

Speech Recognition In Phone Calls Using Mobile Phone Cloning

U.C.Azhagu Raaj¹, S.Gowtham Viswanath²

¹Information Technology, Anna University / Nandha College of Technology, Erode.

²Information Technology, Anna University / Nandha College of Technology, Erode.

ABSTRACT:

In today's modern trends we are facing major threats in terrorist activities and in illegal activities. Also the drug dealings have become more frequent and privatized. It has become difficult for us to bring those activities into our bounds. These activities are found to take part using the mobile phone calls. Hence it is necessary to set some definite constraints for those mobile calls. These mobile calls can be tracked using the ethical mobile phone cloning. The cloned phones can be tracked by using speech recognition. By using the ESN and MIN numbers from the mobile phones it can be cloned and by using the Hidden Markov Model (HMM) algorithm, the speech from the mobile can be tracked. We use suspect words from the user to track their mobile phones. When ever a suspect word is encountered, the mobile phone will get tracked and found for any illegal activities. Here not only the call gets recorded but also the location of the callers will be noted. Thus implementing this technique of speech recognition and mobile phone cloning we can prevent large disasters and find out the culprits easily.

KEYWORDS:

ESN, MIN, HMM, phoneme, utterance, viterbi algorithm.

1. INTRODUCTION:

Speech recognition is one of the most abruptly developing domains in recent trends. It also has placed an undisputable role in Digital Signal Processing. This technology takes impart in all the audio functions in the future. It considered being one the most important core technology in the developed countries such as U.S.A. Also mobile phone cloning has placed an important revolution in the mid 1990's. It is also playing a major changeover in the recent trends.

It is considered to be a threatening threat for the mobile phone users.

Mobile phones can be cloned by using the Digital Data Interpreters and large routers. The ESN and MIN numbers of mobile phones in the particular zone can be received using the DDI. This will get the track of all the phones in that zone. Thus every phone in that zone will be tracked automatically. Now if the phone uses suspect words, then that particular pair of phones will get tracked and their call will be recorded by using the Hidden Markov Model (HMM) algorithm in speech recognition. The secondary filters in the call tracking center will filter the received speech signal for the second time.

2. MOBILE PHONE CLONING:

2.1 components required:

- Digital Data Interpreter (DDI)
- Large router

The phones can be cloned by installing a large router and a Digital Data Interpreter with the router. Reverse engineering can be applied in the router in order to receive the ESN and MIN numbers from all the mobile phones in that zone. Since a large router is installed the router can receive all the signals from that zone. This router sets a particular range in an area and this bounded region will be in total control of the router and the Digital Data Interpreter. The recorded MIN and ESN numbers from the mobile phones will be stored in the external storage in the mobile tracking center.

Both GSM and CDMA phones can be tracked from this method. The phones of all the networks can be tracked by using the router and a DDI. The cloned model can be developed from the original phone by the copy cat box. This

www.ijreat.org

copy will be used to create a replica of the phone and can be used to track the callers.

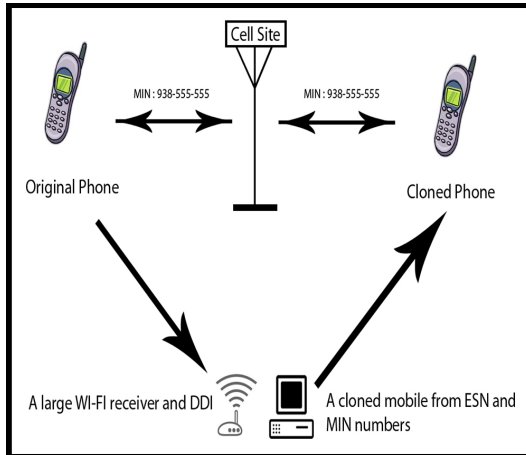


Fig 1. Mobile cloning mechanism

3. RECOGNIZING SPEECH IN CLONED PHONE:

Components required:

- Sound card
- Phoneme
- Acoustic model
- Language dictionary
- Filter

In this model we use speaker independent speech recognition technique. It will not be based on the particular speakers since the phone call can be made by any one else. Since we will not be dealing with single word processing, it uses continuous speech recognition and spontaneous speech recognition for recognizing the speech in a high speed rate. In spontaneous speech recognition, even little utters in the speech can be recognized, Like “ums” and “ahs”. This will also be used in analyzing the natural and non rehearsed sounds.

The received speech signal will be synthesized by the sound cards in the system. The received signal will be in analog format and it will get converted into digital format by the sound cards. Since the computers can access only 0's & 1's the signal is converted into digital format. The signal then will use the phonemes present and can used to bring the exact use of the letter in a word.

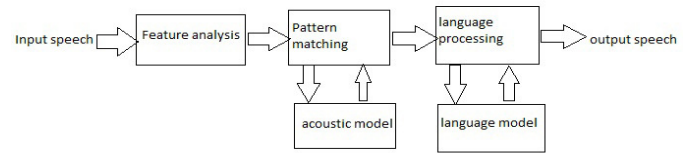


Fig 2. Speech processing

3. PHONEME:

A phoneme is the digital format of the speech. It is also said to be the way of pronunciation of a letter in a word. In English there are about 42 phonemes and Tamil has around 100 phonemes in pronunciation. The phonemes are being governed by the International Phonetic Alphabet (IPA) over 100 years. It manages all phonemes in all the languages all over the world.

4. ACOUSTIC MODEL:

An acoustic model is the digital format of the speech signal. It will be used in processing the speech and giving out the synthesized speech signal. The acoustic model will have the phoneme for analyzing the speech signal. We have many acoustic models to use.

4.1 Artificial neural network:

The artificial neural network will handle specific tasks such as pattern recognition or data classification. Hence ANN can be used in the beginner levels that support all types of Phonemes in any language.

4.2 Recurrent neural network:

A recurrent neural network is used where connections between units form a directed cycle. RNN uses the internal memory to process arbitrary sequence of inputs.

4.3 Gaussian mixture model:

Gaussian mixture models are used for small domains. GMM is used in long time speech processing. That uses 4.2M or more than 3000 hours of speech data. The GMM size is of 1/3 of the floating point. GMM uses more amount of memory in processing data.

4.4 Deep neural network:

www.ijreat.org

The deep neural networks is feed forward, artificial neural network that has more than one layer of hidden unit (j), typically uses logistic function. To reduce the memory consumed by the language model rescoring is performed with a compressed n-gram. They are also used for real time operation on mobile devices.

4.5 Dynamic time warping:

Dynamic time warping is used extensively for measuring similarity between two temporal sequences which may vary in time or speed. This can be used in different speech speeds. DTW has been implemented in audio, video, graphics and indeed any data that can be turned into a linear representation can be analyzed with DTW. This DTW is also used in speaker recognition and in online signature recognition.

5. ALGORITHM USED:

5.1 Hidden Markov Model (HMM):

All the modern systems use Hidden Markov Model as the acoustic model. It is used to view piecewise stationary signal. i.e., this model is used to process the signal in short time. It uses just 10 millisecond in processing a signal. HMM is trained automatically, simple and feasible to use. For every 10 millisecond there will be a real valued vector. This HMM suits for different types of speaker and different recording conditions. The HMM uses a new technique known as the vocal tract length normalization (VTLN).

It also uses Heteroscedastic Linear Discriminant Analysis (HLDA) which is commonly used in the pattern recognition. This pattern recognition technique will get the exact speech sequence from the input and will have varying speech in it. The VTLN and HLDA thus will give enhancement for the speech to get synthesized.

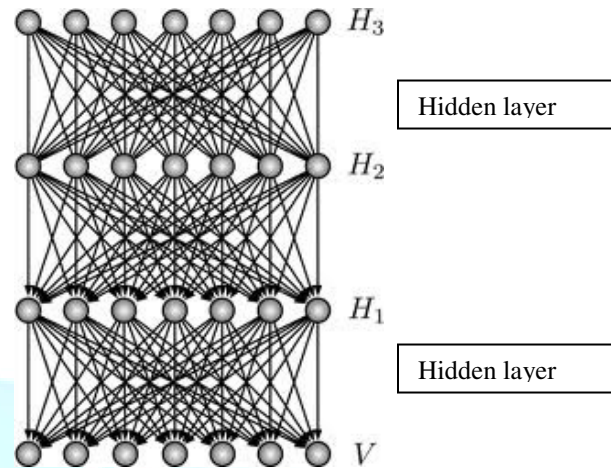


Fig 3. Hidden Markov Model

5.2 DECODING ALGORITHM:

For decoding the signal in the hidden markov model the viterbi algorithm is used. This viterbi algorithm leads to find the best path to solve the problem. In viterbi algorithm a metric is assigned to each surviving path. Then the code is decoded as the surviving path with the smallest metric.

Hence the viterbi algorithm carries out rescoring in the decoding signals that is used in avoiding noise in the signal. The Viterbi algorithm thus uses the best path from the original data for rescoring the speech.

6. LANGUAGE DICTIONARY:

The language dictionary will have all the possible words for the received speech. This language dictionary thus will have the word storage area where all the words will get stored. When ever the speech is received in the language dictionary it will compare for a best suited word and will provide the word in the output. The language dictionary will act as the storage space in the language processing sector. Since it stores large amount of data it needs a very high storage space.

7. PHONE CALL TRACKING CENTER:

The phone call tracking center comprises of a large router and a DDI connected with it. This large router will use reverse engineering technique and will attract all the networks within its boundary. By taking in all the bounds the digital data interpreter now will take the ESN and MIN numbers from the mobile. Thus the mobile gets cloned from the original module. Now the tracks on the mobile will be followed and if the mobile gets any call it

www.ijreat.org

will get tracked for suspect word that is present in the language dictionary. If the word in the language dictionary matches the speech then the system will throw a warning signal and the call will be traced and recorded.

The phone tracking center may not be sufficient enough to track a whole country so that it can be constructed on basis of separate zones in a state. For example, Tamil Nadu has 32 districts and each district can be installed with one center. All the mobiles in that district will be monitored by that particular center. All the district centers can also be monitored by a main center.

Each center can be assisted with a technician and some number of employees, so that it will be easy to work on the suspected network numbers. During suspecting the call the signal has to pass through a special filter that resembles the exact language dictionary. This filter will ensure that there no threat to privacy for users. When there is a secondary filter used in the system it helps the center to work efficiently and the results can be more accurate. The filter will have the same words that are used in the language dictionary. Since we implement continuous speech recognition, it causes imperfection in the filtering process so the secondary filter will bring high accuracy in the signal transfer.

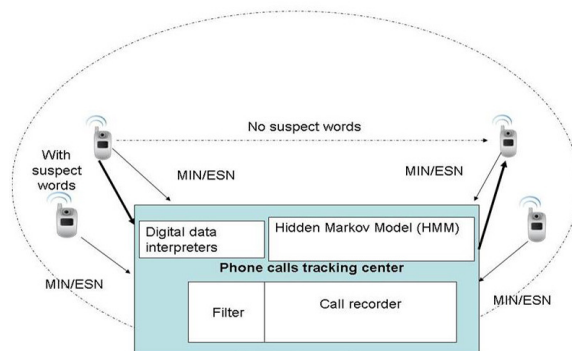


Fig 4. Phone call tracking center

8. FILTERS:

The filters are used to filter the suspect words from the actual synthesized speech. They used to bring in privacy for the ordinary mobile users. Since we in the stage that we cannot be looking over all the speech that are received, hence we use the filters to filter out the suspect word users from the ordinary mobile users. In this proposal we have installed 10 words as

suspect words in order to filter out the results. The words that we used are, 1. RDX, 2. Human bomb, 3. Kidnap, 4. Poison, 5. Naxalite, 6. Sleeper cells, 7. Heroin, 8. Brown sugar, 9. Cannabis, 10. Cocaine. Such words are not ordinarily used and hence these words are tracked to track the illegal activists. Using this filter alone will not get the exact privacy for the users, because the privacy will not be restored in that extent.

9. FUTURE IN PHONE CALLS TRACKING:

For the developing countries like India terrorism is a major threat in economic, political and in all sorts of ways. By using this phone calls tracking system we can prevent the nation's major threats.

Phone calls tracking can also be used to bring in control for the local illegal activities such as drug dealings which will prevent people from dealing with high power drugs and prevents students to get involved in consuming drugs.

When these centers are used effectively crimes can be prevented very soon before it happens.

This can be implemented in all parts of the country so that these kinds of improvised thoughts are more important in a developing country. Still more improvised techniques in these centers can bring a wide detection in terrorist activities and will bring in cost efficient prevention methods.

Ethical mobile phone cloning takes place in these centers and hence we can also track when any mobile phones are cloned. So that black hat hackers can be spotted using phone calls tracking.

10. CONCLUSION:

Thus contributing for the nation's wellness this phone call tracking system should bring peace and development in all sectors of the country. We are now in the middle of a revolution sparked by the terrorist activities all over the country and all over the world. We are in the state to bring immediate preventive measures for the country. Speech processing being the rapidly uprooting technology in recent trends is considered to play a major role in nation's security.

www.ijreat.org

11. REFERENCES:

- [1] Tripti Singh, Saumya Tiwari, Nidhi Gupta, Sarika Tyagi, “Mobile phone cloning”, International journal of advances in computing and communications, 2013.
- [2] Parwinder pal Singh, Er. Bhupinder Singh, “Speech recognition as emerging revolutionary technology”, international journal of advanced research in computer science and software engineering, October 2012.
- [3] Xin lei, Andrew senior, Alexander Gruenstein, Jeffrey Sorensen, “Accurate and compact large vocabulary speech recognition on mobile devices”- Interspeech 2013.
- [4] S.E. Levinson, A. Ljolje, L.G. Miller, “continuous speech recognition from a phonetic transcription”, AT & T Bell laboratories.
- [5] Dr. Joseph Picone, “fundamentals of speech recognition”, Mississippi state university.
- [6] www.google.com
- [7] www.wikipedia.com
- [8] www.ieeexplore.ieee.org.

The logo for IJREAT PRDG features a stylized globe in the background. The word "IJREAT" is written in a large, light blue, sans-serif font across the middle of the globe. Below the globe, the word "PRDG" is written in a larger, bold, light blue, sans-serif font. The entire logo is set against a white background with horizontal lines radiating from the globe.

www.ijreat.org