Proposal on a new framework for big data analysis in cellular telecommunication network in Bangladesh

Md. Ashek-Al-Aziz, Amathul Hadi Shakara, Md. Kawshik Ahmed

Department of CSE and ETE
University of Development Alternative, Dhaka, Bangladesh

Abstract:
World has been being suffering from terrorist activities since the year of 1999. A plenty of massacres happed during last decade worldwide. Also, Bangladesh has suffered from this. It is an obvious that players of such dangerous activity need to have communication and conversation among them before doing such massacres. If their conversations are understood prior to anything happens disastrous, massacre can be prevented. So that a new framework have been proposed. A review has been carried out on telecommunication and speech recognition system before new framework is proposed. Future scope of research work is also discussed.
Key Words: *Big Data; Telecommunication; Speech Recognition.*

I. Introduction

Huge volume of data are being transmitted in telecommunication system specially cell phone networks in every moment. Every Government all over the world is very much concern about authenticated usage of cell phone communication network and cell phone users' identification. Keeping the large volume and variety of data is an essential task of cell phone operators specially the registration information and applying queries on voice records over the network for tracking the terrorists or any kind of intruders. Searching over the voice data has been felt a required job for tracking the harmful activities before any massacre happened since their might be some conversation over telephone on each occurrence. Storing the voice conversations and searching over that is a big data analysis issue. An idea of dedicated framework of such big data analysis has been proposed in this article.

Data are generated from various sources and kept in conventional storage mediums and usually conventional databases and queries are executed on the data for seeking the desired answers and result. Data becomes 'Big Data' when it reaches to such a volume and complexity that overflows the capacity of database, go beyond the limit to be processed by conventional query methods over data structures [1-2]. The solution of this issue is to incorporate the improved distributed computing and machine learning techniques, pattern recognition etc in Internet infrastructure since most data are generated in distributed computing environment over Internet [2-3]. Big data analysis is practiced in different application area like environmental data analysis, nuclear research, and multimedia data analysis and so on. Our aim is to propose a hybrid framework model for storing and analysis of speech data over cellular mobile telephone network for the purpose of crime prevention especially in Bangladeshi context and query over user registration information that might be a big data issue in Bangladesh. Bangladesh has 165.615 million subscribers of cell phone according January 2020 census and Bangladesh has threat of terrorists.

There are three issues concerned in our case of study – (i) Communication Framework through which conversations' voice data are transmitted, sent and received since voice recorders for capturing and storing are to be located in the network, (ii) Application of speech recognition methods over stored conversations and (iii) Legal issue due to breaking the confidentiality of the conversant.

II. Communication Framework

Open System Interconnection (OSI) reference model is the primary concern for every kind of communication system. According to seven layer architecture, data thrown by source end to the network passes through Router towards destination end. Router acts like a switching station directs the data units using Djkstra algorithm toward destination through another cloud [4-5]. In case of computer networks or Internet, Router is the implemented device otherwise it is a switching station for telecommunication network whatever it is in fixed phone or cellular mobile phone network. Geographic areas are divided into many hexagonal cells (unequal in size and non uninformed shape mostly) where each cell consists of one tower at least with many antennas covering 360^0 angles with Base Transmission Station (BTS) which is actually a switching station. A fiber optic

network connecting all the base stations forms a large mesh connecting through Mobile Telephone Exchanges (MTX) strengthens the communication efficiency and a central switching station or control room is located at a central location for the whole network administration [6-7-8]. Switching, routing and congestion control is automatically carried out using software driven automated system at each BTS and MTX. Few cell phone operators in Bangladesh like Grameen Phone Limited, Robi Axiata, Teletalk Bangladesh and Banglalink implements same technology in same fashion throughout Bangladeshi territory.

III. Speech Recognition System

Usually there are two types of speech data – voiced and unvoiced among which we are concerned about voiced data.

Basic model of speech recognition gathers acoustic speech data, analyze and make decision what has been told. Speech signal as soon as it is captured can be proceed for Training and Recognition. The sequence of Training and Recognition steps are – (i) Speech Detect with Framing and Overlapping (ii) Pre-emphasis (iii) Hamming Window (iv) Feature Extraction (v) Fast Fourier Transformation (vi) Matching and (vii) Taking Decision. When speech is detected as signal wave, frames are stripped each of 20 mille second in length where frames can be overlapped up to certain limit. Existence of utterance is checked measuring the strength of the signal and zero crossing rates in frames that crosses the threshold. If threshold value is crossed then frames are buffered until contiguous buffers exceed the predefined buffer length. A pre-emphasis filter acts as high pass filter removes the DC component of the wave and makes the sound wave sharper. Hamming window integrates the closest frequency components and Fast Fourier Transformation (FFT) transforms the time domain of the signal into frequency domain. Then features of the signal like amplitude, energy, pitch, intensity, velocity, acceleration, vibration, etc are extracted using Mel Frequency Cepstral Coefficient (MFCC), Gaussian Mixture Model (GMM) and Linear Predictive Coding (LPC) model applied in combination. Such features are actually extracted by matching the feature matrix and reference matrix. Decision is made which word it is uttered depends on the closeness of the feature and reference matrix of dictionary word [9]. A single word is recognized using the method discussed. A dictionary with all words' utterance should be made before a sample is fed into the system to be checked which word has been uttered or a dictionary with utterance can be built dynamically. Obviously, the words should be from Bengali and other relevant language like Hindi, Urdu, Pashtu, Arabic and English of course. This is actually a pattern recognition based approach since the spectrum extracted features are compared with the words' features already stored in dictionary database.

A widely known neural network model so called Hidden Markov Model (HMM) is also applied in this approach. HMM can suitably applied for recognition of a word, phrase or a sound less that a word. A direct comparison can be made between the input speech which is to be matched or recognized and the possible patterns learned in the training stage to identify the input unknown speech on the basis of how good the pattern is matched with the reference pattern [10-11].

Sometimes a mobile phone handset registered by one person can be handed over to another person and misused by unethical communication. In such case voice recognition may be essential as well as speech is recognized. Such a model consists of speech recognition and speaker recognition methods applied. Speech recognition already described is likely to be – (i) Feature Extraction using Mel Frequency Cepstral Coefficient (ii) Codebook Generation using Vector Quantization Algorithm and stored in database. When an input acoustic speech in given in the model, it is compared and identified using Hidden Markov Model. Beside speaker recognition system uses one more method of Feature Matching for speakers' voices using DISTMIN algorithm [12].

Also, these models are not sufficient to recognize the context in which the conversant may have their discussion over the phones. Moreover, almost every speech recognition model receives acoustic voice data as input and analyze to detect the utterance so that methods discussed working upon analog signal whereas cellular network and modern Public Service Telephone Network (PSTN) is dependent upon digital transmission in which analog signal is converted into digital form at the receiver's handset and back to analog form at the destination's ear only. Therefore, digital speech is needed to be converted into discrete time analog samples again to be analyzed for speech recognition. And the voice conversations are needed to be captured in both intermediate and main switching stations.

IV. A new framework

Our proposed idea consists of a combination of all above models described. Mobile handset users usually talks moving from one position to another even from one cell to another. While capturing speakers' speeches, both BTS and MTX are to be supported by extra storage backup to capture the speeches of conversant. While the conversant are in one cell, speeches may be captured in respective BTS but while either or both are in another cell, the speeches will be captured by BTS of that respective cells. An accumulator or dedicated crawler can collect the split conversations and kept in single database table. A database with the

attributes like sender and receiver's mobile number, call duration and voices are needed to be built in which the captured voices are to be stored. Such a database is beyond the limit of conventional database since variable length audio is to be stored in database whereas cell phone numbers consist of integer number for dialing nationwide except a special character is added before international dialing code, call duration is also numerical value and phone users' names are string type data. Variable length audio files for each conversation can be linked by pointer with respective record of cell phone users' details. Digital form of audio data should be converted into analog signal or discrete time analog samples and analyzed using model as described in early paragraphs to detect the words and linguistics for forming the sentences. Here the audio files are being converted into text files by which audio files can be replaced to reduce the multimedia database into conventional database. The text form of conversations can be searched using the desired keywords which may mean the possible unethical activity involvement of the cell phone users in case of sufficient match found in deduced search result. Crawler software is needed to be executed round the clock to catch all the conversations, store in database, convert audio speeches into text files and search the text conversation to hunt unethical activity.

V. Conclusion and Recommendations

A picture has been tried to be drawn in this article in which speakers' speeches have been captured and stored in computer storage medium in multimedia database, speeches converted into analog since to be analyzed as acoustic data and words are extracted and sentences are found. Searching is carried out on the text files of conversion speeches later on. Dedicated crawler software is proposed to carry out the job.

Still there is scope of development for recognition of the context of conversation which is essential because it will make easy to understand about the topic discussed between two people what their intention for doing in upcoming times. Moreover, a legal framework is also required since confidentiality of peoples' conversations over telephone network are taken into storage which is actually an act like intruders and the purpose is to disclose though to authorized persons of respective authority like law order enforcing agencies or government intelligences agency of Bangladesh. Also, it is unpredictable the volume of storage requirements to hold the conversations in database. A research should be carried out to estimate the requirement of storage and how long the conversations are to be existed in the database is necessary.

References

[1]. Andrej Trnka, Big data analysis, European journal of science and theology, Vol. 10, Suppl.1, 143-148, October 2014
[2]. Manish Kumar Kakhani, Sweeti Kakhani and S.R. Biradar, Research Issues in Big Data Analytics, International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 2, Issue 8, August 2013
[3]. Travis B. Murdoch, Allan S. Detsky, The inevitable application of big data to health care, Journal of American Medical Association, Vol. 309, No. 13, April 2013
[4]. William Stallings, Data and Computer Communication, 5th Edition, ISBN 81-203-1240-6, Prentice Hall, 1997
[5]. Andrew S. Tanenbaum, Computer Networks, Ch 1, pp 1-76, 3rd Edition, ISBN 81-203-1165-5, Prentice Hall, 1997
[6]. Saleh Faruque, Cellular Mobile System Engineering, ISBN 0-89006-518-7, Artech House, 1996
[7]. Kaven Pahlavan, Prashant Krishnamurthy, Principles of Wireless Networks – A unified approach, ISBN 81-203-2380-7, Prentice Hall, 2002
[8]. William C. Y. Lee, Mobile Cellular Telecommunications – Analog and Digital Systems, 2nd Edition, ISBN 0-07-113479-4, McGraw-Hill, 1995
[9]. A. M. Ali, M. Hossain, M. N. Bhuiyan, Automatic Speech Recognition Technique for Bangla Words, International Journal for Advanced Science and Technology, Vol. 50, January 2013
[10]. M.A. Anusuya, S.K. Katti, Speech Recognition by Machine: A Review, (IJCSIS) International Journal of Computer Science and Information Security, Vol. 6, No. 3, 2009
[11]. Preeti Saini, Parneet Kaur, Automatic Speech Recognition: A Review, International Journal of Engineering Trends and Technology-Vol. 4, Issue 2, 2013
[12]. Suma Swamy and K.V Ramakrishnan, An Efficient Speech Recognition System, Computer Science & Engineering: An International Journal (CSEIJ), Vol. 3, No. 4, August 2013