

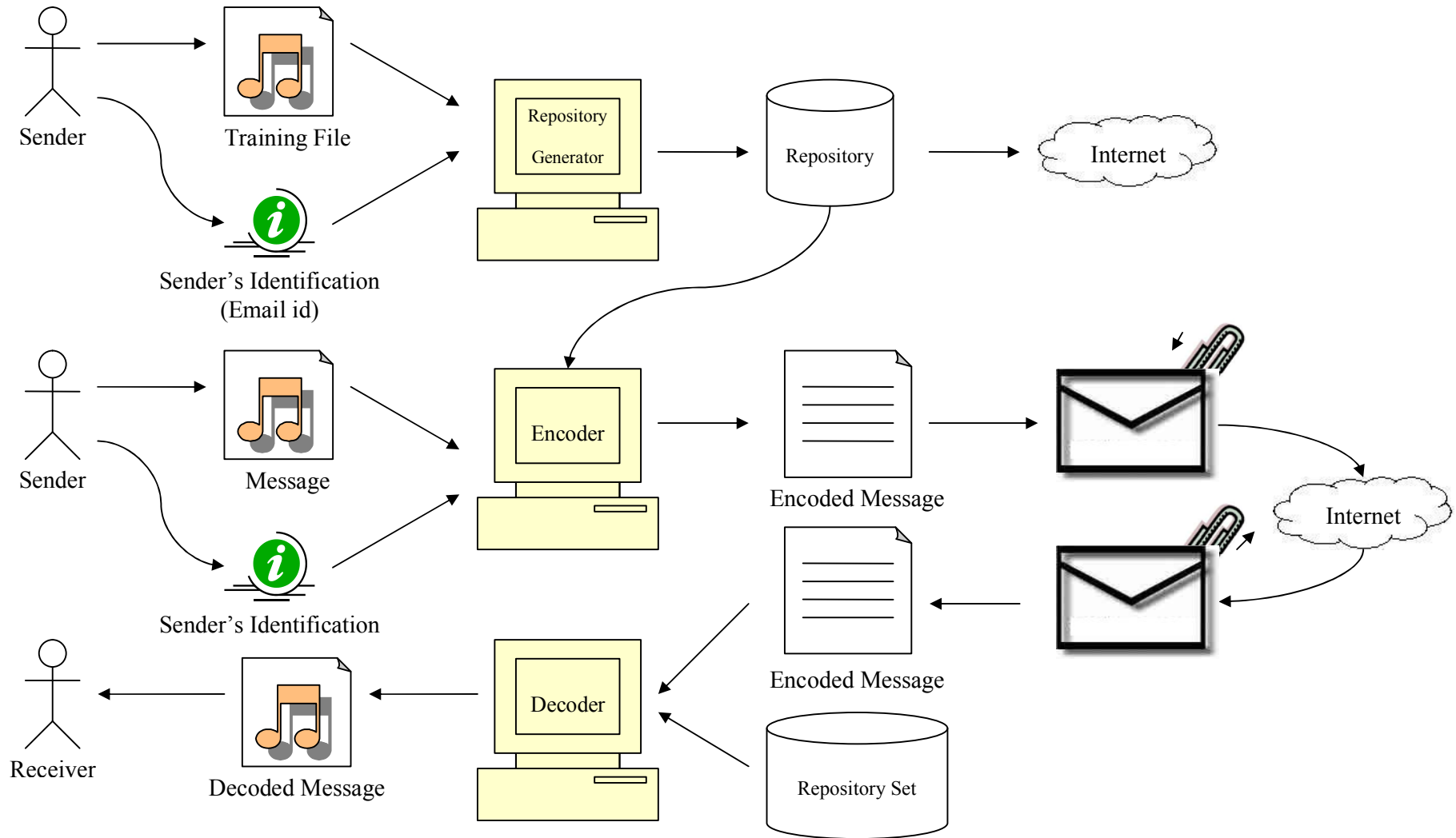
# Speech coding using personalized speech repository

Manish S. Gupta

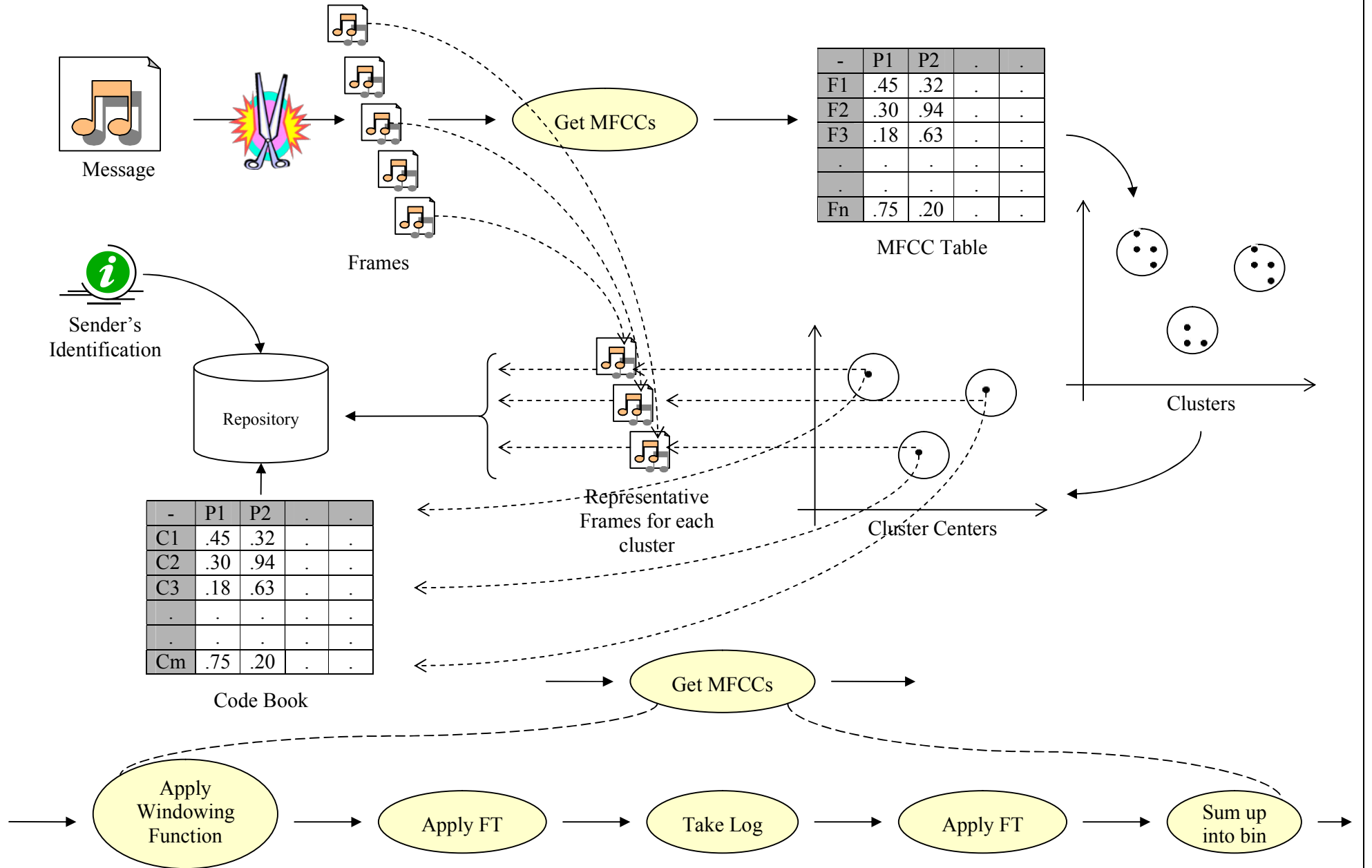
Apoorv B. Shrivastava

Sumeet D. Mehta

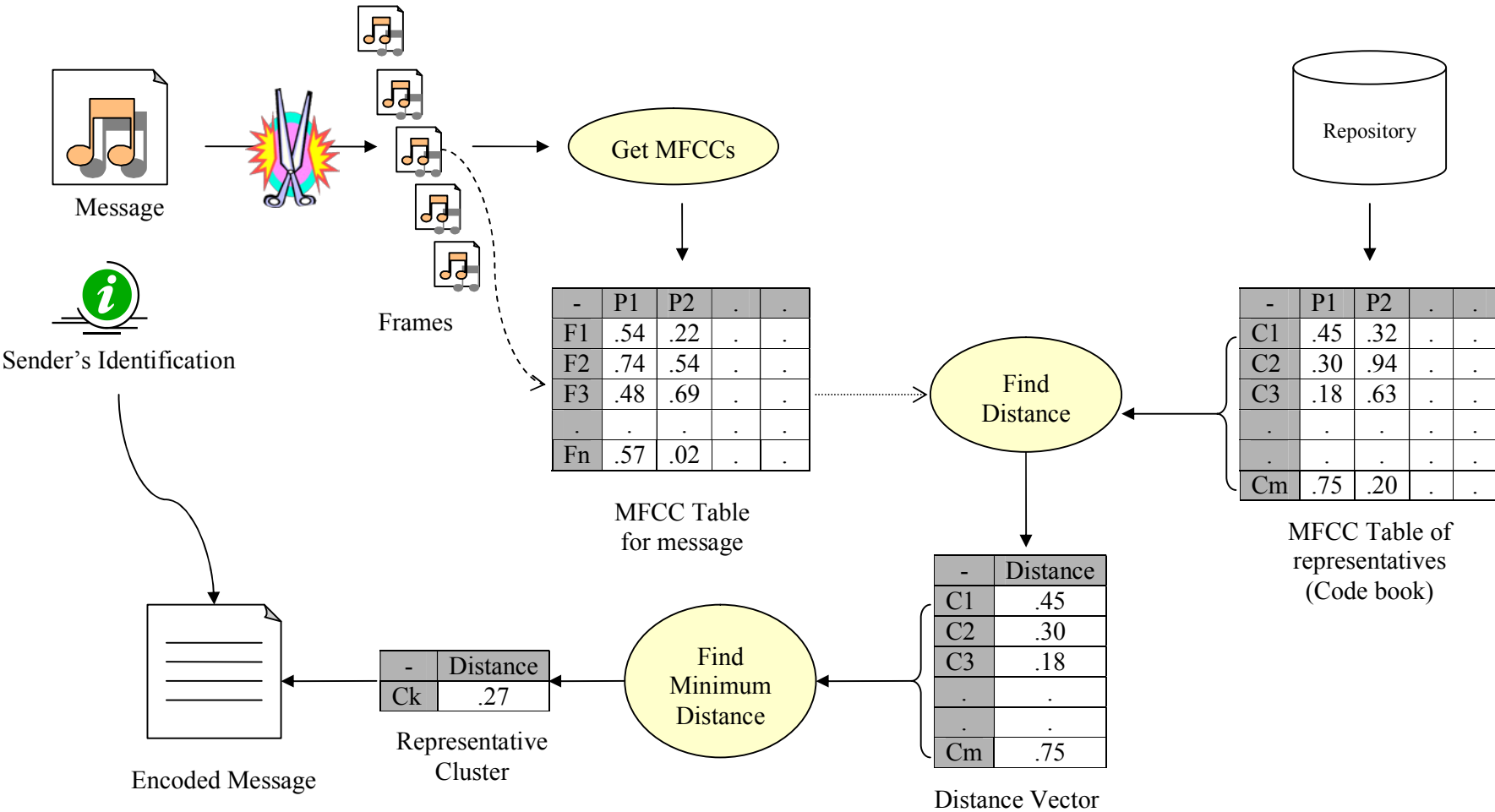
The project involves building a system for exchanging voice messages over mail using speech coding, resulting in tremendous speech compression. The sender will record his voice message and transform it into the coded, compressed file using the **encoder module**. The coded file is transferred as an **email attachment**. The receiver passes the attached file through the **decoder module**, which reproduces the original speech. Both the encoder and decoder will use a **repository** of speech segments created using the **repository generator**. The repository may be transported by CDs, or may be made available for download.



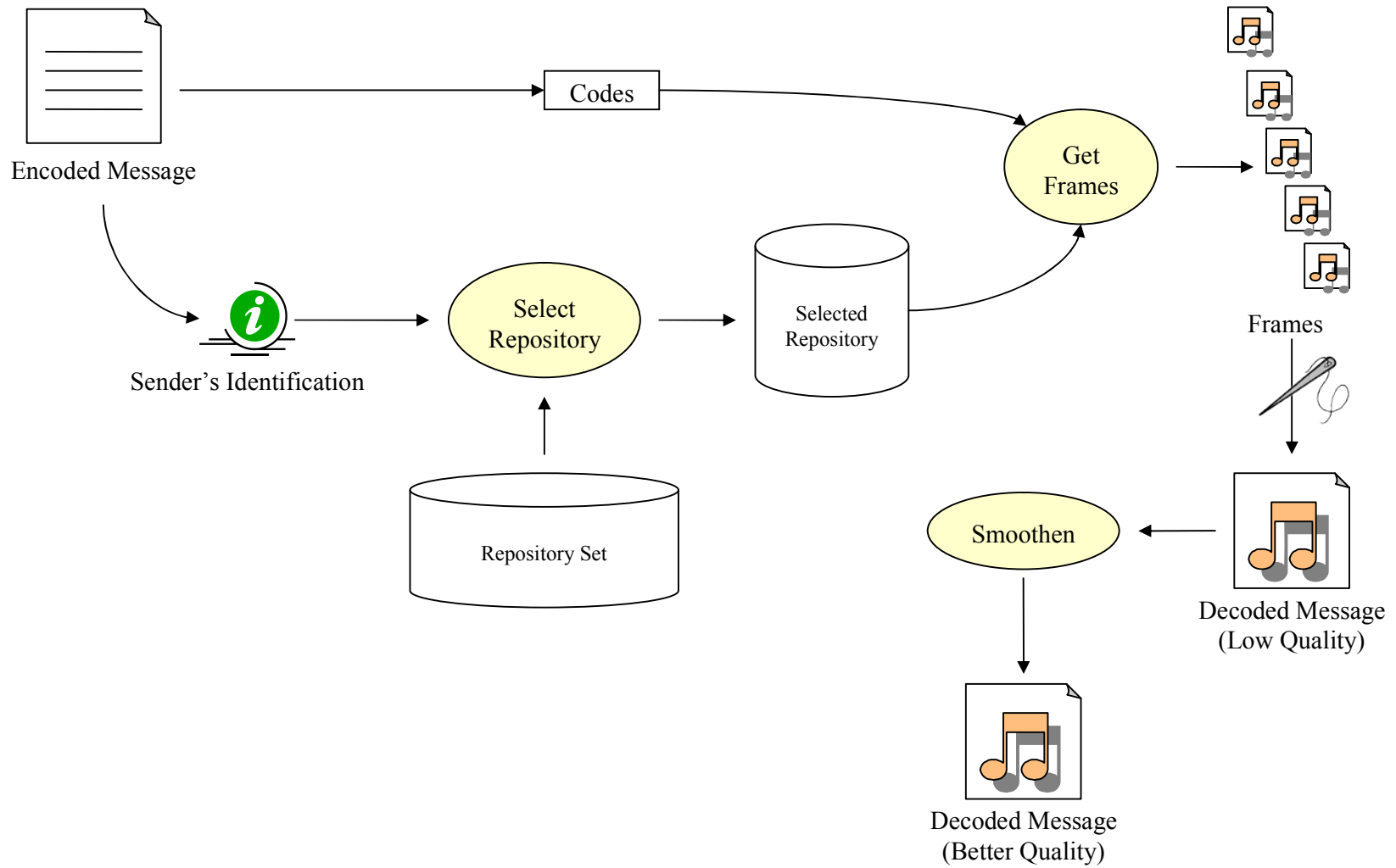
# REPOSITORY GENERATOR



# ENCODER






























# DECODER






## CLASSES IDENTIFIED






### wavefile

 ChunkID : char\* = "RIFF"  
 ChunkSize : unsigned long int  
 Format : char\* = "WAVE"  
 Subchunk1ID[4] : char\* = "fmt "  
 Subchunk1Size : unsigned long int  
 AudioFormat : unsigned short int  
 NumChannels : unsigned short int  
 SampleRate : unsigned long int  
 ByteRate : unsigned long int  
 BlockAlign : unsigned short int  
 BitsPerSample : unsigned short int  
 Subchunk2ID : char\* = "data"  
 Subchunk2Size : unsigned long int  
 data : char\*  
 path : char\*  
 curloc : unsigned long int  
 locinc : unsigned long int





 wavefile()  
 wavefile()  
 wavefile()  
 getFirstFrame()  
 getNextFrame()  
 appendFrame()  
 commit()  
 toMono()  
 ~wavefile()  
 makePerm()













### frame m f c t a b l e

 frame table : frame \*\*  
 m f c t a b l e : double \*\*  
 count : int\*



 frame m f c t a b l e()  
 getFrame()  
 addFrame()  
 delFrame()  
 ~ frame m f c t a b l e()







### clustermanager

 current : long int  
 dist[k] : double\*  
 cluster\_centers : double \*\*  
 fm tbl : frame m f c t a b l e



 clustermanager()  
 read\_frame()  
 init\_centroids()  
 distance()  
 minimum()  
 recalculate1()  
 showcenters()  
 addtorepository()  
 ~clustermanager()  
 start()  
 getallclustercenters()  
 compare()

### filemanager







 list : char\*\*  
 count : int




 newTempFile()  
 checkTempFile()  
 delTempFile()  
 makePerm()  
 filemanager()  
 ~filemanager()

### frame





 frame length : unsigned long int  
 getMFCC()








### code file

 path : char\*  
 file ptr : FILE\*  
 emailid : char\*  
 hash : unsigned int  
 codestream : unsigned int  
 curloc : unsigned long int

 append()  
 read()  
 reademailid()

### repositorymanager

 cluster\_centers : double \*\*  
 filename : char\*\*  
 count : unsigned int  
 repositoryname : char\*

 getcode()  
 makeNewFileName()  
 getfilename()  
 repositorymanager()  
 getclustercenter()  
 ~ repositorymanager()  
 selectrepository()

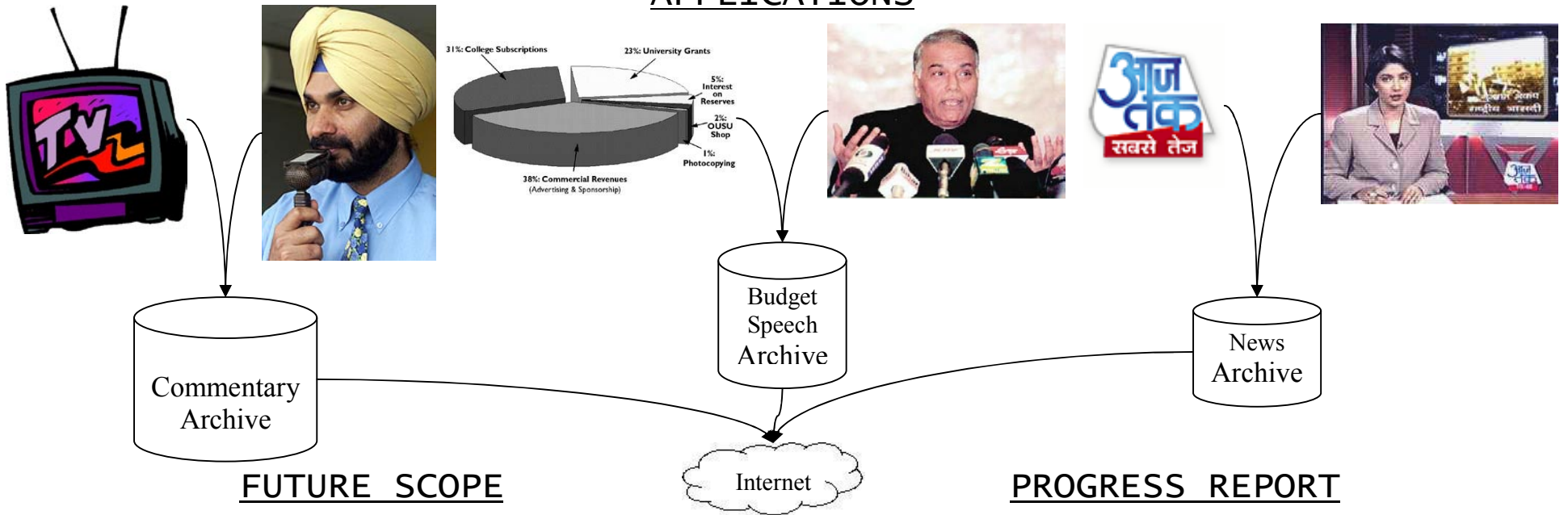
## ADVANTAGES

- Efficient Bandwidth Usage
- Clarity Of Communication
- Easy to use package
- Reusable as a shared library

## LIMITATIONS

- Single Speaker Voice allowed
- Heavy resource utilization
- Fixed Frame Length used
- Empirically decided cluster cardinality

## APPLICATIONS



## FUTURE SCOPE

- Streaming (i.e. real time communication)
- Segmentation (Variable length frames)
- Conferencing (handling speech of more than 1 person)
- Better quality using better prosody modification techniques
- Replacement of WAV format

## PROGRESS REPORT

- |             |             |
|-------------|-------------|
| ● Analysis  | completed   |
| ● Designing | completed   |
| ● Coding    | in progress |
| ● Testing   | to be done  |

Visit us at <http://vox.sourceforge.net>

Project Guides: Mrs. Vijayalakshmi & Mr. Chetan Vaity