расходов на разработку прототипа (Таблица 4.1.4). Данная цена не является конечной. При подсчёте не сделана оценка затрат материала для сборки корпуса и платы.

Таблица 4.1.4 – Экономическая составляющая модуля

Тип	Описание	Цена, руб.	Количество, шт.	Всего, руб.
Блок питания	Напряжение 5 В, сила тока 1 А.	820	1	820
Микрофон	Электретный микрофон МАХ9814 с усилением до 60 дБ и автоматической регулировкой усиления	64	1	64

Продолжение таблицы 4.1.4

Микроконтроллер	Микроконтроллер ESP32 с встроенным модулем Wi-Fi	930	1	930
Динамик	Динамик с громкостью 30 дБ	20	1	20
Итого				1 834

4.1 РЕАЛИЗАЦИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

4.1.1 Программное обеспечение модуля

На момент запуска модуля модуль должен сканировать сети Wi-Fi и предлагать подключиться к сети в которой предполагается работа устройства.

Чтобы выбрать сеть, к которой должно подключаться устройство, пользователь должен выбрать её в специальном веб-интерфейсе.

Для того чтобы пользователь мог подключиться к веб-интерфейсу выбора сети, устройство создаёт временную Wi-Fi сеть, подключившись к которой, пользователь вводит IP-адрес в браузере устройства и переходит на страницу выбора сети.

На рисунке 4.1.1.1 показана временная Wi-Fi сеть «esp32ap», которую создаёт модуль.

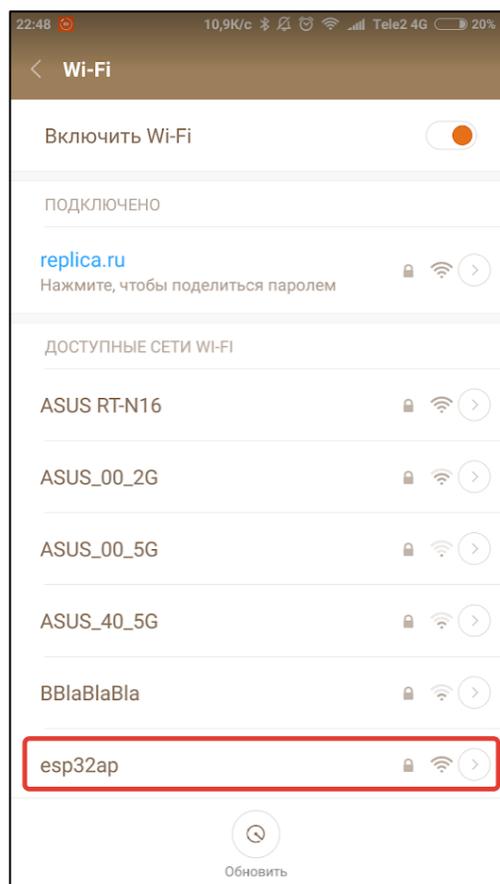


Рисунок 4.1.1.1 – Временная Wi-Fi сеть «esp32ap»

Далее, подключившись к сети, автоматически открывается страница подключения к Wi-Fi сети, где можно выбрать нужную сеть для подключения. Выбор сети показан на рисунке 4.1.1.2

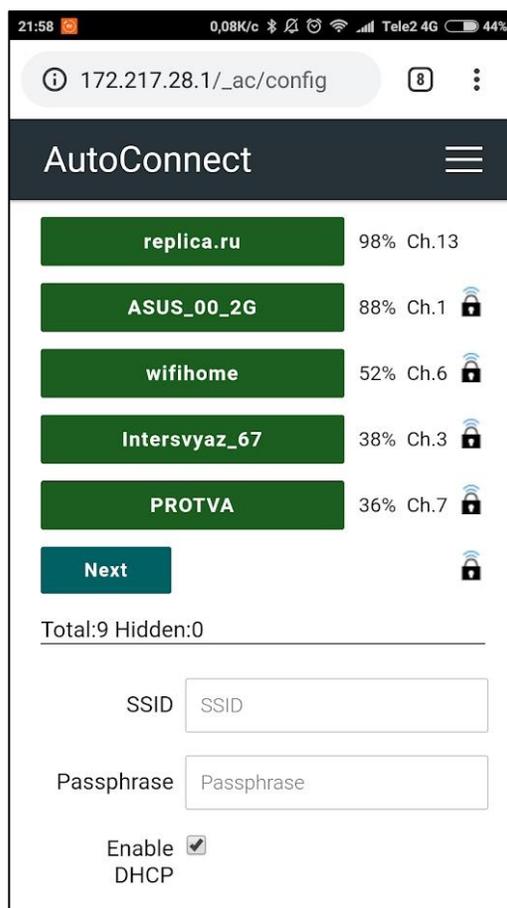


Рисунок 4.1.1.2 – Выбор Wi-Fi сети.

После подключения к сети, устройство готово записывать аудио и отсылать его на сервер для распознавания речи. Экран подключения к сети показан на рисунке 4.1.1.3.

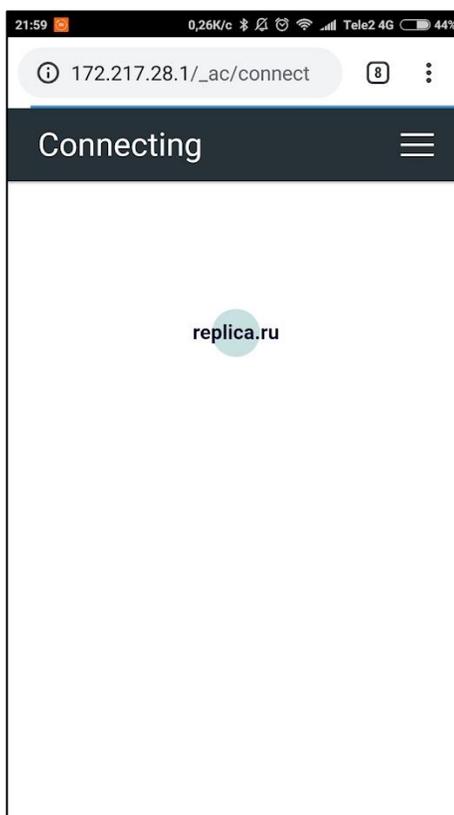


Рисунок 4.1.1.3

Реализация кода программного обеспечения модуля приведена в листинге А.1 приложения А.

4.1.1 Сервер

Сервер распознаёт аудиопоток, который приходит на заданный порт хаба и распознаёт в них голосовые команды. Программное обеспечение сервера прилагается в листинге Б.1 приложения Б.

4.2 ВЫВОД

В данной главе было реализовано программное обеспечение модуля для подключения к сети Wi-Fi и передачи данных на хаб умного дома. А так же реализована серверная часть комплекса.

5 ТЕСТИРОВАНИЕ ПРОГРАММНО-АППАРАТНОГО КОМПЛЕКСА

5.1 МЕТОДОЛОГИЯ ТЕСТИРОВАНИЯ

Для тестирования программно-аппаратного комплекса было проведено альфа-тестирование. В результате тестирования было уделено большее внимание серверной части по части распознавания голоса.

Под альфа-тестированием понимается как имитация реальной деятельности программно-аппаратного комплекса, так и реальная работа с системой потенциальными пользователями.

Основная цель альфа-тестирования – получить более подробную информацию о готовом продукте, в частности узнать о дефектах и предоставить организации-разработчику обратную связь и информацию о дефектах от группы людей, превосходящей тестеров в количестве, без изменения репутации пока не выпущенного продукта. [16].

5.2 ПРОВЕДЕНИЕ ПРОЦЕДУРЫ ТЕСТИРОВАНИЯ

Проведём тестирование серверной части приложения с помощью набора записей телефонных разговоров, что по качеству звука и по количеству шумов аналогично записываемого с модуля звука. Общая продолжительность записей – 30 часов.

Распознавание файлов происходило на тестовом сервере с помощью python-скрипта.

Листинг 5.2.1 Тестирование

```
#!/usr/bin/python
import os
import sys
import wave
import json
```

```

from vosk import Model, KaldiRecognizer

model_dir = os.path.join('../models/kaldi-ru-0.7/')
wavs_folder_dir = os.path.join('../data/raw/calls/')
rseults_folder = os.path.join('../data/interim/vosk_kaldi-
0.7_results_of_recognition/')

model = Model(model_dir)

for wav in os.listdir(wavs_folder_dir):
    wav_path = os.path.join(wavs_folder_dir + wav)
    wf = wave.open(wav_path, 'rb')
    rec = KaldiRecognizer(model, wf.getframerate())

    recognised_str = ''
    while True:
        data = wf.readframes(1000)
        if len(data) == 0:
            break
        if rec.AcceptWaveform(data):
            recognised_str += json.loads(rec.Result())['text'] + ' '

    file_txt_path = os.path.join(rseults_folder + wav[0:-4] + '.txt')
    with open(file_txt_path, 'w') as file_txt:
        file_txt.write(recognised_str)

df = pd.read_csv('../data/raw/filtered_final.csv', sep=',')
df.drop(df.columns[[0,1,2,3,4,7]], axis=1, inplace=True)
df['kaldi'] = ""

# Путь до папки с результатами распознавания
results_folder =
os.path.join('../data/interim/results_of_recognition(kaldi-0.7)(vosk)/')

for file_name in os.listdir(results_folder):
    file_path = os.path.join(results_folder, file_name)
    with open(file_path, 'r') as file_txt:
        recognized_str = file_txt.readline()
        recognized_str = recognized_str.replace('\n', ' ').replace(r'\s',
' ').strip()

        wav_name = file_name[0:-3]+"wav"
        if (not df[df['path'] == wav_name].empty):
            df.loc[df[df['path'] == wav_name].index[0], 'kaldi'] =
recognized_str

df.to_csv('../data/processed/text_kaldi-0.7(vosk).csv')

```

Итоговое значение WER = 0.4955446865273455, что достаточно качественно для распознавания команд.

Пример ошибки:

Ошибочная строка:

«здравствуйте у меня вот такая проблема вчера дверь меняли входную и как бы повредили кабель от телевизора можно как то в страницу проблему»

Реальная фраза:

«здравствуйте у меня вот такая проблема вчера дверь меняли входную и как бы повредили кабель от телевизора можно как то устраниться проблему»

5.3 ВЫВОД

Весомых отклонений не наблюдается, значит тестирование распознавания речи прошло успешно.

ЗАКЛЮЧЕНИЕ

В ходе выполнения представленной выпускной квалификационной работы было проделано следующее:

- проведен анализ существующих устройств распознавания голосовых команд умного дома;
- определен набор основных функций и требований к разрабатываемой системе;
- разработана архитектура и функциональный состав системы;
- выбраны средства реализации и платформа для функционирования системы;
- разработана серверная часть системы;
- разработана клиентская часть системы;
- произведена сборка прототипа устройства;
- проведено тестирование работы серверной части;

Результатом выполненной работы является спроектированная система распознавания речи и управления устройствами умного дома.

В настоящее время разработка продолжается, с целью расширения функционала программного-аппаратного комплекса, улучшения качества распознавания речи. Основные функции, заявленные в данной работе, подготовлены к эксплуатации.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- 1 Google Home. –
https://store.google.com/us/product/google_home_speaker?hl=en-US. – Дата обращения: 01.05.2020.
- 2 Apple Homepod. – <https://www.apple.com/homepod/>. – Дата обращения: 01.05.2020.
- 3 Яндекс.Станция. – <https://station.yandex.ru/>. – Дата обращения: 01.05.2020.
- 4 Amazon Echo. – <https://www.amazon.com/gp/product/B00X4WHP5E/>. – Дата обращения: 02.05.2020.
- 5 OpenHab. – <https://www.openhab.org/docs/installation/linux.html/> . – Дата обращения: 08.05.2020
- 6 python-openhab 2.13.0. – <https://pypi.org/project/python-openhab> . – Дата обращения: 08.05.2020
- 7 Тампель, И. Б. Автоматическое распознавание речи : учебное пособие / И. Б. Тампель, А. А. Карпов. — СПб: НИУ ИТМО, 2017. — 152 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/110433> (дата обращения: 25.05.2020).
- 8 Habr. Почему Kaldi хорош для распознавания речи? –
<https://habr.com/ru/post/470696/> . – Дата обращения: 08.05.2020
- 9 Беленко, М. В. СРАВНИТЕЛЬНЫЙ АНАЛИЗ СИСТЕМ РАСПОЗНАВАНИЯ РЕЧИ С ОТКРЫТЫМ КОДОМ / М. В. Беленко, П. В. Балакшин // Международный научно-исследовательский журнал. — 2017. — № 04 (58) Часть 4. — С. 13—18. — URL: <https://research-journal.org/technical/sravnitelnyj-analiz-sistem-raspoznavaniya-rechi-s-otkryтым-kodom/> – Дата обращения: 01.05.2020

- 10 Системы распознавания речи с открытым исходным кодом. –
<https://lvee.org/en/abstracts/273/> . – Дата обращения: 20.05.2020
- 11 AutoConnect» An Arduino library for ESP8266/ESP32 WLAN configuration at run time with web interface. –
<https://play.google.com/store/apps/details?id=com.handcent.app.nextsms> . –
Дата обращения: 08.05.2020
- 12 ЧИП и ДИП. - <https://www.chipdip.ru/>. –Дата обращения: 23.05.2020.
- 13 AliExpress. – <https://aliexpress.com>. –Дата обращения: 23.05.2020
- 14 ESP32. – <https://www.espressif.com/en/products/socs/esp32/overview>. – Дата обращения: 10.06.2020
- 15 MAX9814. – <https://datasheets.maximintegrated.com/en/ds/MAX9814.pdf>. –
Дата обращения:10.05.2020
- 16 Генельт, А. Е. Учебно-методическое пособие по дисциплине «Управление качеством разработки ПО» : учебно-методическое пособие / А. Е. Генельт. — Санкт-Петербург : НИУ ИТМО, 2007. — 187 с. — Текст : электронный // Лань : электронно-библиотечная система. – Дата обращения: 10.05.2020