Anryze

Distributed Network

Speech Recognition Platform with Distributed Computer Network

Table of Content

1. Introduction

Anryze Distributed Network is a peer-to-peer distributed computer network for speech recognition and neural network education would allow users to transcribe audio files without reliance on a third party provider like Google or Amazon. The removal of central controls would mitigate most traditional data failures and provide the highest accuracy of a speech transcription due to neural networks education through distributed system.

Over the past 50 years people tried to make computers understand human language in order to make them more responsive to our needs and helpful, and to create a bigger amount of useful features. Starting from Bell Labs and IBM experiments, the most sufficient value in speech recognition was introduced by Google with the innovative Speech API. But all these solutions were built on the principles from 50 years ago and soonly will reach their limits. Thus, now it is time to embrace the brand new solutions and take computer speech recognition on the next level.

Why speech recognition matters? As soon as we educate computers to understand natural human language, it will become a revolution in a different areas of work. Developing, educating, and creating new neural network models is vital nowadays, because of the possibility to make computer actually think like a human being. The accurate speech recognition is the very first step to it. After all, this is how we communicate. Our mission is to create the best Speech Recognition System: highly educated, thinking like us, understanding all our words and intention, and while doing that, we want to get as much people as we can to participate. Afterwards, we will be proud to say that this is the collaborative act aimed on the education of the Computer that became possible not only due to the efforts of the small group of scientists but also with the contribution of usual people. We want them to feel as a part of this revolutionary process.

2. Speech Recognition

Basic Concept

Speech is a complex phenomenon. People rarely understand how is it produced and perceived. The naive perception is often that speech is built with words, and each word consists of phones. The reality is unfortunately very different. Speech is a dynamic process without clearly distinguished parts. It's always useful to get a sound editor and look into the recording of the speech and listen to it.

All modern descriptions of speech are to some degree probabilistic. That means that there are no certain boundaries between units, or between words. Speech to text translation and other applications of speech are never 100% correct. That idea is rather unusual for software developers, who usually work with deterministic systems. Therefore, it creates a lot of issues specific only to speech technology.

Structure of speech

In current practice, speech structure Anryze use as follows:
Speech is a continuous audio stream where rather stable states mix with dynamically changed states. In this sequence of states, one can define more or less similar classes of sounds, or phones. Words are understood to be built of phones, but this is certainly not true. The acoustic properties of a waveform corresponding to a phone can vary greatly depending on many factors - phone context, speaker, style of speech and so on. The so-called coarticulation makes phones sound very different from their "canonical" representation. Next, since transitions between words are more informative than stable regions, developers often talk about diphones -

parts of phones between two consecutive phones. Sometimes developers talk about subphonetic units - different substates of a phone. Often three or more regions of a different nature can easily be found.

The number three is easily explained. The first part of the phone depends on its preceding phone, the middle part is stable, and the next part depends on the subsequent phone. That's why there are often three states in a phone selected for speech recognition.

Sometimes phones are considered in context. Such phones in context are called triphones or even quinphones. For example "u with left phone b and right phone d" in the word "bad". It sounds a bit different from the same phone "u" with left phone b and right phone n" in word "ban". Please note that unlike diphones, they are matched with the same range the waveform as just phones. They just differ by name because they describe slightly different sounds.

For the computational purpose it is helpful to detect parts of triphones instead of triphones as a whole, for example, to create a detector for a beginning of triphone and share it across many triphones. The whole variety of sound detectors can be represented by a small amount of distinct short sound detectors. In general we use 4000 distinct short sound detectors to compose detectors for triphones. We call those detectors senones. A senone's dependence on context could be more complex than just left and right context. It can be a rather complex function defined by a decision tree, or in some other way.

Next, phonemes build subword units, like syllables. Sometimes, syllables are defined as "reduction-stable entities". To illustrate, when speech becomes fast, phones often change, but syllables remain the same. Also, syl-

lables are related to intonational contour. There are other ways to build subwords - morphologically-based in morphology-rich languages or phonetically-based. Subwords are often used in open vocabulary speech recognition.

Subwords form words. Words are important in speech recognition because they restrict combinations of phones significantly. If there are 40 phones and an average word has 7 phones, there must be 40^7 words. Luckily, even a very educated person rarely uses more than 20k words in his practice, which makes recognition way more feasible.

Words and other non-linguistic sounds, which we call fillers (breath, um, uh, cough), form utterances. They are separate chunks of audio between pauses. They do not necessarily match sentences, which are more semantic concepts.

On top of this, there are dialog acts such turns, but they go beyond the purpose of the document

Recognition process

The common way how Anryze recognizes speech is the following: we take the waveform, split it on utterances by silences then try to recognize what is being said in each utterance. To do that we want to take all possible combinations of words and try to match them with the audio. We choose the best matching combination. There are few important things in this match.

First of all, it is a concept of features. Since a number of parameters is large, we are trying to optimize it. Numbers are calculated from speech usually by dividing speech on frames, then, for each frame of length typically 10 milliseconds we extract 39 numbers that represent the speech.

That's called feature vector. The way to generate numbers is a subject of active investigation, but in simple case it is a derivative from spectrum.

Second, it is a concept of the model. Model describes some mathematical object that gathers common attributes of the spoken word. In practice, for audio model of senone is gaussian mixture of its three states; to put it simple, it is a most probable feature vector. From concept of the model the following issues are raised: how good does model fits practice, can model be made better of its internal model problems, to what extent the model is adaptive to the changed conditions.

The model of speech is called Hidden Markov Model or HMM, it is a generic model that describes black-box communication channel. In this model, process is described as a sequence of states which change each other with certain probability. This model is intended to describe any sequential process like speech. It has been proven to be really practical for speech decoding.

Third, it is a matching process itself. Since it would take a long time to compare all feature vectors with all models, the search is often optimized by many tricks. At any point, we maintain best matching variants and extend them with time, producing best matching variants for the next frame.

Models

According to the speech structure, three models Anryze used in speech recognition to do the match:
An acoustic model contains acoustic properties for each senone. There are context-independent models that contain properties (most probable feature vectors for each phone) and context-dependent ones (built from senones with context).

A phonetic dictionary contains a mapping from words to phones. This mapping is not very effective. For example, only two to three pronunciation variants are noted in it, but it's practical enough most of the time. The dictionary is not the only variant of mapper from words to phones. It could be done with some complex function learned with a machine learning algorithm.

A language model is used to restrict word search. It defines which word could follow previously recognized words (remember that matching is a sequential process) and helps to significantly restrict the matching process by stripping words that are not probable. Most common language models used are n-gram language models-these contain statistics of word sequences-and finite state language models-these define speech sequences by finite state automation, sometimes with weights. To reach a good accuracy rate, your language model must be very successful in search space restriction. This means it should be very good at predicting the next word. A language model usually restricts the vocabulary considered to the words it contains. That's an issue for name recognition. To deal with this, a language model can contain smaller chunks like subwords or even phones. Please note that search space restriction, in this case, is usually worse and corresponding recognition accuracies are lower than with a word-based language model.

Those three entities are combined together in an engine to recognize speech.

What Anryze optimized

When speech recognition is being developed, the most complex issue is to make search precise (considering as many variants to match as possible)

and to make it fast enough to not run for tremendously long periods of time. There are also issues with making the model match the speech since models aren't perfect.

Mostly the system is tested on the test database that is meant to represent the target task correctly.

The following characteristics are used:
** Word error rate.** Let we have original text and recognition text of length of N words. From them the I words were inserted D words were deleted and S words were substituted Word error rate is

WER = (I + D + S) / N

WER is usually measured in percent.

Accuracy. It is almost the same thing as word error rate, but it doesn't count insertions.

Accuracy = (N - D - S) / N

Accuracy is actually a worse measure for most tasks, since insertions are also important in final results. But for some tasks, accuracy is a reasonable measure of the decoder performance.

Speed. Suppose the audio file was 2 hours and the decoding took 3 hours. Then speed is counted as 1.5xRT.

ROC curves. When we talk about detection tasks, there are false alarms and hits/misses; ROC curves are used. A curve is a graphic that describes the number of false alarms vs number of hits, and tries to find optimal point where the number of false alarms is small and number of hits matches 100%.

Comparison Test with Google and IBM

Methodology of testing the recognition system.

Speech recognition systems used for testing:

- MS Azure
- Google Speech
- IBM Watson
- Anryze STT V.b0.53

For testing the following characteristics were applied:

- 3 male voices
- 2 female voices
- Seperate speech (there is a pause after every word)
- Conjoint speech
- Generals topics
- Special word usage (there are international names, scientific and technological etc.)

Tests description:

1. 1 speaker, conjoint speech, general topic
2. 1 speaker, separate speech, general topic
3. 1 speaker continuous speech, special words
4. 1 speaker, separate speech, special words
5. 2 speakers, conjoint speech, general topic
6. 2 speakers, separate speech, general topic
7. 2 speakers, conjoint speech, general topic
8. 2 speakers, separate speech, special words

Prior to testing, we have prepared 10 different text sets for each characteristic (separate speech, monologue, dialogue, special words etc.)

Testing process

Each speaker have read out particular texts according to test type. Recordings were made in record studio in order to get clean sound. After recording sound files were converted into appropriate formats for each recognition system on tests. With the API results were obtained.

Technologies description

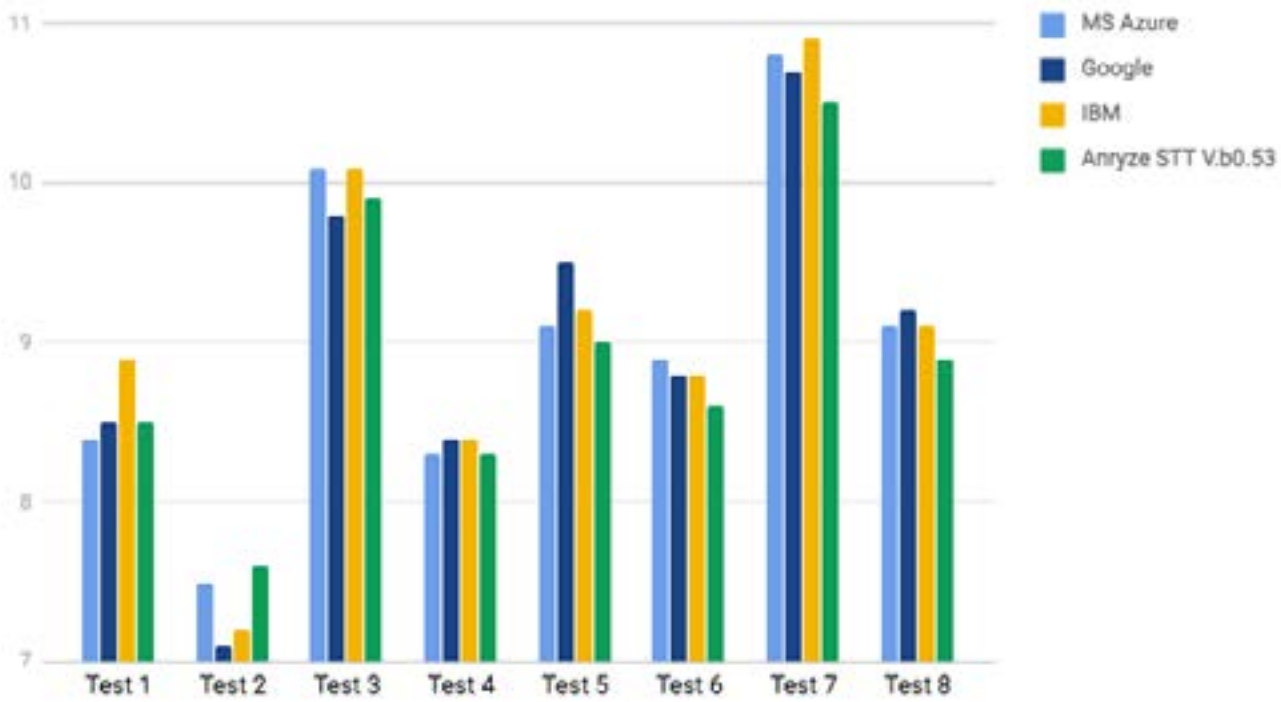
While developing Anryze STT we have applied following technologies and solutions. Here are some of them:

- Wavelet transform - allows to improve recognition by reducing loss of data.
- Recognition using fractal code descriptor - reconstructing the signal while any sampling frequency.
- Multi-Objective Learning for Deep Neural Network Based Speech Enhancement - responsible for constructing extended speech signals.

- Invariant Representations - increases the stability of acoustic variability.
- Highway Connections in Convolutional Recurrent Deep Neural Networks is an extension of the CLDNN model by integrating connections that provides a direct stream of information from the cells of the lower layers to the cells of the upper layers.
- Recognition with the use of distributed capacity – allows to reduce equipment costs and obtain resistance to changing loads.
- Acoustic models based on long short-term memory – is a recurrent neural network architecture that has been designed to address the vanishing and exploding gradient problems of conventional RNNs

Results

Graphic shows the percentage of mistake while recognizing.

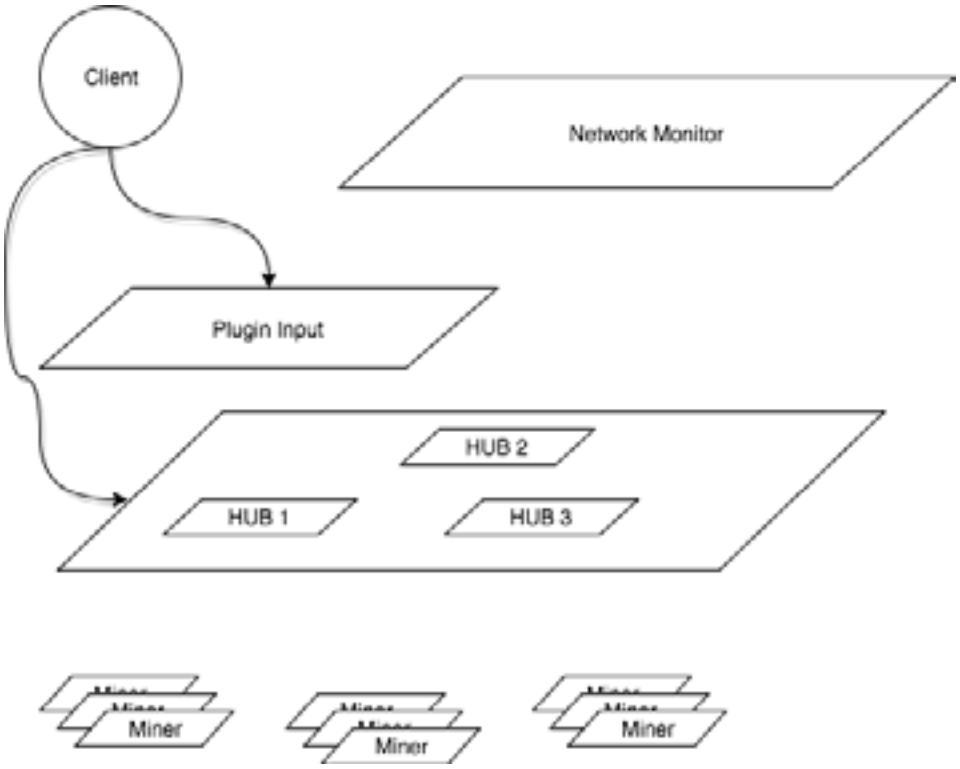


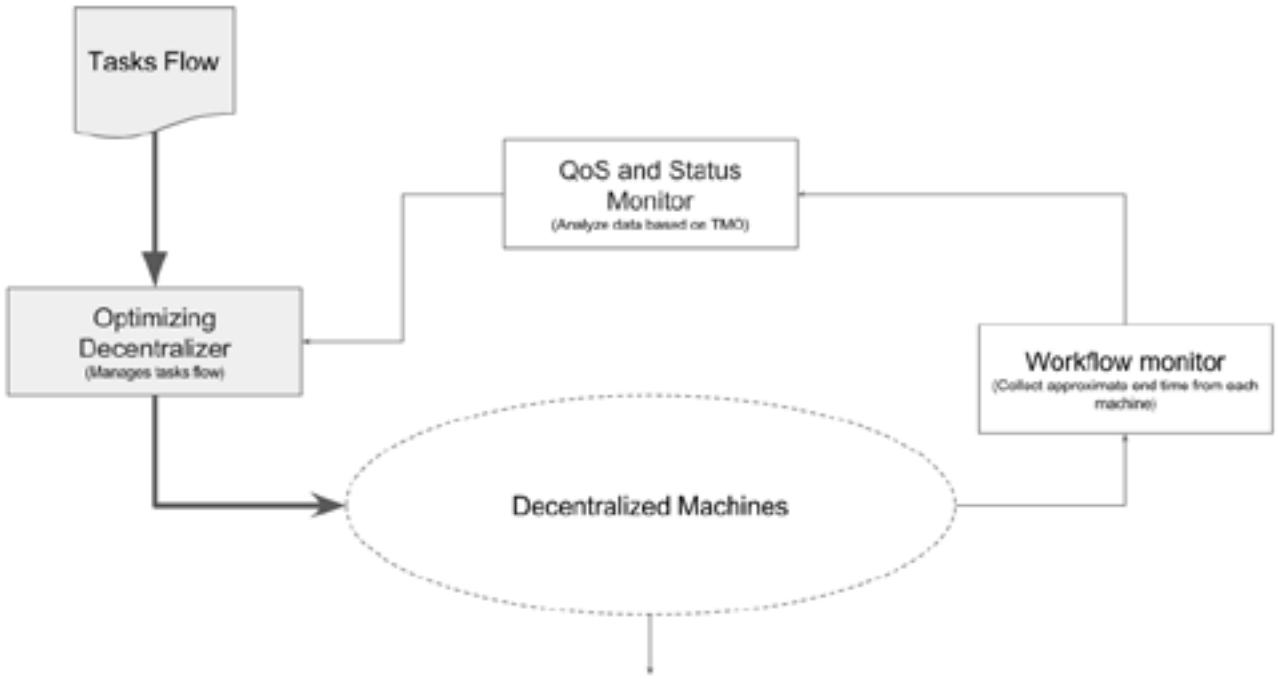
3. Decentralized Network

Anryze is a decentralized worldwide fog computer for AI (speech recognition) purpose. Anryze computing power exchange is the free market, so malicious hubs and users will shortly be ignored by buyers and miners due to their bad reputation. We expect Anryze to be the smartest, cheapest and largest decentralized computing system due to self-learning intelligent agents. Hubs retranslate tasks from users to miners. Hubs unite miners and distribute tasks among them. Priority is given to miners with a better connection to hubs servers and with bigger computing power.

1. User app detects suitable hub (based on access speed and apparent capacities).
2. Form smart contract and pay for service
3. Start data transfer.
4. After several minutes user receive transcribed speech in text format.

In case user upload a prepared audio file, the request is processed in the hub server and then data transfers to the miner. If user uses audio stream, data stores on hubs servers the same time retranslating stream to miner. It ensures the system against possible loss of data in the event of a connection failure.


Client
Network Monitor
Plugin Input
HUB 2
HUB 1
HUB 3
Miner
Miner
Miner


Tasks Flow
QoS and Status Monitor
(Analyze data based on TMO)
Optimizing Decentralizer
(Manages tasks flow)
Workflow monitor
(Collect approximate end time from each machine)
Decentralized Machines

4. Product

Anryze Distributed Network is a next generation speech recognition platform that enables real-time voice recognition alongside with traditional requests.

The Product consists of 2 parts - Miner's part and User's input plugin.

User's input plugin is a web application and API solution where the user decides to upload audio through web browser or upload directly from his server through API.

After files are uploaded, Network Monitor analyzes the audio, calculates number of minutes, and finds the best suitable miner.

5. Market

The problem is that most communications between people happen over the voice and even if somebody records the voice - you need to spend the same amount of time to understand what they were talking about. So a lot of data is just recorded and stored.

Speech recognition brings a huge value for people and businesses because text is easy to analyze.

The global voice recognition market size was valued at USD 51.09 billion in 2015 and is expected to grow at 11% yearly.

The adoption of voice-enabled applications for mobile devices and other smart voice-enabled consumer products is increasing the scale of the consumer/mobile category, challenging the traditional voice technology engine of growth, namely contact centers. Legal compliance issues will continue to ensure the healthcare sector's dependence on voice recognition solutions that can help the industry meet electronic medical record (EMR) standards while providing efficient healthcare.

In the enterprise and healthcare sectors, voice recognition technologies are allowing physicians, soldiers, and other users to accomplish more by reducing the time consumed by written communications. The military sector, the industry's largest investor, uses the technology to increase operational efficiency and for precision driving. North America remains the leading market in that sector, nonetheless, Asia-Pacific countries are rapidly increasing their market shares.

First Clients

First Clients will be brought to the platform by Anryze. Most of them are US companies in such industries as financial compliance, banking, and sales. Here is the list of the companies we are already working with.

Mast Mobile – provides the mobile sales productivity platform that embeds CRM into your sales team's day-to-day activities, automatically loading all their client communications into the database. Reps can focus on selling, not data entry, and the company gets value from its investment in CRM. It solves the biggest challenge for CRM systems (manual data entry) which is increasing as team move to mobile devices.

https://www.mastmobile.com/

Weeden & Co – broker firm founded in 1922 in New York. It provides brokers service for institutional and private investors on US markets (as NYSE, Nasdaq, etc.) with more than 300 broker agents who talk to clients daily.

http://www.weedenco.com/

Bino – mystery shopping platform that allows restaurants, cafes, bars, shops and stores to measure customer experience with help of mystery shopper visits. Most of the visits happen with voice recording. Bino works on Eastern European and US markets.

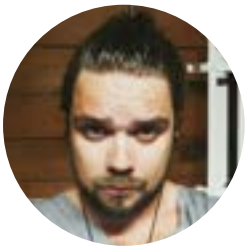
http://getbino.com/

CohereComm – vendor of telecom systems and software for financial institutions as broker firms, financial advisers, banks, and compliance outsourcing companies. CoheComm works with 385 broker firms that generate more than 500k minutes of voice recordings daily.

https://www.coherecomm.com/

---

6. Founders



Anton Gera
CEO and Founder
https://www.linkedin.com/in/anton-gera-6b932a1a/

Mike Ezhov
CMO and Founder
https://www.linkedin.com/in/mike-ezhov/

Oleg Zaichuk
CTO and Founder
https://www.facebook.com/oleg.zaichuk

Max Kudymets
Project Manager
https://www.linkedin.com/in/maksymkudymets/

Oleg Fedosenko
Machine Learning and AI development
https://www.linkedin.com/in/oleg-fedosenko-b185038/

Alex Pogrebenniy
Back-end development
https://www.facebook.com/pogrebenniy.alexandr

Oleksii Leontiev
PhD in applied mathematics, University of Tokyo
https://www.linkedin.com/in/oleksii-leontiev-8a010952/

Sergey Levashkin
PhD in applied mathematics, Moscow State University
https://www.linkedin.com/in/sergey-levashkin

Partners

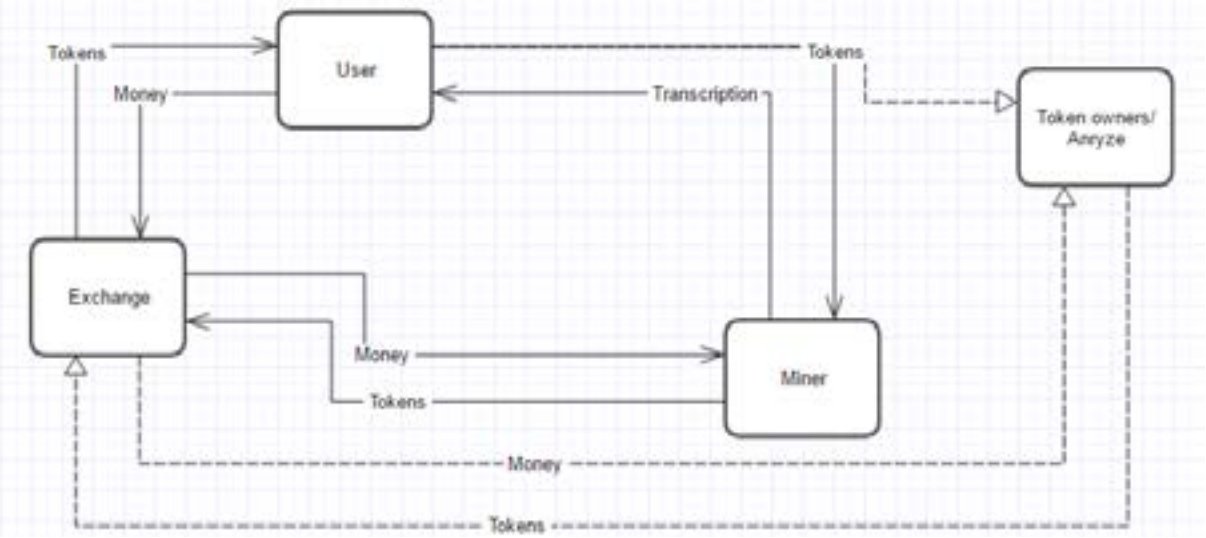
We are proud to be supported by our advisers, partners, and investors.

Starta Accelerator (NYC based accelerator) with focus on
Eastern European tech companies.
http://www.startaaccelerator.com/

Waves Platform
Waves is an open blockchain platform designed for ease of use
and mass adoption.
https://wavesplatform.com/

7. Monetization

1. Anryze Distributed computing power based on blockchain charges users for transcribing a certain time of the audio file. Therefore, users pay for service, miners (computing capacities) get tokens, and token holders with Anryze get commission.



Miners get 70% of tokens in exchange for computing power, other part goes equally to Anryze and token holders (10% and 20%). If the demand on service will be bigger than capacity, the price of tokens on the exchange will rise.

Anryze gets tokens from transcription to develop our network, introduce voice verification and other AI services. While token owners get them as compensation for their investments.

2. Each minute of the processed audio is equated to a token. We use the market price model, so the current cost of a minute will depend on transcription demand and mining supply in a system and also on token demand and supply on the crypto exchange. Initially, the rate will be

1:1 (one minute for one token) and 1:0,01 (one token for $0,01). As the quantity of tokens is limited and in the nearest future it would not be able to meet transcription demand, so the price of the token will rise the same as the number of the minutes per token. Such model will provide greater demand and reward investors. Tokens can be bought on ICO or after ICO through crypto exchange.

Basic market price model:

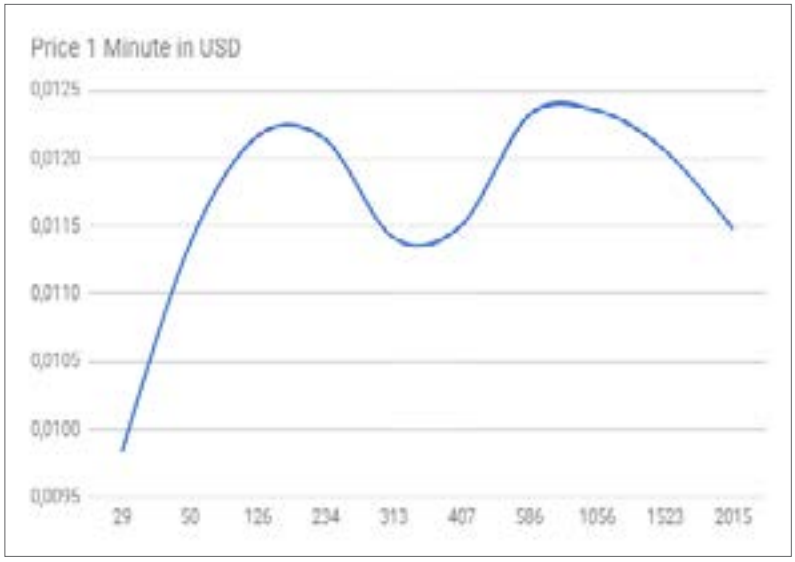
$$D_n = \alpha * D_{n-1} * \exp\left(\frac{B_n - P_{n-1} * \min(D_{n-1}; S_{n-1})}{B_n + P_{n-1} * \min(D_{n-1}; S_{n-1})}\right)$$

$$S_n = \beta * \left(\frac{P_{n-1}}{P_{n-2}}\right) * \exp\left(\frac{D_{n-1} - S_{n-1}}{D_{n-1} + S_{n-1}}\right)$$

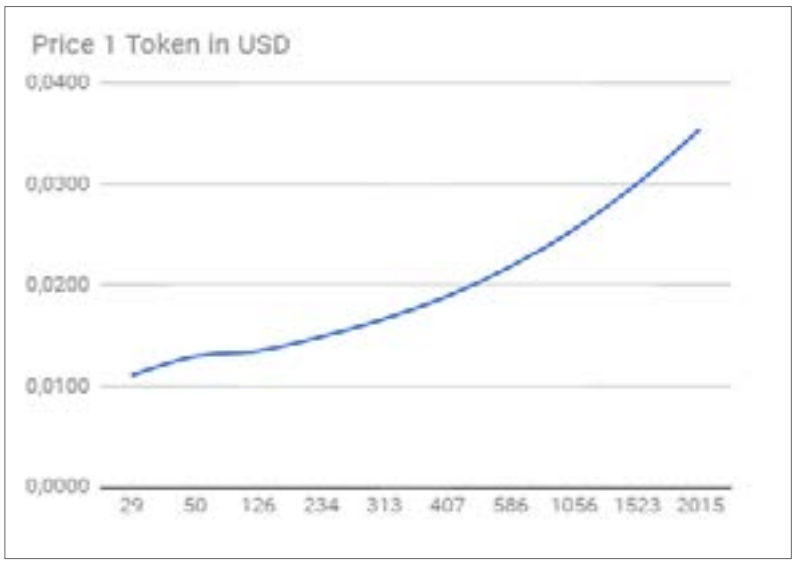
$$P_n = \gamma * P_{n-1} * \exp\left(\frac{D_n - S_n}{D_n + S_n}\right)$$

Where:
D_n-demand in n-period; S_n - supply in n-period; P_n-price in n-period
α,β,γ - Coefficient of adaptation / base increase; B_n-budget in the period

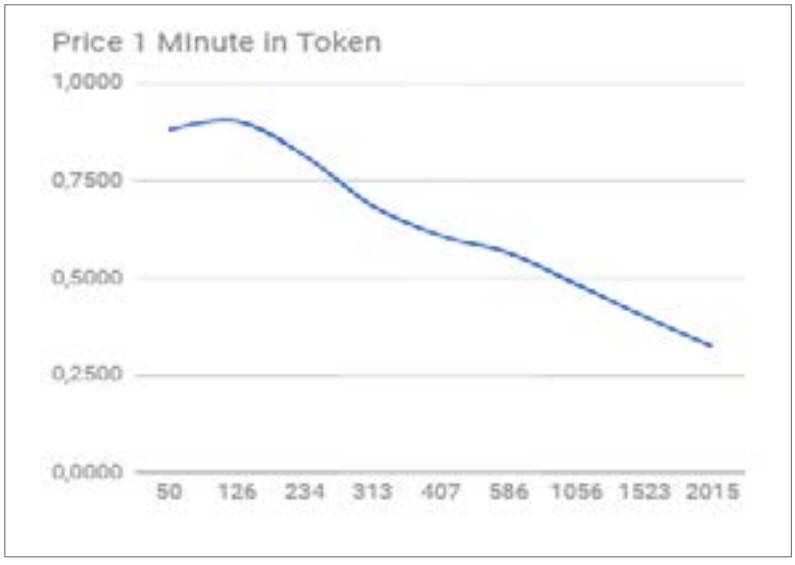
The charts below show the forecast of the cost of tokens and minutes (The x-axis indicates the volumes of transcribed audio in millions of minutes)



The real price of transcription (minute per USD) will be volatile but thanks to the market mechanism, a high price increases supply and reduces demand, which leads to lower prices, and vice versa



While amount of tokens is limited, transcription volumes will rise which due to market will increase demand for tokens and their price as well.



High prices for tokens will not affect transcription costs, therefore user can transcribe more minutes per 1 token. Price for one minute in token will decrease.

Money-Minute market		2Q	3Q	4Q	5Q	6Q	7Q	8Q
Demand	minutes	29 008 000	67 631 976	144 513 069	234 317 477	313 404 535	413 827 435	672 567 483
Supply	minutes	30 000 000	50 488 004	126 199 320	235 367 674	354 567 633	407 377 224	586 180 710
Transcibed	minutes	29 008 000	50 488 004	126 199 320	234 317 477	313 404 535	407 377 224	586 180 710
Price for minute	$	0,0098	0,0114	0,0122	0,0121	0,0114	0,0115	0,0123
Money-Token market		338 300 000	338 300 000	338 300 000	338 300 000	353 300 000	353 300 000	353 300 000
Miners tokens		18 071 543	31 100 315	79 799 993	133 829 277	150 489 424	173 726 790	231 931 228
Demand	$	285 244	574 012	1 535 217	2 844 110	3 576 712	4 685 831	7 221 487
Supply	$	233 501	418 723	1 412 952	2 329 177	2 856 998	3 633 382	5 408 341
Tokens turnover		25 816 489	44 429 021	113 999 990	191 184 682	214 984 892	248 181 129	331 330 326
Price for token	$	0,0110	0,0129	0,0135	0,0149	0,0166	0,0189	0,0218
Token-Minute market								
Demand		25 816 489	44 429 021	113 999 990	191 184 682	214 984 892	248 181 129	331 330 326
Supply		30 000 000	50 488 004	126 199 320	235 367 674	354 567 633	407 377 224	586 180 710
		29	50	126	234	313	407	586
Price for minute		0,8900	0,8800	0,9033	0,8159	0,6860	0,6092	0,5652

8. ICO

Anryze is planning to raise between $1 – $3 million USD in the next round of financing by selling tokens. This money will help to accelerate market penetration and product development.

Anryze is an official service provider and crowd-funding initiator.

A special token (RYZ) will be created on the Etherium Platform for crowdfunding.

Soft cap: 180 million RYZ

Hard cap: 330 million RYZ

The initial rate will be 1 USD = 100 RYZ.

The expected amount of tokens to be offered is 330 million RYZ* (an additional 6,6 million RYZ will be issued for the bounty program, and 15 million RYZ will be issued to the option pool for rewarding current and future employees. These tokens cannot be sold for the next 12-month period and will be under lock contract). Total amount of tokens will be 353 300 000 RYZ tokens.

The tokens will be sold at a discount to early buyers at the rate sliding from 150 RYZ tokens per $1 to 100 RYZ tokens per $1. Exchange rates are shown in the table below:

331,7 million tokens to sell out	
Cumulative number of USD committed	Price per 1 USD
1 - 120 000	150
120 000 - 250 000	140
250 000 - 500 000	130
500 000 - 800 000	120
800 000 - 1 500 000	110
1 500 000 - 3 000 000	100

Tokens will be sold through several currencies in the equivalent of Ethereum

The creation of new tokens will stop after an equivalent of $3 million USD is raised or after the ICO expiration date.

All unsold tokens will be distributed among all investors in accordance with their deposits.

No new tokens will be created after ICO.

All money raised will be exchanged for Anryze distributed network audio transcription or sold through crypto exchange.

In the case of fund raising less than the planned smart contract will refund all money back to investors

.

9. Escrow

All money raised will be deposited to an escrow account. Management will be able to access this money only if a minimum of $1 million USD or more is raised.

In the case of fund raising less than the planned smart contract will refund all money back to investors. In any case, our team will continue to develop the product and introduce it to the market.

10. Compensation to token owners and miners gains

RYZ will be traded on the crypto exchange.

Blockchain will be transferring 20% tokens from each transcription of the audio to the RYZ token owners in the proportion to the amount of token held. It will start doing this at the same time blockchain of de-centralized transcription capacities starts working. Amount of com-pensation and miners revenue are shown in a spreadsheet below:

		2Q	3Q	4Q	5Q	6Q	7Q	8Q	9Q	10Q
Token owners gains										
for 1000$	$	19	39	100	206	295	435	722	1 524	2 588
Investment balance	$	1 105	1 292	1 346	1 487	1 663	1 887	2 179	2 552	3 004
Total balance	$	1 124	1 330	1 447	1 693	1 958	2 322	2 901	4 076	5 592

		2Q	3Q	4Q	5Q	6Q	7Q	8Q	9Q	10Q
Miners gains										
Minutes transcribed		29 008 000	50 488 004	126 199 320	234 317 477	313 404 535	407 377 224	586 180 710	1 056 137 677	1 523 020 806
Revenue in tokens		18 072 677	31 102 802	79 814 721	133 862 398	150 534 368	173 785 639	232 016 011	357 859 568	427 809 714
Revenue in USD	$	199 671	401 808	1 074 652	1 990 877	2 503 698	3 280 082	5 055 041	9 131 471	12 852 743
Cost of minig	$	14 504	25 244	63 100	117 159	156 702	203 689	293 090	528 069	761 510
Income	$	185 167	376 564	1 011 552	1 873 718	2 346 996	3 076 393	4 761 951	8 603 402	12 091 233

11. Current project's status

Presently, the application is working in a beta and is integrated to several clients. Integration with new B2B users will be completed in a few months. Also, technical and legal work for entering the US market is in progress. All money raised from the ICO will be used for a fully functional application development and the expansion into new markets.

Anryze website: https://anryze.com

The project team is responsible for making the results open to the public and for using all available resources to inform about the project.

- We will publish a report about current development results and issues at least twice a month, as well as a monthly financial report about capital expenditures.
- The Report will contain current project needs and issues.
- All major breakthroughs will be communicated with interested mass media and spread at the major community forums like BitcoinTalk and CryptoCoin Talk.

12. Preliminary financial plan

Preliminary plan is based on ICO = $2M assumption

	1Q	2Q	3Q	4Q	5Q	6Q	7Q	8Q
Revenue, total, USD	0	83 275	111 825	218 684	393 151	571 636	732 763	934 008
COGS	0	11 408	17 611	31 509	52 781	74 173	95 045	117 120
Traffic & Maintenance	0	8 910	14 256	24 948	40 986	57 024	73 062	89 100
Processing	0	2 498	3 355	6 561	11 795	17 149	21 983	28 020
OPEX	94 020	122 750	133 930	150 960	155 580	162 480	164 610	167 200
Marketing & PR	13 560	24 890	33 670	49 500	54 120	61 020	63 150	65 740
Legal & advisers	3 600	4 800	4 800	6 000	6 000	6 000	6 000	6 000
Administrative expenses	11 460	11 460	11 460	11 460	11 460	11 460	11 460	11 460
Salaries OPEX	65 400	81 600	84 000	84 000	84 000	84 000	84 000	84 000
EBITDA	-94 020	-50 883	-39 716	36 216	184 790	334 983	473 108	649 688
CAPEX	173 500	155 500	184 500	166 500	148 500	130 500	130 500	130 500
Salaries CAPEX	83 500	83 500	112 500	112 500	112 500	112 500	112 500	112 500
Hard&Soft purchase	90 000	72 000	72 000	54 000	36 000	18 000	18 000	18 000
EBIT	-267 520	-206 383	-224 216	-130 284	36 290	204 483	342 608	519 188
Taxes	0	0	0	0	0	0	0	93 216
Expenditures Total	267 520	289 658	336 041	348 969	356 861	367 153	390 155	414 820
NET PROFIT	-267 520	-206 383	-224 216	-130 284	36 290	204 483	342 608	342 664
Salary & Benefits, total	137 400	138 600	156 000	156 000	156 000	156 000	156 000	156 000
Personnel Headcount	21	23	31	31	31	31	31	31
C-level Salaries, a month	3	3	3	3	3	3	3	3
CEO (NYC)	9000	9000	9000	9000	9000	9000	9000	9000
CMO	9000	9000	9000	9000	9000	9000	9000	9000
CTO	6000	6000	6000	6000	6000	6000	6000	6000

	1Q	2Q	3Q	4Q	5Q	6Q	7Q	8Q
Sales & BD	2	3	3	3	3	3	3	3
VP of Sales	24000	24000	24000	24000	24000	24000	24000	24000
Sales Representative	15000	15000	15000	15000	15000	15000	15000	15000
Sales Representative	0	15000	15000	15000	15000	15000	15000	15000
Product development	14	14	20	20	20	20	20	20
Senior Fulstack Developer	6000	6000	6000	6000	6000	6000	6000	6000
Senior Fulstack Developer	0	0	6000	6000	6000	6000	6000	6000
Backend Developer	4500	4500	4500	4500	4500	4500	4500	4500
Backend Developer	4500	4500	4500	4500	4500	4500	4500	4500
Backend Developer	4500	4500	4500	4500	4500	4500	4500	4500
Backend Developer	4500	4500	4500	4500	4500	4500	4500	4500
Backend Developer	4500	4500	4500	4500	4500	4500	4500	4500
Frontend Developer	3000	3000	3000	3000	3000	3000	3000	3000
Voice Analyst	9000	9000	9000	9000	9000	9000	9000	9000
Voice Analyst	0	0	9000	9000	9000	9000	9000	9000
Data Analyst	0	0	9000	9000	9000	9000	9000	9000
Data Analyst	9000	9000	9000	9000	9000	9000	9000	9000
Product manager	3000	3000	3000	3000	3000	3000	3000	3000
Designer	3000	3000	3000	3000	3000	3000	3000	3000
Fulstack Developer	6000	6000	6000	6000	6000	6000	6000	6000
Fulstack Developer	6000	6000	6000	6000	6000	6000	6000	6000
Fulstack Developer	0	0	6000	6000	6000	6000	6000	6000
Backend Developer	0	0	4500	4500	4500	4500	4500	4500
Backend Developer	0	0	4500	4500	4500	4500	4500	4500
Frontend Developer	4500	4500	4500	4500	4500	4500	4500	4500
Other staff	2	3	5	5	5	5	5	5
Support	2400	3600	6000	6000	6000	6000	6000	6000