Compressed Sensing based Audio Sensor Data Classification in IoT

Submitted By-

Mandeep Kumar

(MIT2020019)

Submitted To-

Dr . Manish Kumar

(Associate professor)

OUTLINE

1. Abstract
2. Introduction
3. Problem Statement
4. Literature review
5. Proposed Methodology
6. Dataset
7. Results
8. References

Abstract

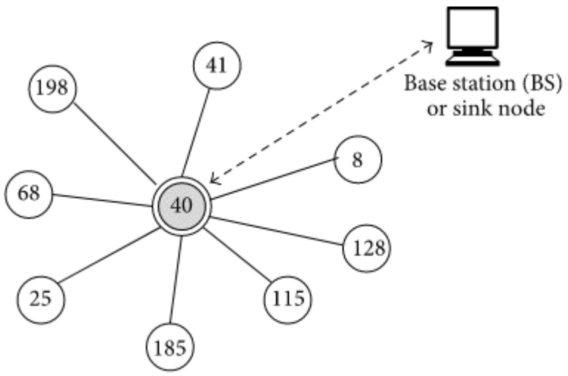
- Forests play a vital part in our daily lives because they have numerous advantages such as oxygen, shelter for animals, and so on. However, illegal deforestation is now increasing forest destruction. For forest safety, IoT sensors are used.
- IoT has been used in a wide range of applications, namely industrial automation, military, transportation, and environmental monitoring.
- Since these sensors have little resources,as the number of sensors grows, so does the need for new mechanisms to mitigate parameters like power usage, expense, latency, and traffic congestion.
- Compressed Sensing (CS) theory can help with these statistics.The CS demonstrates how sparse signals can be precisely reconstructed using just a few random linear measurements.

Introduction

The Internet of Things (IoT) is a new connectivity model that envisions a world in which "things," or computers equipped with sensors, actuators, electrical devices, transceivers, apps, and appropriate protocol stacks, would be able to connect with other devices with the goal of enhancing users' living standards.

Sensors

- Sensors are compact with minimal processing and computational power. This sensor nodes will detect, quantify, and gather environmental data, and then transfer that data to the user based on a local decision making process.
- Sensor units that are strategically placed in an inside or outside environment.The aim of the Sensor based IOT is to collect data from the environment.
- One of the major constraints of IoT sensors is power dissipation during signal transmission.As in Sensors power and storage available is less which leads to many challenges like a large amount of storage and transmission.
- WSN was the first to use compressed sensing. after it was determined that reducing the payload was a successful process.Many applications naturally produce from this kind of sparse (or compressible) structure.

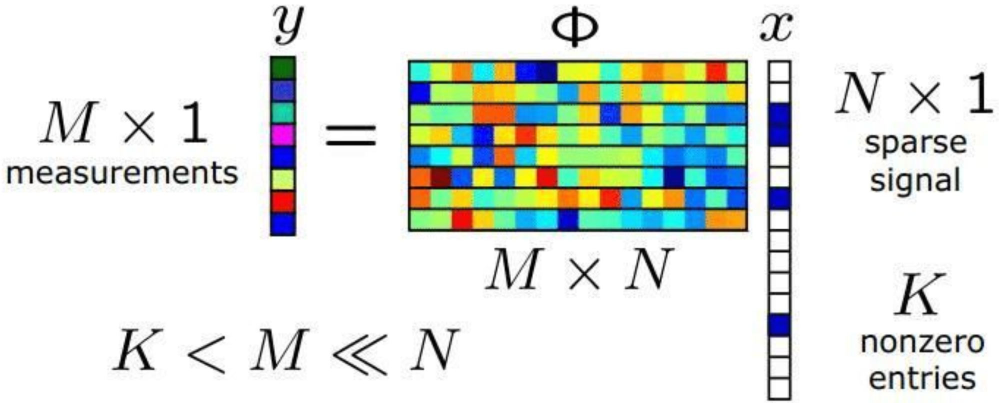


Compressive Sensing

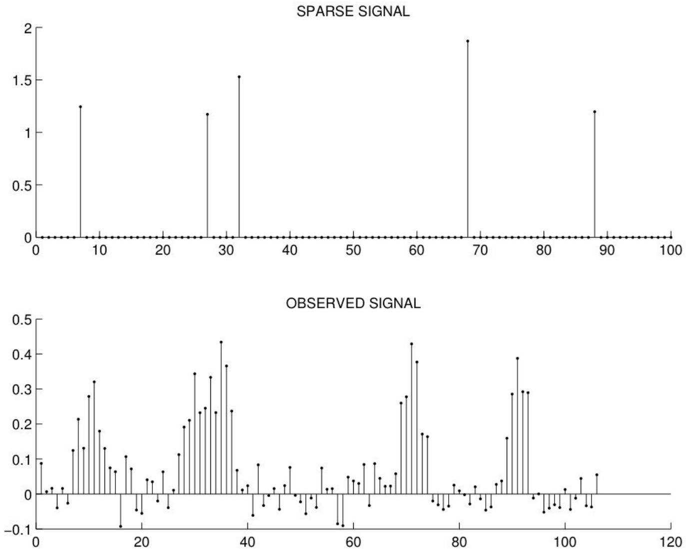
CS relies on two principle:

1. Sparsity: which pertains to the signal of interest.

2. Restricted Isometry Property



Sparsity and Sparse Representation

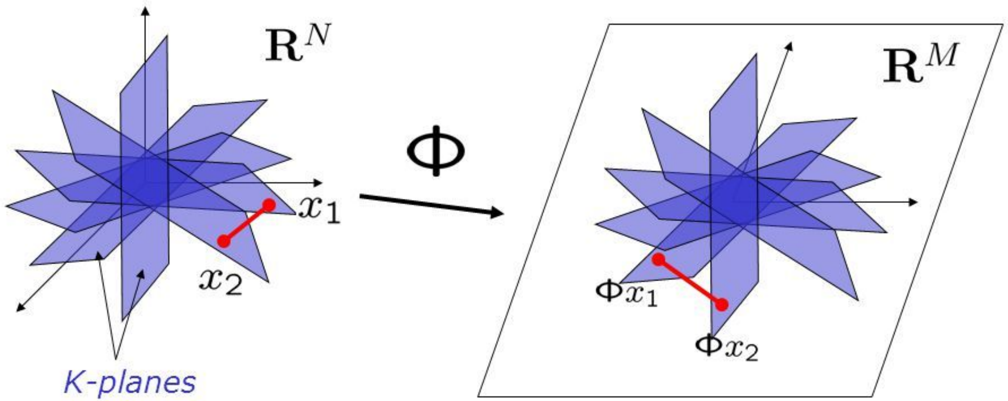


Sparsity of a signal is defined as the number of non-zero elements in the signal under a certain domain.

Restricted Isometry Property(RIP)

- Preserve the structure of sparse/compressible signals
- RIP of order *2K* implies: for all *K*-sparse x_1 and x_2

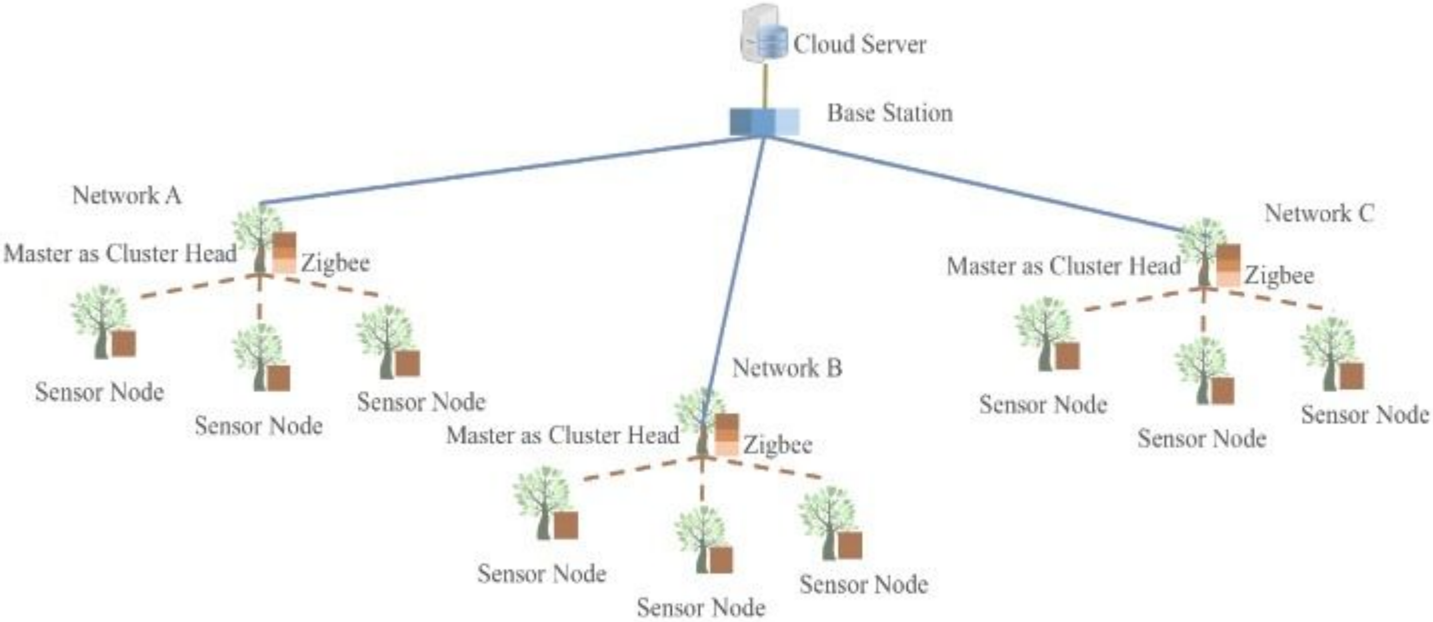
$$(1-\delta_{2K}) \leq \frac{\|\Phi x_1 - \Phi x_2\|_2^2}{\|x_1 - x_2\|_2^2} \leq (1+\delta_{2K})$$



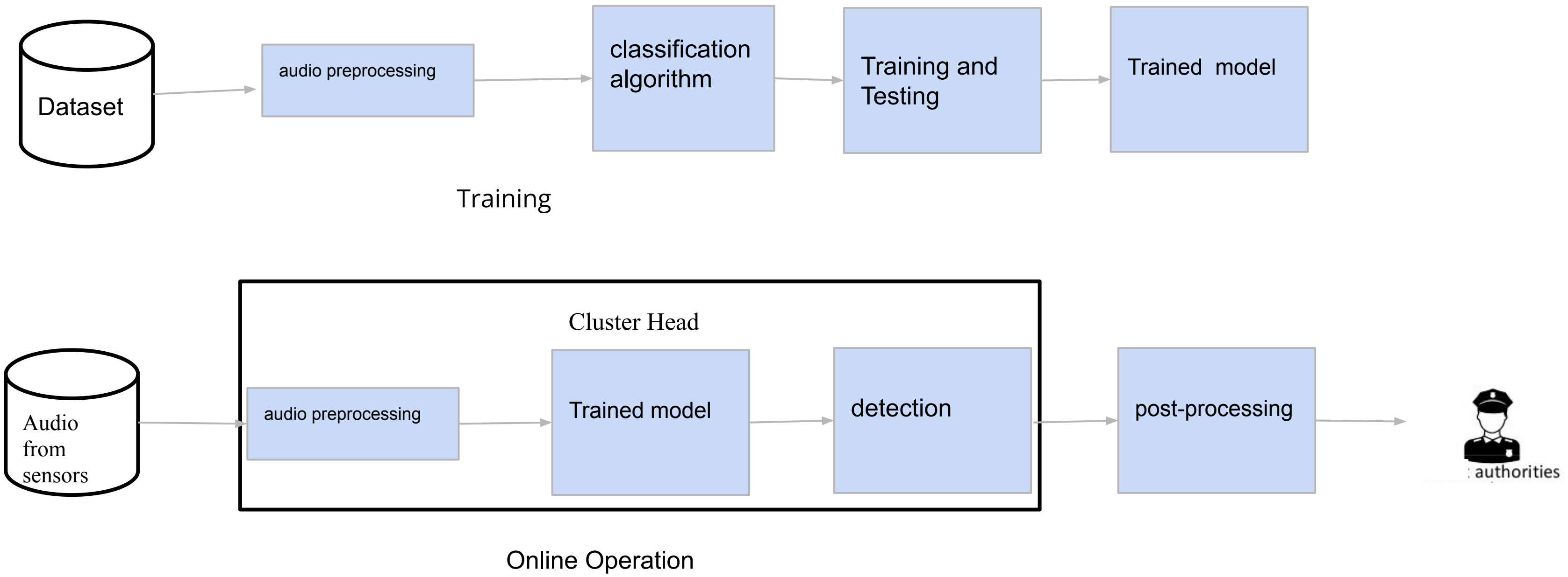
Problem Statement and Motivation

- In the present paradigm data(signal) received from the sensors is sent to the cluster head.
- Here most power is consumed while transmitting the data,in order to do it more efficiently we can use the technique in which sampling and compression is done in one step that helps in eliminating the unnecessary data and store only required data which can be reconstructed when needed.
- CS technique is helpful as the structure of data is the same as before compression ,so we can **directly Classify Compressed Signals**.
- In this work the objective is to classify different audio signals which are present in the audio sample.Here audio sample is the preprocessed output of compression sensing,when compressive sensing is applied on original audio.
- Compressive sensing is able to generate audio samples from original audio on the basis of some of the audio features that makes our audio sample equivalent to original audio in terms of characteristics but less in terms of size.
- So our first challenge is to determine those features among all the audio features like amplitude,mfcc,Spectral Bandwidth,Spectral Roundoff etc.
- After determining those features we will apply preprocessing on the original data to compress data. And we train our model on the preprocessed data and then try to classify data.

Network Model



Proposed Method



Literature Survey

Chen, J.-T.; Lin, C.-B.; Liaw, J.-J.; Chen, Y.-Y. et. al [1], addressed logging detection and implemented a tool for detecting illegal tree logging in mountains using vibration . They used simple subtraction of two data to obtain differential signal intensity as a vibration functionality in their work. The results of this study showed that the method could identify between sawing wood disturbances and human physical vibrations. The authors' vibration signal designs that used sound amplification demonstrated a clear better accuracy, as well as performance improvement all in all.

In Prasetyo, D.C.] et. al [2], the authors presented a prototype system aimed at detecting illegal logging, which was focused on the use of both sound levels sensors and vibration sensors were used to sense chainsaws, and vibration sensors were used to identify tree falling in forests. The Arduino Nano model was used, and GSM modules were used to provide information to the forest guard patrols.According to the findings, the value of 63.4 dB for chainsaws, as well as the vibration sensor threshold of 4400, is appropriate for detecting illegal logging.

In [3] Ahmad and Singh suggested a technique for detecting tree slicing in forests that relied on insulative impedance and the difference between parameters. properties that were based on the distance between parameters, and also utilized Appl. Sci. 2020, 10, 7379 3 of 12 Gaussian mixture measures (GMM), principal component analysis (PCA), and k-means clustering.Their strategy worked well, with precision of up to 92 percent in thick forests and up to 76 percent in open forests.

In [4] a device focused on wireless sensor networks and multiple sensors was developed to detect and identify illegal tree cutting. Tone and vibration sensors were used in the network nodes. The Xbee Pro S2C module was utilized as a communication medium and the Arduino Nano was used for data processing procedures. The system introduced by the authors was tested in small forest and open area scenarios and the findings showed that the authors' work was cost-efficient and had a promising performance.

In [5] Based on the extraction of Haar-like characteristics, a method chainsaw sound tracking was developed.Using frequency-domain function extraction, the approach aimed to interpret and distinguish signals from audio sources. Haar-like characteristics were identified from the spectrogram. To distinguish chainsaw from non-chainsaw sounds, the system used a two-stage thresholding technique. The results of the study indicated that Appl. Sci.2020,10, 73794 of 12the method was very effective in recognizing chainsaw sounds and that it could effectively perform this discrimination in forests.

Dataset

Link: https://urbansounddataset.weebly.com/urbansound8k.html

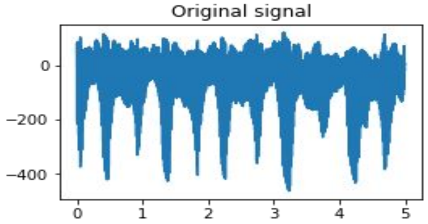
Dataset used in this task contains around 8000+ sounds.
Length of each audio clip is less than equal to 4s. This dataset classes are then mapped to unique number.The number of folds (1-10) that this file has been assigned to.

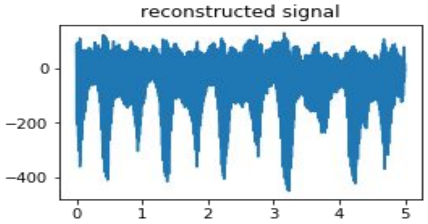
Id to class mapping:

air conditioner	=>0
car horn	=>1
children playing	=>2
dog bark	=>3
Drilling	=>4
engine idling	=>5
gunshot	=>6
jackhammer	=>7
Siren	=>8
music	=>9

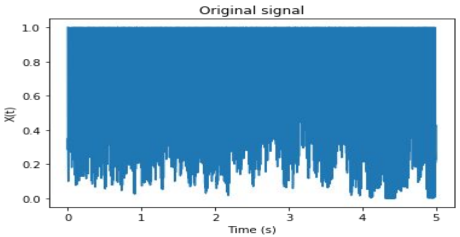
Results

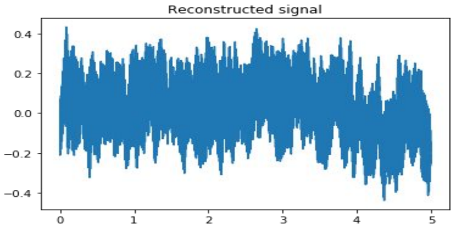
Mel-Frequency Cepstral Coefficients(MFCCs)





Chroma feature





- MFCCs
- Chroma feature
- Spectral Centroid
- Spectral Rolloff
- Spectral Bandwidth
- Zero-Crossing Rate

SVM VS CNN

	SVM		CNN	
	Normal	After CS	Normal	After CS
Accuracy	70%	34%	63%	61%
Training time	12519s	9519s	22843s	20466s

Training Results

With CS

```
Epoch 499/500
3360/3360 [==============================] - 41s 12ms/step - loss: 0.4538 - accuracy: 0.8446
Epoch 500/500
3360/3360 [==============================] - 41s 12ms/step - loss: 0.4490 - accuracy: 0.8414
2240/2240 [==============================] - 10s 4ms/step
('Test score:', 2.2377496985452514)
('Test accuracy:', 0.6102678775787354)
Training took: 20466 seconds
```

Without CS

```
Epoch 497/500
3360/3360 [==============================] - 45s 13ms/step - loss: 0.3981 - accuracy: 0.8634
Epoch 498/500
3360/3360 [==============================] - 45s 14ms/step - loss: 0.4246 - accuracy: 0.8530
Epoch 499/500
3360/3360 [==============================] - 45s 13ms/step - loss: 0.4246 - accuracy: 0.8527
Epoch 500/500
3360/3360 [==============================] - 45s 13ms/step - loss: 0.4036 - accuracy: 0.8607
2240/2240 [==============================] - 9s 4ms/step
('Test score:', 1.8935086045946394)
('Test accuracy:', 0.6312500238418579)
Training took: 22843 seconds
```

Classification Report

	precision	recall	f1-score	support
0	0.31	0.52	0.39	21
1	0.17	0.32	0.22	19
2	0.07	0.09	0.08	22
3	0.33	0.50	0.40	2
4	0.00	0.00	0.00	13
5	0.33	0.04	0.08	23
6	0.00	0.00	0.00	18
7	0.25	0.28	0.26	32
8	0.53	0.58	0.55	50
9	0.33	0.14	0.20	7
micro avg	0.29	0.29	0.29	207
macro avg	0.23	0.25	0.22	207
weighted avg	0.27	0.29	0.26	207

Confusion Matrix

predicted label \ true label	0	1	2	3	4	5	6	7	8	9
0	224	27	46	18	45	26	27	29	40	1
1	1	197	5	1	0	2	4	31	10	0
2	9	16	164	9	73	26	10	6	4	0
3	0	0	1	7	2	1	4	0	1	0
4	2	8	14	6	86	5	10	5	14	0
5	12	13	58	8	48	174	18	10	1	1
6	14	13	7	18	13	4	183	16	54	1
7	2	26	0	1	3	0	13	160	12	0
8	1	9	0	3	4	0	4	13	166	0
9	5	0	1	3	15	1	4	1	8	63

Conclusion

We proposed a method in which compression technique is used to compress the audio singal result obtained shows that SVM is not good for classification of compressed signals whereas CNN gives far better results.

References

1. Chen, J.-T.; Lin, C.-B.; Liaw, J.-J.; Chen, Y.-Y. Japan, 26–28 November 2018; Springer: Cham, Switzerland, 2018; pp. 212–219.
2. Prasetyo, D.C.; Mutiara, G.A.; Handayani, R.Motion and chainsaw sound sensor for uncontrolled logging The 2018 International Conference on Communication, Electronics, Renewable Energy, and Communications has released its proceedings.(ICCEREC), Bali, Indonesia, 5–7 December 2018; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2018; pp. 93–98.
3. Ahmad, S.F.; Singh, D.K. Using acoustic properties, automatically detect tree cutting in woods.. J. King Saud Univ. Comput. Inf. Sci. 2019.
4. Mutiara, G.A.; Suryana, N.; Mohd, O. To detect human cutting, several sensors are clustered on a wireless sensor network. Int. J. Adv. Sci. Eng. Inf. Technol. 2020, 10, 164–170.
5. Czúni, L.; Varga, P.Z. Time Domain Audio Features for Chainsaw Noise Detection Using WSNs. IEEE Sens. J. 2017, 17, 1.
6. Mohammad Abu Alsheikh ;Machine Learning in Wireless Sensor Networks:Algorithms, Strategies, and Techniques in Wireless Sensor Networks2014.
7. G. Simon, M. Maroti, A. Ledeczi, G. Balogh, B. Kusy, A. Nadas, G. Pap, J. Sallai, K. Frampton, Sensor network-based countersniper system, in:Second International Conference on Embedded Networked Sensor Systems Proceedings (Sensys), Baltimore, MD, 2004.
8. http://internetofthings.electronicsforu.com/2014/12/building-blocks-iot-getting-started/
9. https://www.coursera.org/learn/iot-augmented-reality-technologies/lecture/8ZlnC/iot-architecture
10. http://www1.cnnic.cn/ScientificResearch/LeadingEdge/wlw1/

THANK YOU