Noise Sources Classification in Map4Noise*: a crowd-sensing assisted approach

Yuan Lu §

The Department of Computer Science, New Jersey Institute of Technology

Abstract: Inspired by the increasingly serious noise issues in American cities, a crowd-sensing assisted real-time environmental noise classifier is implemented in the project of Map4Noise. The real-time noise classification can not only enhances the awareness of citizens to their noise exposure, but also helps the city planners to focus on the major noise sources and further reduce theses health threatens. The currently labeled noise databases are too small and too specific to train a robust probability based classifier for real-time noise classification. To handle with the shortage of labeled noise data, a crowd-source labeling function is designed in this project by introducing the idea of “human as sensor”. Mel-frequency cepstral coefficients (MFCCs) are extracted as features from original noise signals to protect users' privacy and balance the networking traffic between users and server. The probability based classifier in this project is implemented by hidden Markov models (HMMs) in parallel. Taking the advantage of crowd-source labeling, the HMMs also have the ability of algorithmically self-evolving. The re-estimation of inaccurate HMMs can be triggered with the accumulation of error rate and required no computation over other accurate HMMs.

Keywords: Hidden Markov model, Crowd-sensing, Human as sensor, Mel-frequency cepstral coefficient, parallel computation

---

* Project website: http://map4noise.njit.edu

§ Email address: yl768@njit.edu

I. Introduction

1.1 Motivation

Variously clinic evidences [1][2][3][4][5] indicate the exposure to environmental noise can cause both auditory and non-auditory health effects like hearing loss, annoyance, sleep disturbance, cardiovascular disease, cognitive impairment and even psychological symptoms. What's worse, noise pollution is pervasive in modern life. Approximately, 5% of the world population is suffering from noise-induced hearing loss [6]. In the European Union states, just the traffic noise makes 54% of population (56 million) exposed to unhealthy acoustic climate [7]. Noise also results great economic cost. In the US, noise-induced hearing loss is the most frequent occupational disease. About 22 million Americans are working in an environment with hazardous noise, and about $242 million dollars are spent on the compensation for occupational hearing loss per year [8].

A noise map can present measured or predicted noise exposure levels over a specified area. This visualized noise distribution will raise citizens' awareness of environmentally acoustic quality and also enhance the ability of city planners to successfully manage noise issues and further reduce noise-induced health threats [9]. Generally, the technical implementations of mapping noise can be classified into two categories, namely the noise simulation and the noise measurement. The computational predication is a feasible and inexpensive approach. The procedure of performing the simulation requires input parameters including noise source data (e.g. traffic flow and composition, average speed) and propagation environment (e.g. 3D digital terrain, meteorological conditions). Even though the simulation has merits, it also has obvious shortcomings such as the lack of accuracy and temporal dynamics which would be significant to evaluate annoyance and sleep disturbance. In contrast, the approach based on measurement can provide higher temporal solution. A grid of stationary devices is necessary to provide measured sound pressure and associated data. For example, Lille, France, deployed a noise measurement network, which is consisted by more than 80 distributed devices [10].

The noise measurement network can monitor the acoustic climate in real-time by fusing the data from every devices, and provide alerts for potential noise threats with small temporal granularity. However, to achieve higher spatial resolution, a much denser sensor network will be necessary, which greatly increases the difficulty of deployment and the cost of maintenance.

Although the traditional noise measurement network is neither feasible nor affordable for a large geographic area [11], the emerging sensor-enabled smart phones introduce a novel and promising technical approach to "install" a noise measurement network based on the mobile sensing [12][13]. Inspired by that, a few works [14][15][16][17][18][19][20] have been done on the topic of the mobile sensing based noise climate monitoring. The New York City 311 service [21] also implemented the function of noise monitoring since it utilizes "human as a sensor" to record noise events instead of quantified noise data. These applications are far from full-fledged crowd-sensing based noise monitoring apps, since they only focused on the basic service of measuring the personal and/or community exposure, and didn't emphasize on solving the challenges of noise source recognition, which can help to measure the seriousness of different kinds of noise pollution.

In our Map4Noise project, a real-time noise recognition application is implemented based on the crowd-source sensing with a focus on solving the known challenge. Comparing with the traditional implementations, the crowd sensing can greatly enhance temporal and spatial resolution for environmental noise monitoring, which makes the real-time noise alert possible. The idea of "human as sensor" is used to collect qualitative data that are failed to be automatically classified. Moreover, the labeled noise data by human users makes the on-line noise automatic recognition feasible and is foundation part of the self-evolving classifier.

1.2 Goals of Project

1.2.1 Basic function

The most basic function of this project is measuring the personal/community noise exposure dose.

Through monitoring the personal exposure, a user can know the level of noise that she/he exposed to, and his/her heath experts can make professional recommendation to prevent further health impairment. Also, a distribution map of noise level will be produced based on crowd-sourcing data with the correction of acoustic propagation model. This map can visualize the community exposure dose and help city managers to control noise issues. What's more, recently happening noise events will be presented on the map to provide risk alerts in real-time.

1.2.2 Noise classification

The noise classification algorithm is crucial for the noise monitoring. The components of noise pollution can be delicately analyzed with the results of noise classification, and the major pollution sources can be further determined. In this project, hidden Markov models (HMMs) are used as the classifier. HMMs are probability based models which can be trained separately, which means the re-estimation of one specific HMM will not require any computation performed on other unrelated HMMs.

1.2.3 Human as sensors (Crowd-source label)

Although the noise event classification algorithms can be powerful, users would inevitably be involved in annotating sensing data with associated metadata describing the acoustic sampling. Since the classification tasks could be difficult to do algorithmically, especially, when the mobile end has limited bandwidth and computation resources. What is worse, raw acoustic data is often too noisy, and the size of training set cannot be large enough at initial phase because of the known scalable sound sensing problem [22]. As a result, the reliability of the "best" classification algorithm, if there is, would be doubtable. The user participation and assistance are necessary and can simplify the issue. Furthermore, the HMM based acoustic event classifier could be algorithmically evolved and get better performance with accumulated annotations from users.

1.3 Related works

The Gaussian mixture model based HMM has been intensively studied in the past decades. As a welldeveloped probability based algorithm, whose performance has been proved in the domain of speech recognition, HMM was also used for acoustic environment classification [23][24] with MFCCs[25] features. However, both of these works didn't handle with the large scale issue in acoustic classification.

II. Implementation

2.1 Data collection

2.1.1 Mobile Application Design

A mobile crowd-source application is implemented at Android 6.0 system for this project. The project website (https://map4noise.njit.edu/) has a download link for its beta version. Fig.1 shows the user interface. It is clear the application implement the functions of crowd-source sensing and crowd-source labeling. In addition, the application also has a function of “silent mode” at which phase unlabeled data points can be automatically collected by a foreground service.

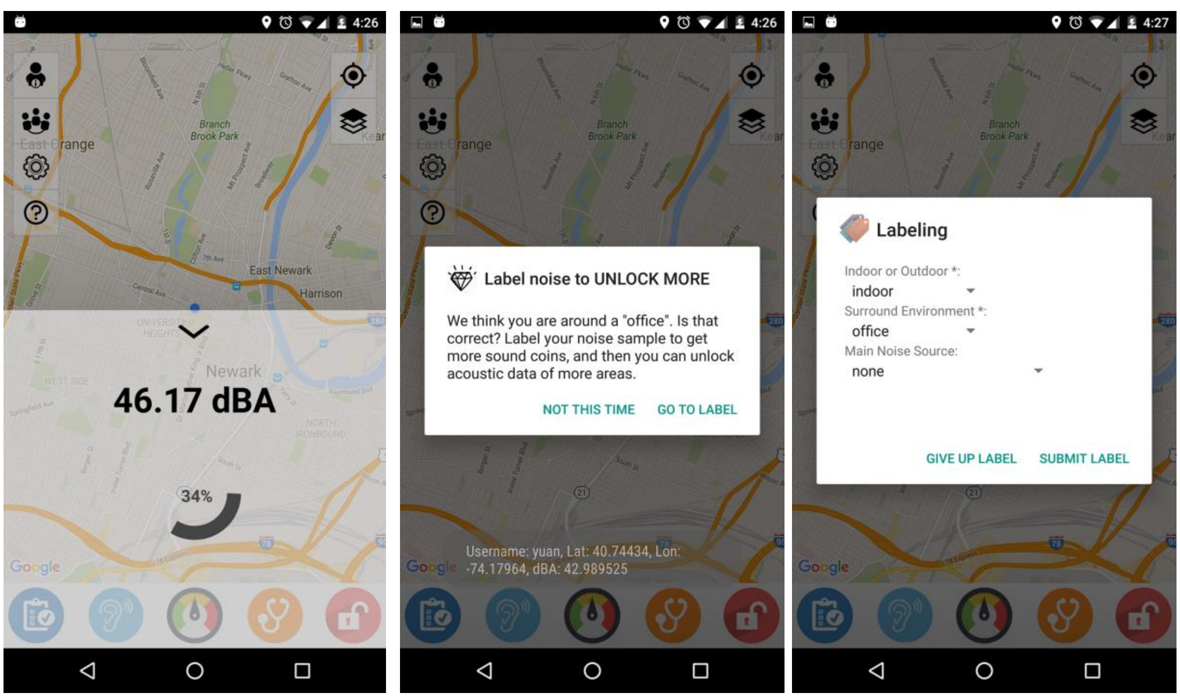

Fig.1. The UI of crowd-sensing and crowd-source labeling functions.

2.1.2 Features Extraction

Mel-frequency cepstral coefficient (MFCC) is used as features with the purpose of dimensionality reduction.

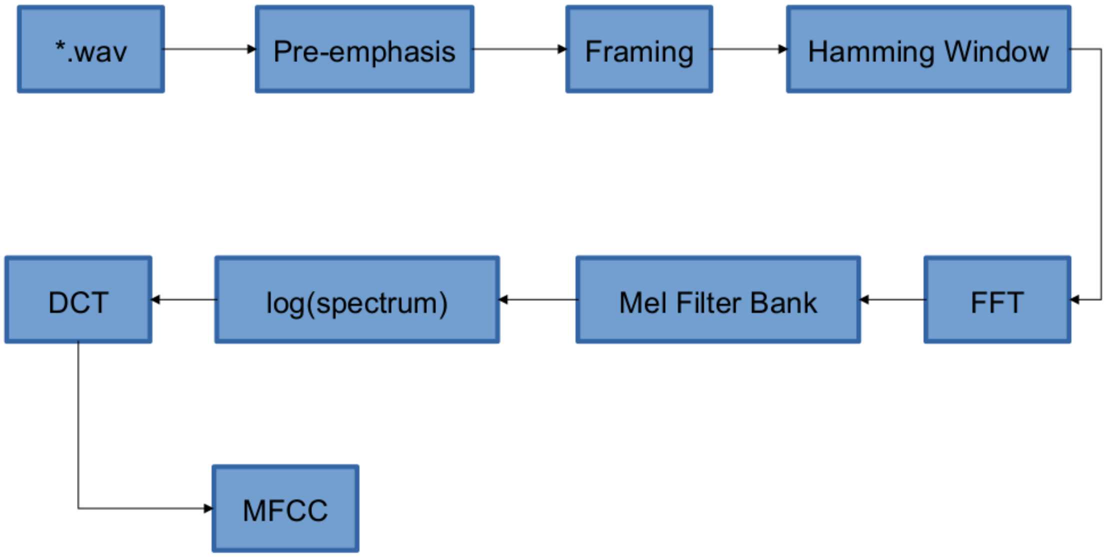

Fig.2. The procedures of feature extraction (MFCC).

As shown in Fig.2, MFCCs are calculated from the original wav files by the standard MFCC procedures with pre-emphasis and Hamming window. In this project, each sample is 10 seconds with a sampling rate of 36kHz. When extracting MFCCs, the length of frame is chosen as 20ms. 39 MFCCs, including 13 static MFCCs, 13 1st-order MFCCs and 13 2nd-order MFCCs, are used to span the feature space.

2.1.3 Server end

The server end is implemented in Java2EE combined with MySQL database.

2.1.4 Parallel implementation of classifier

The library of HTK is used to implement the HMM training and testing functions. HTK is an old but efficient toolbox for GMM-HMM based speeching recognition. This project also takes the advantage of

Perl ForManager library to develop parallel training controller, which greatly reduce the time spent on training. Table.1 shows the dependency of time usage on the maximum number of threads for same training and testing sets. It is shown that the parallel computation can save more than 64% time when allowing 8 processes at an Intel i7 CPU. Even only apply 2 processes in training, about third time can be saved. Table.1 also shows eight threads cannot enhance the performance too much than that of four threads, which follows up the basic limitation of parallel computation: the marginal effect of increasing thread number will lead up to higher overhead and lower efficiency.

Table.1. Performance of PARALLEL computation.

Max Thread#	8	4	2	1
real	1m35.553s	1m55.075s	2m43.949s	4m39.029s
user	6m11.148s	4m58.920s	4m41.812s	4m37.400s
sys	0m0.884s	0m0.744s	0m0.736s	0m0.704s

2.2 Classification

2.2.1 Initial training

As shown in Fig.3, one directional HMM is used to present noise sample. A small training set as shown in Fig.4 is used to initiate HMMs. For every individual HMM, there are 15 hidden states including 1 starting state and 1 ending state. There are 12 labels used in training. These labels and their associating meaning are listed as following:

"drive": user is driving – 29 samples;

"street": off a street – 55 samples;

"highway": off a highway – 20 samples;

"siren": a siren of emergency – 6 samples;

"lecture": someone is speaking slowly – 63 samples;

"dialog" : some persons are talking – 45 samples;

"loudTalk": people are talking loudly, like a bar or restaurant – 9 samples;

"machine": around machine – 17 samples;

"sleep": quite silent bedroom – 25 samples;

"office" - office in GITC at NJIT – 45 samples;

"blowMe" – wind is blowing the microphone – 8 samples;

"walkBag" – mobile phone is in the pocket or bag which is moving – 9 samples.

The last two labels only used for filtering anomaly data in "silent mode", since when a cellphone is placed in a moving pocket or is blown by wind, the recorded sound cannot measure the real noise at all. These harmful data has to be filtered otherwise the measured acoustic distribution cannot represent the threatens of noise pollution.

"Dialog" samples are collected from six different scenarios; "Siren" samples are from two different scenarios; "Lecture" samples are from three; "Highway" are from two different situations; "Sleep" samples are contributed by 3 users; "Office" data is collected from 3 various offices; "LoudTalk" data comes from 2 restaurants and 1 bar. However, "Drive" and "Machine" samples are got from unique data sources.

Fig.3. One directional HMM.

2.2.2 Test over a small set

A small test set is used to verify the performance of the initial HMMs as noise classification. The confusion matrix of testing is shown in Table.2. It is clear that the performance of HMMs as noise classifier is acceptable.

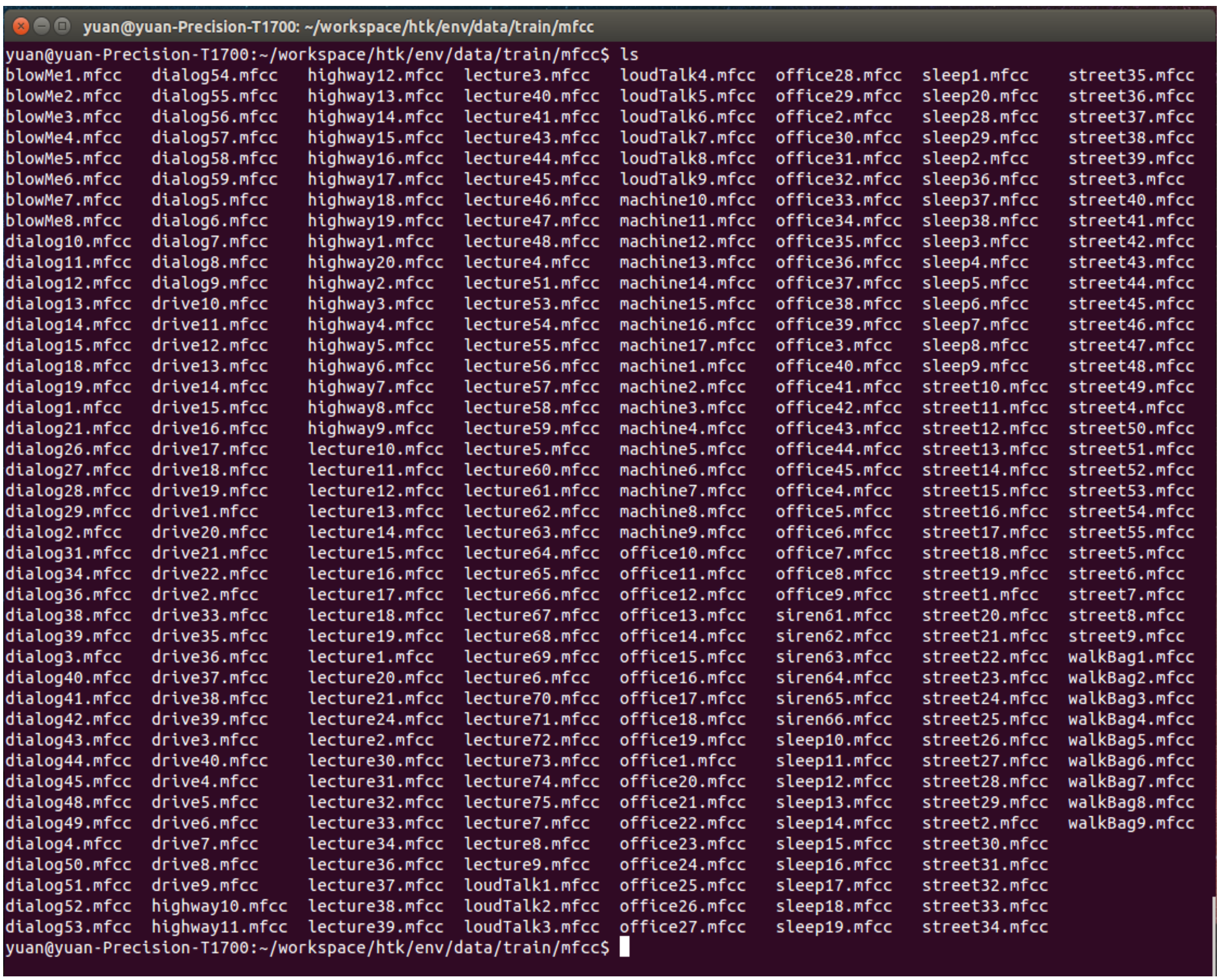

Fig.4. The training set for initially generating HMMs.

Table.2. The confusion matrix of a small testing set.

		Recognized As									
		Drive	Street	Highway	Siren	Lecture	Dialog	LoudTalk	Machine	Sleep	Office
Label	Drive	30	1								
	Street	1	18								
	Highway		1	11							
	Siren				4						
	Lecture					11	1				1
	Dialog					2	15				
	LoudTalk						1	6			
	Machine								5		
	Sleep									15	
	Office						1				16

2.3 Algorithmically self-evolving of HMMs

Since the initial training set is definitely too small when considering the scale issue of environmental acoustic recognition. The recognition error rate will dramatically increase with more scenarios and larger scale of problem. As shown in Fig.5, this project proposes a self-evolving HMMs to handle with this unavoidable problem.

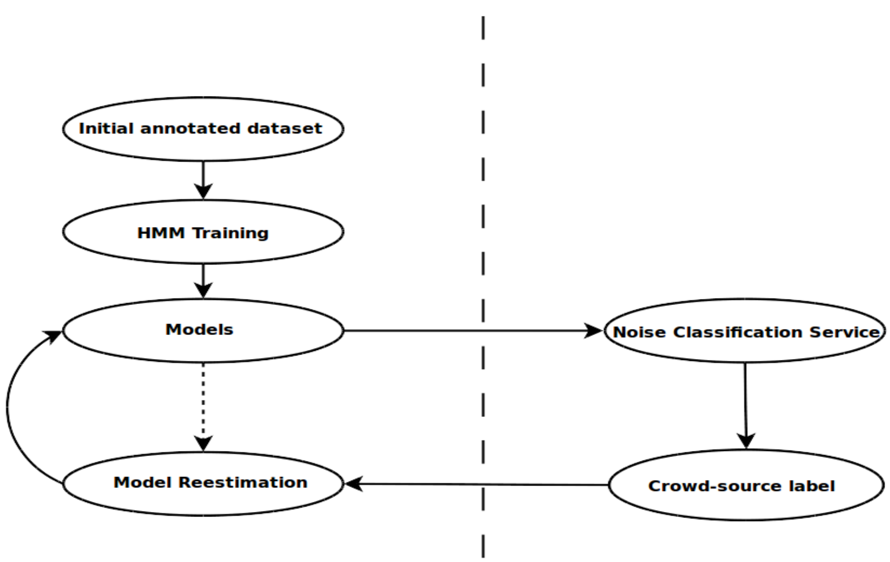

Fig.5. The self-evolving architecture.

2.3.1 Enlargement of training set (Re-estimation case 1)

The original training set of “loudTalk” is collected in Giovanni Pizza & Grill. The HMM of “loudTalk” is trained specifically for this scenario. As shown in Table.3, when new data which is got from Gyu-Kaku, almost half of them are wrongly classified. The strategy of enlarging of training set is suitable for his situation. As shown in Table.4, when put 7 correctly classified new data samples into training set and do the model re-estimation for “loudTalk”, the six out of the seven previously wrongly classified samples are correctly classified.

Table.3. Data from a new scenario results in errors by not enough trained HMMs.

	… ...	Lecture	Dialog	LoudTalk
LoudTalk	… ...	3	4	7
	… ...	… ...	… ...	… ...

Table.4. Put 7 “loudTalk” samples into training set, and the error decreases.

		Lecture	Dialog	LoudTalk
LoudTalk	… ...	0	1	6
	… ...	… ...	… ...	… ...

2.3.2 Increment of model complexity (Re-estimation case 2)

Although the enlargement of training set works for a lot of situations, it could confront with the limitation of model complexity. With restricted model complexity, the samples which are near in feature space cannot be reasonably handled with. Table.5 indicates HMMs with 10 hidden states are not complicated enough to reliably classify “lecture” and “dialog”. Table.6 clarifies HMMs with 15 hidden states can distinguish “lecture” and “dialog”.

Table.5. HMMs with 10 hidden states CANNOT distinguish “lecture” and “dialog”.

		Lecture	Dialog
Lecture	… ...	8	5
Dialog	… ...	4	13
… ...	… ...	… ...	… ...

Table.6. HMMs with 15 hidden states CAN distinguish “lecture” and “dialog”.

		Lecture	Dialog	Office
Lecture	… ...	11	1	1
Dialog	… ...	2	15	… ...
… ...	… ...	… ...	… ...	… ...

2.3.3 Split of model (Re-estimation case 3)

Too complicated probability models will yield the problem of over-fitting problem. As a result, it is not feasible to increase the performance of noise classification by adding too much extra complexity. The strategy of “split of model” can be employed to handle this issue. For example, Table.7 shows 4 “street” samples are incorrectly classified as “dialog”. Then, the wrongly recognized samples are used to train a nominal label named “Street1”. The nominal “Street1” label is manually checked and confirmed as “siren”. Table.9 indicates the new “Siren” model enhances the performance of classier.

Table.7. Four “street” samples are wrongly classified as “Dialog”.

		Street	Dialog
Street	… ...	14	4
… ...	… ...	… ...	… ...

Table.8. Take the advantage of the 4 incorrectly classified “street” samples to train a new model “street1”.

		Street	Street1 (Siren)	Dialog
Street	… ...	12	2	0
… ...	… ...	… ...	… ...	… ...

Table.9. Newly collected “Siren” data is accurately classified.

		Street	Street1 (Siren)	Dialog
Street	… ...	12	0	0
Siren	… ...	… ...	5	… ...

III. Summary

In the United States, few efforts have been made on noise pollution monitoring, even though it was estimated that 104 million Americans are under the risk of hearing loss since their continuous 24-hour average exposure levels are greater than 70 A-weighting decibel (dBA) [26]. Taking the advantage of HMMs, this project implements a crowd-sourcing sensor network based real-time noise classification system, which is a low-cost and promising technical approach to face the challenge of increasingly severe noise issues in the metropolitan areas of the United States.

Reference:

[1] World Health Organization. "Burden of disease from environmental noise-quantification of healthy life years lost in Europe. 2011." WHO Regional Office for Europe, Bonn.

[2] Popescu, Diana, and Iuliana Moholea. "Monitoring the reaction and response of people to urban noise." Archives of Acoustics, 35, 2 (2010): 237-244.

[3] Basner, Mathias, et al. "Auditory and non-auditory effects of noise on health." *The Lancet* 383.9925 (2014): 1325-1332.

[4] Haralabidis AS, et al. Acute effects of night-time noise exposure on blood pressure in populations living near airports. Eur Heart J 29(5):658–664 (2008);

[5] Stansfeld, Stephen A., and Mark P. Matheson. "Noise pollution: non-auditory effects on health." *British medical bulletin* 68.1 (2003): 243-257.

[6] Oishi, Naoki, and Jochen Schacht. "Emerging treatments for noise-induced hearing loss." *Expert opinion on emerging drugs* 16.2 (2011): 235-245.

[7] Babisch W. Exposure to environmental noise: risks for health and the environment. Workshop on “sound level of motor vehicles”, Directorate General for Internal Policies of the European Parliament, Brussels, 2012.

[8] NIOSH [2014a]. NIOSH safety and health topic: noise and hearing loss prevention. http://www.cdc.gov/niosh/topics/noise/

[9] Murphy, Enda, and Eoin A. King. "Strategic environmental noise mapping: Methodological issues concerning the implementation of the EU Environmental Noise Directive and their policy implications." Environment international 36.3 (2010): 290-298.

[10] Chopard, F., et al. "The fundamental contribution of noise measurement networks to the application of European Directive 2002/49/EC." *Proc. INTER-NOISE*. 2007.

[11] S. Santini and A. Vitaletti. Wireless Sensor Networks for Environmental Noise Monitoring. In 6. GI/ITG Workshop on Sensor Networks, Aachen, Germany, 2007.

[12] Lane, Nicholas D., et al. "A survey of mobile phone sensing." *Communications Magazine, IEEE* 48.9 (2010): 140-150.

[13] Santini, Silvia, Benedikt Ostermaier, and Robert Adelmann. "On the use of sensor nodes and mobile phones for the assessment of noise pollution levels in urban environments." *Networked Sensing Systems (INSS), 2009 Sixth International Conference on*. IEEE, 2009.

[14] Ruge, Lukas, Bashar Altakrouri, and Andreas Schrader. "Soundofthecity-continuous noise monitoring for a healthy city." *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013 IEEE International Conference on*. IEEE, 2013.

[15] D'Hondt, Ellie, Matthias Stevens, and An Jacobs. "Participatory noise mapping works! An evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring." *Pervasive and Mobile Computing* 9.5 (2013): 681-694.

[16] Kanjo, Eiman. "NoiseSPY: a real-time mobile phone platform for urban noise monitoring and mapping." *Mobile Networks and Applications* 15.4 (2010): 562-574.

[17] Rana, Rajib Kumar, et al. "Ear-phone: an end-to-end participatory urban noise mapping system." *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks*. ACM, 2010.

[18] Leao, Simone, Kok-Leong Ong, and Adam Krezel. "2Loud?: Community mapping of exposure to traffic noise with mobile phones." *Environmental monitoring and assessment* 186.10 (2014): 6193-6206.

[19] L.Agr, D. Casali and A. Brancotti. http://www.widetag.com/widenoise/

[20] http://citylab.inria.fr/urban-civics/

[21] Zheng, Yu, et al. "Diagnosing New York City's noises with ubiquitous data." *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2014.

[22] Lu, Hong, et al. "SoundSense: scalable sound sensing for people-centric applications on mobile phones." *Proceedings of the 7th international conference on Mobile systems, applications, and services*. ACM, 2009.

[23] Couvreur, Christophe, et al. "Automatic classification of environmental noise events by hidden Markov models." *Applied Acoustics* 54.3 (1998): 187-206.

[24] Ma, Ling, Ben Milner, and Dan Smith. "Acoustic environment classification." *ACM Transactions on Speech and Language Processing (TSLP)* 3.2 (2006): 1-22.

[25] Chazan, Dan, et al. "Speech reconstruction from mel frequency cepstral coefficients and pitch frequency." *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on*. Vol. 3. IEEE, 2000.

[26] Hammer, Monica S., Tracy K. Swinburn, and Richard L. Neitzel. "Environmental noise pollution in the United States: developing an effective public health response." *Environ Health Perspect* 122.02 (2014): 115-119.