

## **СЕТЕВОЙ МНОГОПОЛЬЗОВАТЕЛЬСКИЙ ЧАТ С ГОЛОСОВЫМ АССИСТЕНТОМ**

Дубоделов А.В.

Научный руководитель - Белова С.В.

Цель работы – реализация сетевого чата на основе многопользовательского сервера с использованием технологии сокетов; а также внедрение голосового помощника для упрощения работы с приложением.

Разработка осуществлялась на языке программирования Python.

Данная тема выбрана неслучайно, информационный поток в современном мире растет с каждым днем, а вместе с этим возрастает потребность в обмене информацией, в удобном общении через современные средства коммуникации. С задачей обмена информацией хорошо справляются социальные сети и мессенджеры, чья востребованность отображается в количестве пользователей. Так в 2017 году сообщество WhatsApp преодолело отметку в 1 миллиард пользователей. Технологию голосового помощника в том виде, в котором он представлен сейчас, можно считать инновационной, так как первые современные ассистенты были выпущены во втором десятилетии XXI века, а их продажи уже составляют миллионы долларов. Так ассистент Google был запущен в 2016 году, а его продажи за 4 квартал 2018 года составили 11.5 миллионов штук [1].

Серверная часть программы представляет собой многопользовательский сервер. Он организывает единообразную систему принятия и рассылки сообщений клиентам. Для подключения клиента к серверу, второй принимает запрос на соединение от удаленного хоста(клиента). Каждому подключенному клиенту выделяется новый поток, что позволяет говорить о многопоточном сервере для данного приложения. Прием сообщений осуществляется путем прослушивания всех подключенных адресов. Перед отправкой сообщение кодируется по системе utf-8 и передается частями по 1024 байта.

Клиентский модуль представляет собой интерфейс с возможностью регистрации и входа в систему, отправки текстовых сообщений и смайликов (рисунок 1).

В программе также присутствует ряд проверок на авторизацию, существование пользователя в базе, попытку отправки сообщений неавторизованным пользователем и т.д.



Рисунок 1 – Интерфейс приложения

Для защиты от несанкционированного доступа к информации перед отправкой и при получении сообщения выполняется шифрование/расшифрование с использованием классического симметричного алгоритма. В последующих доработках и обновлениях предполагается внедрение более сложного вида шифрования с достаточной длиной ключа для обеспечения криптостойкости.

Чатом могут пользоваться и люди с ограниченными возможностями, например, с плохим зрением. Для этого все действия в программе снабжены звуковым сопровождением: звуками входа, выхода, отправки сообщения. А также внедрен голосовой помощник (рисунок 2) реализующий ряд дополнительных функций:

- чтение последнего сообщения в чате;
- запуск сторонних приложений, например, браузера;
- констатация текущего времени.



Рисунок 2 – Кнопка вызова голосового помощника

Голосовой помощник управляется только посредством звуковых команд. Реализация основана на технологии Google Cloud Speech Recognition API [2]. Что обеспечивает голосовому помощнику наивысшую оценку качества работы, а также позволяет распознать 120 языков и до четырех различных языков в одном запросе [2]. Используемая технология дает возможность внедрения machine learning и обучения помощника новым командам [2].

Фраза, произнесенная человеком, воспринимается помощником как звуковой сигнал с чередующимися гласными и согласными звуками. Поэтому для

распознавания речи голосовой помощник делит слова на звуковые фрагменты – фонемы. Для распознавания фонема, ассистент работает с их числовым представлением. Так, после записи запроса и его отправки на сервер, он делится на множество мелких фреймов длиной в сотые доли секунды. Далее фреймы проходят математическое преобразование, результатом этих преобразований является коэффициент, описывающий частотные характеристики фрейма. На основании этих данных подбирается фонем, которому принадлежит полученный фрейм.

Данная технология позволяет распознавать голос(аудио) и возвращать текст в режиме реального времени.

Также для лучшего распознавания речи были внедрены две технологии:

- задержка записи звука;
- нечеткий поиск.

Первая из них позволяет помощнику определить звук фона, дабы лучше вычлнять речь говорящего. Реализация данной функции осуществляется через предварительное прослушивание фона, которое занимает 1 секунду времени, до начала записи голоса диктора.

Технология нечеткого поиска позволяет распознавать нечеткую речь. Основывается данная технология на алгоритме расстояния Леванштейна, заключающегося в поиске наиболее схожего варианта действия, путем нахождения минимального количества изменений в данных, полученных от пользователя [3], для получения схожего варианта в системе помощника.

В результате реализации данного приложения был получен простейший сетевой чат, подходящий для использования в небольших компаниях, группах, обществах. После двухэтапного тестирования: программного, физического (тестировщиками), - можно говорить о готовности приложения первому релизу.

## Литература

1 BlogDTI. [Электронный ресурс] / Голосовые помощники: эволюция, устройство и основные игроки. – Режим доступа: <https://blog.dti.team/voice-assistants-1/> .– Дата доступа 13.05.2020, свободный – Загл. с экрана. – Яз.рус.

2 Google Cloud. [Электронный ресурс] / Speech-to-Text powered by machine learning. – Режим доступа: <https://cloud.google.com/speech-to-text?hl=ru> .– Дата доступа 12.05.2020, свободный – Загл. с экрана. – Яз.англ.

3 Habr. [Электронный ресурс] / Нечеткий поиск в тексте и словаре. – Режим доступа: <https://habr.com/ru/post/114997/> .– Дата доступа 12.05.2020, свободный – Загл. с экрана. – Яз.рус.