BARC : Television Audience Measurement

A presentation by Team IIT Kanpur

Understanding the Problem

We propose the following four modules for solving the 5 listed problems effectively:

Acoustic Fingerprinting

Remote IR Decoding

Speech-to-Text

PIR Room Occupancy Detection

Fingerprinting Algorithm

Fingerprint Hashing

By looking at our spectrogram peaks and combining peak frequencies along with their time difference between them, we can create a hash, representing a unique fingerprint for a audio file.

Fingerprinting Channels

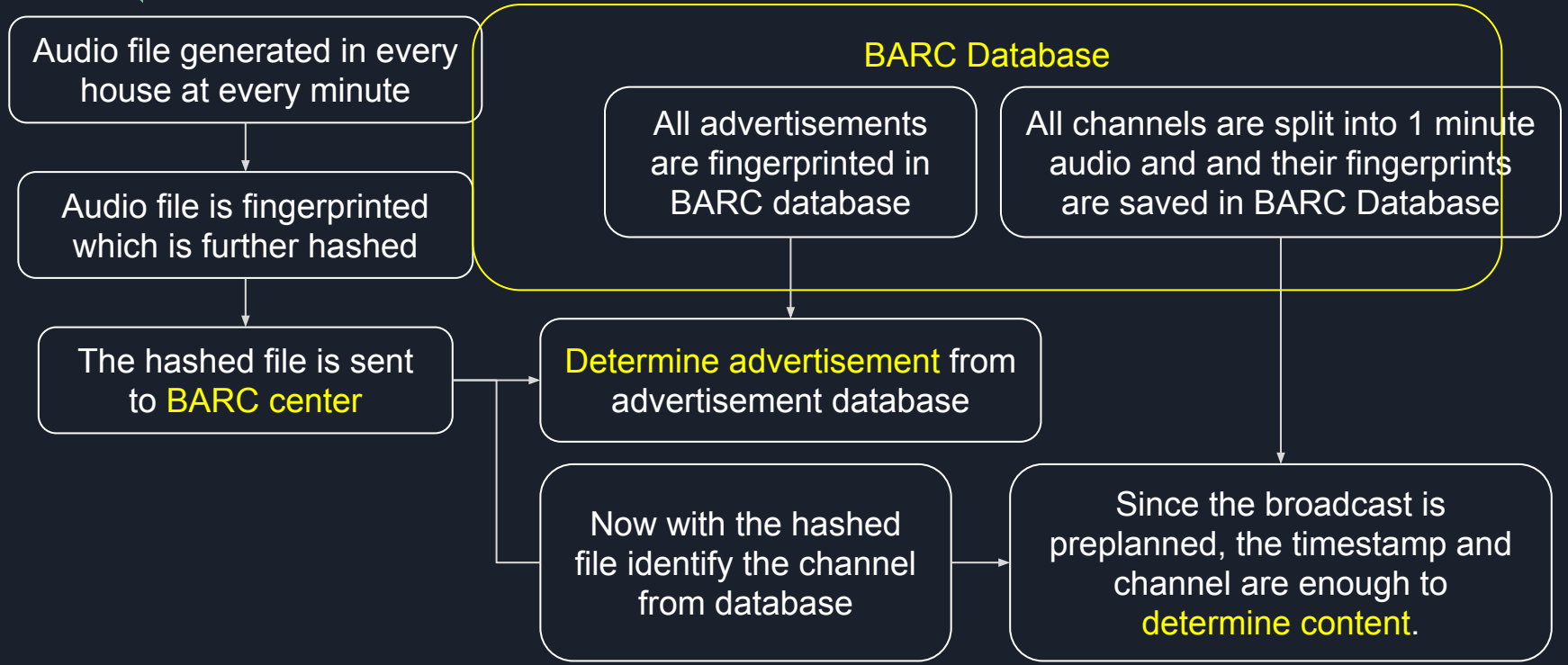
The MySQL Database Schema contains two tables viz.

- ➔ Fingerprints
- ➔ Name of the Channel

Fingerprint Alignment

The hashes we extract out of the sample will have an offset that is relative to the start of the sample. But this offset does not matter as long as the track is being played and sampled at the same speed it was recorded and released from the broadcaster, which is the case in Television.

Overview of the Solution



Speech-to-Text

A Speech-to-Text module has been implemented which records audio, splits into 1 minute intervals and converts it into text and send only the keywords of the advertisements along with the time stamp.

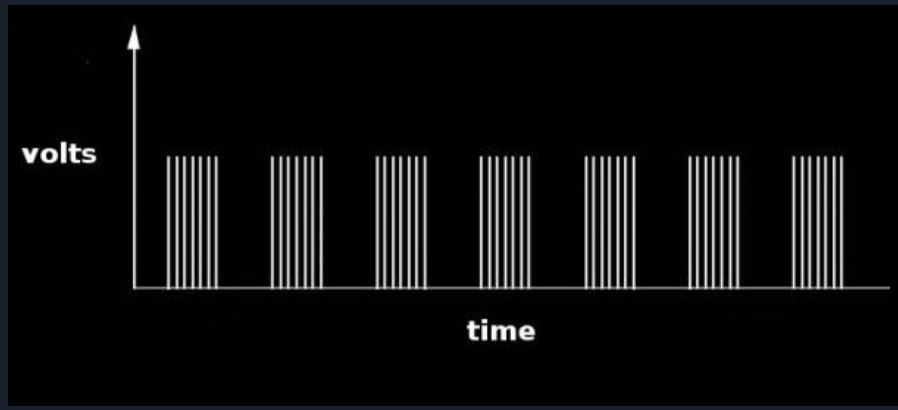
The keywords like “Airtel”, “Vodafone”, “Cadbury Dairy Milk” , “Amul” have been added to our simulation of BARC Database, which is a one-time effort.

This also helps estimate the channel being watched to some extent along with the exact advertisement being watched.

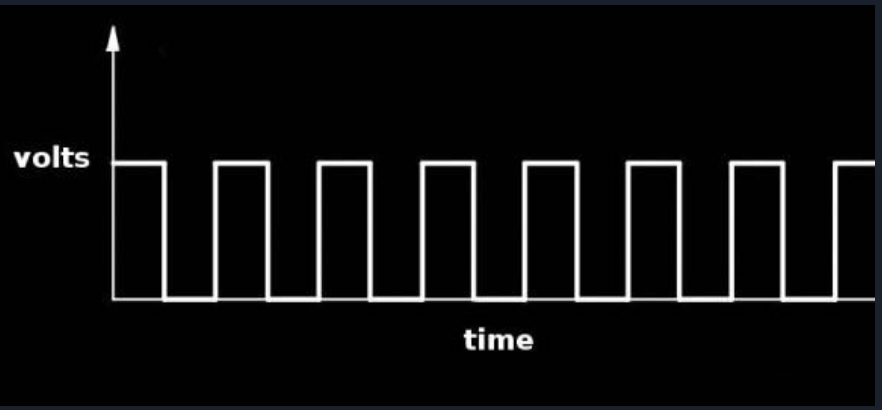
Remote IR Decoder

The IR remote generates an IR signal upon button press as shown below. The receiving end demodulates this signal and outputs a binary waveform that can be read by a microcontroller.

Output of TV Remote IR



Signal received by TSOP



Decoding these signals into Hex, we get the channel number pressed on the remote to map TV Navigation and identify channel being watched and can easily be implemented on TV Guide based TV navigation.

Who is watching?

Analysis of technologies available for person identification

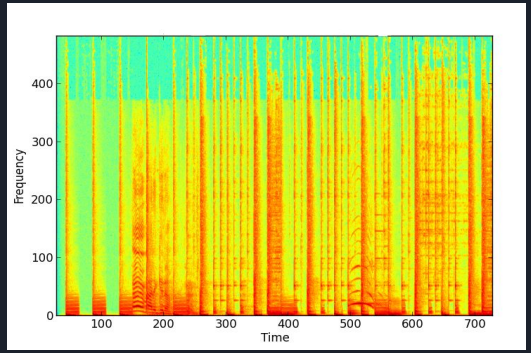
	Technology	Remarks
1	Computer Vision	Invades privacy
2	Device sniffing	Needs Smartphone and WiFi connectivity
3	Proximity sensors (Beacon technology)	Needs Smartphone and WiFi connectivity
4	CSI technology(WiFi based person identification)	WiFi connectivity, requirement of routers and costly CSI extraction supporting receivers
5	**Infrared Sensors (Passive infrared sensors)**	**Cheap technology, non intrusive, easy setup and preserves privacy**

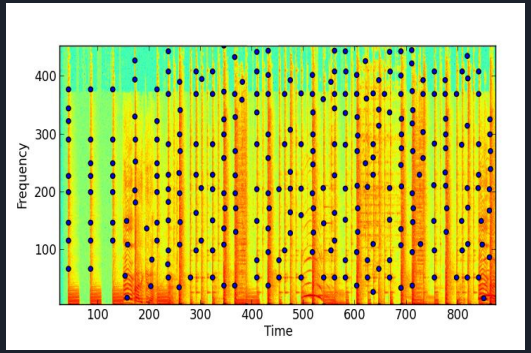
Acoustic Fingerprinting

01 Digitally encoded audio is stored as a list of numbers. An uncompressed 3-minute .wav audio file has almost 16 million of such samples which are a signal of sorts, hence Fast Fourier Transform over small time steps to create its spectrogram.

02 From the spectrogram, peaks are found in amplitude. Other (time, frequency) pairs around these peaks are lower in amplitude, and are thus less likely to survive noise.

03 Once these noise-resistant peaks are extracted, we have found points of interest in a audio that identify it. We are effectively squashing the spectrogram down once we've found the peaks.



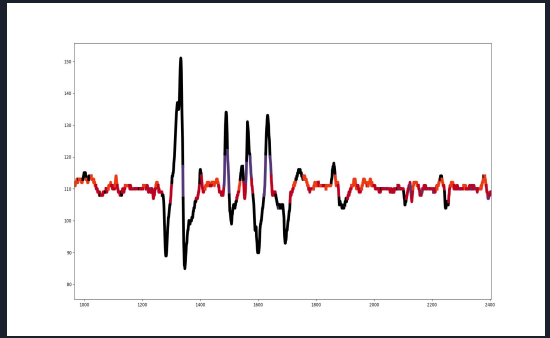


Room Occupancy Counting with a single PIR sensor

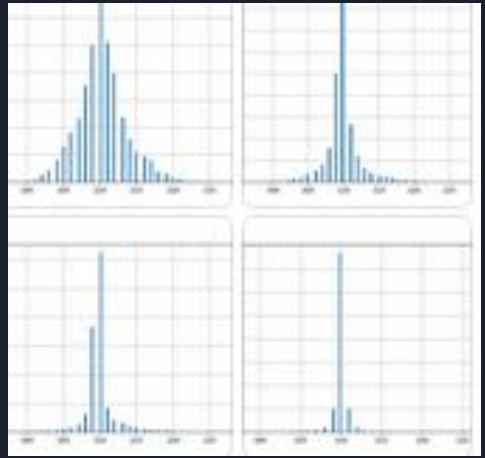
A PIR Sensor is employed to estimate the number of people watching the television with the use of iHMMs

PIR or Passive Infra-Red Sensor detects the body heat radiations to detect motion around it and is sensitive to direction.

The hidden Markov model (HMM) is a widely used probabilistic model for segmentation of time series data and has been successfully used to model behavioral patterns of sequential data from different sources.

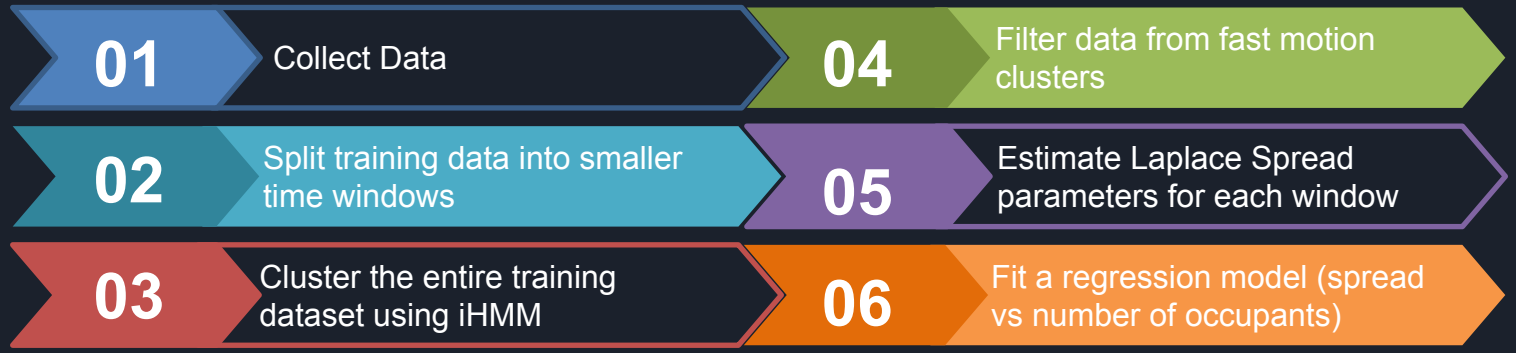


Motion Classification using iHMMs
Red : small motion behaviour
Black : fast motion (filtered out)



Variation in Histograms for 1,2,3,4 people in a room.

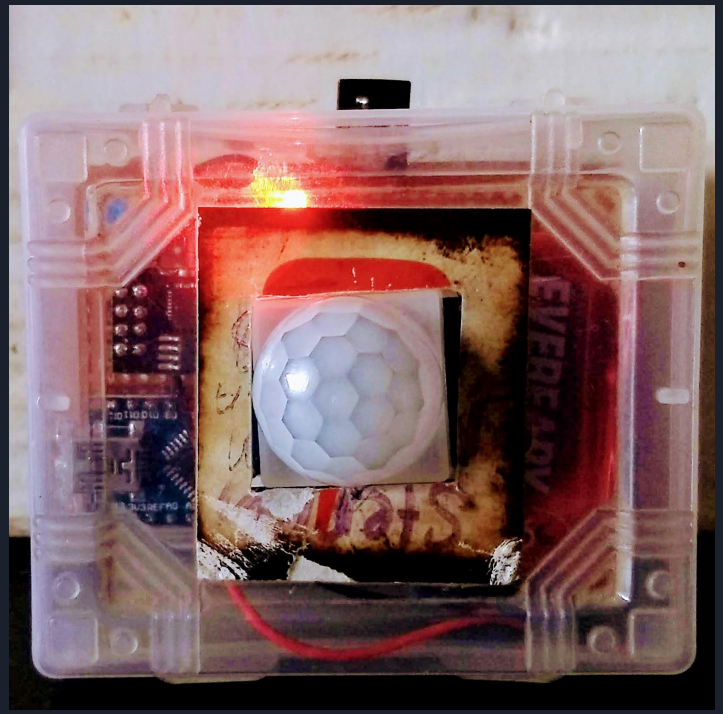
Architecture of the proposed room-occupancy counting system



- Large temporally local fluctuations in the PIR output reflect some temporally local human behavior and will bias the occupancy count estimate.
- The iHMM groups together motions that are temporally similar, so larger movements would be clustered in separate groups and we can easily filter them out.

Using Multiple Sensors, Gait Movements of people can be detected and can be classified into Age Groups and Gender. Thus we know, “Who is watching?”

Prototype I - Stand Alone Device



Remote IR Decoding and Room Occupancy Counter are incorporated in this prototype. This is a stand-alone device.

Component	Cost per Unit	Quantity Needed	Total Cost
PIR Sensor	INR 95	1	INR 95
TSOP	INR 25	1	INR 25
Arduino Nano	INR 250	1	INR 250
ESP8266	INR 150	1	INR 150
Miscellaneous	INR 30	-	INR 30
Total Cost			INR 550

Overall Cost Analysis

Audio Fingerprinting and Speech-to-Text are accomplished by just a Generic Microphone and an on-board computer (Raspberry Pi Zero) in the prototype. Cost for this prototype is INR 1000

For the final product, the audio out from the TV will be directly forwarded to BARC Servers using a GSM based connection, hence we get rid of the ESP8266.

These prototypes are currently built on hobby devices, (Arduino and Raspberry Pi). Building these on a final product will surely reduce extensively and then finally mass production will reduce the costs further by 7%.

Cost Effect	Cost Reduced	Effective Cost
Prototype	-	INR 1550
Final Product	INR 1015	INR 535
Mass Manufacturing	INR 35	INR 500
Overall Cost	-	INR 500