Methods for Sensing Urban Noises

Tong Liu[1,2], Yu Zheng[1], Lubin Liu[2], Yanchi Liu[3], Yanmin Zhu[2]
[1]Microsoft Research, Beijing, China
[2]Shanghai Jiao Tong University, Shanghai, China
[3]Information Systems Department, New Jersey Institute of Technology, Newark, NJ, United States
{v-tongli, yuzheng, v-yilwan}@microsoft.com; {sjtu_eric, yzhu}@sjtu.edu.cn; yanchilyc@gmail.com

ABSTRACT
Many people live in cities are suffering from noise pollution, which may impair the physical and mental health. A first step towards the understanding of urban noises is to measure real noise levels. In this technical report, we introduce two categories of methods to measure urban noise levels, discussing their advantages and disadvantages. Finally, we apply one of mobile phone-based methods to an in-the-field study that explores the noise situation in New York City (NYC) [10]. The data collected and the mobile client used in the study are available for download at [2].

INTRODUCTION
The rapid progress of urbanization is leading to serious noise pollution [11], which may cause stresses, sleep losses, high blood pressures and even heart attacks [4]. Tackling noise pollution has been attracted a wide range of attention [5][6][8][9]. For example, European Union planned to build up three-dimensional noise maps of all major cities by 2007 [4]. In addition, since 2011, NYC has opened the platform where people can call 311 to complain whatever makes them uncomfortable, including noises [1].

To build a noise map based on real noise measurements [5][9][10], it is necessary to collect the real-measured noise data of different places. In this paper, we introduce a few measuring methods, discussing their applications to different scenarios that are concerns with money, time and human resources. Using one of the proposed methods, we conducted a study to measure the real noise level of 36 locations in Manhattan. The results of the experiment are used to explore the correlation between real-measured noise levels and people's complaints about urban noises [10].

METHODS
In this section, we introduce two categories of sensing methods to collect real-measured noise levels, according to the measuring devices. One is based on professional sound level meters; the other is using mobile phones.

Professional Device-Based Methods
The first category of methods employs a professional noise-measuring device, named sound level meter [3]. A standard sound level meter, as shown in Figure 1 A), gives a readout of equivalent continuous sound level in decibels (dBs), converted from the voltage signal sensed by its microphone.

**The paper was done when the first author was an intern in Microsoft Research under the supervision of the second author. The mobile app was developed by the third author in Shanghai Jiao Tong University.*

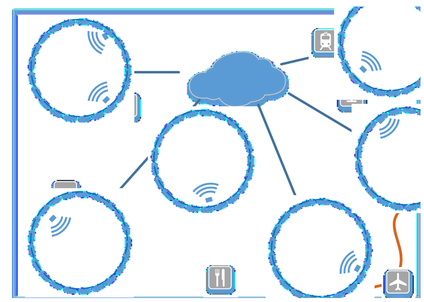


Figure 1. Sound level meter **Figure 2. A network of sound level meters**

A straightforward method is to deploy professional devices permanently in different locations with a certain density e.g., a device per 0.25 square kilometers. The data collected by these devices is then transmitted to a backend system through wired or wireless communication channels, as shown in Figure 2. While this method can produce a long period of high quality noise measurements, it requires a high cost of deployment and maintenance, e.g. power supply, a cover protecting the device from rain, and data transition channels. Additionally, the coverage of this approach is also a concern as we cannot really deploy such sets of devices anywhere.

Without deploying these professional devices permanently, another method is to employ many people, each of which holds a professional device to simultaneously measure the noises in different locations. While this method is more flexible than a fixed deployment, carrying a sound level meter, which is about 30×10 centimeters, is not a fun for a user. Thus, this method is difficult to scale up to thousands of users. The coverage of the locations that can be sensed through this approach is still a problem.

Mobile Phone-Based Methods
To solve the inconvenience of the aforementioned methods, mobile phones are used as substitutes whose microphones are employed to detect the sound power of the surrounding environment. Though the accuracy is slightly lower than a professional sound level meter, the wide availability of mobile phones in end users significantly increases the spatial and temporal coverage of sensing spaces.

1) In an ideal mobile phone-based method, many people hold a mobile phone to measure their ambient noises and then send the noise data back to a centralized place for a further processing. According to some studies [7][9], however, the noise level measured by a mobile phone has a certain deviation from the value measured by a professional sound level meter. Moreover, different mobile phones have

different deviations. Consequently, the data collected by different phones is not directly comparable. To overcome the problem, a calibration is needed for each phone to obtain accurate noise data. This is quite time consuming and reduce the feasibility of a large-scale deployment.

Some calibration algorithm has been proposed to do a self-calibration. For instance, an intervention–free calibration algorithm [7] uses a linear function to model the relationship between phone measurements and real values, as shown in Equation 1:

$$R = \alpha M + \beta + \varepsilon, \qquad (1)$$

where R and M denote real values and phone measurements, respectively; α and β are undetermined coefficients; ε is the random error. α and β can be estimated automatically based on the noise levels collected in a quiet indoor environment.

2) Another method is to use the same mobile phone to measure the noise levels of different locations consecutively. When we only need to evaluate the ranking among different locations in terms of noise level, this method is very lightweight and agile. According to the study in [7], the measurements of a mobile phone and the true noise have a linear relationship. Thus, the measurements of the same mobile phone can reveal the relative ranking between the noise levels of different locations, even if a mobile phone is not calibrated. The aforementioned algorithm can be applied to calibrating a mobile phone for a better measurement.

EXPERIMENT

We applied the second mobile phone-based method in a research project [10] that explores the noise situation in NYC based on the 311 complain data†. In this project, we modeled urban noises as a tensor with three dimensions denoting regions, noise categories and time slots, respectively. We divided a city into disjoint regions by major roads and segmented time of day into 1-hour slots. Each entry of the tensor stores the number of 311 complaints about a particular noise category in a particular region and a particular time slot. As there are not always people reporting the ambient noise anywhere and anytime, the 311 data is very sparse, resulting in many entries in the tensor without values. So, we supplemented the missing entries in the tensor, using the 311 complaint data together with social media, road network data, and Points of Interests. To validate if the inferred values (also called noise indicators) of these missing entries align with the true noise situations in the corresponding locations, we performed an experiment, where a single mobile phone was employed to sense the noise level of 36 locations in Manhattan in different time of day.

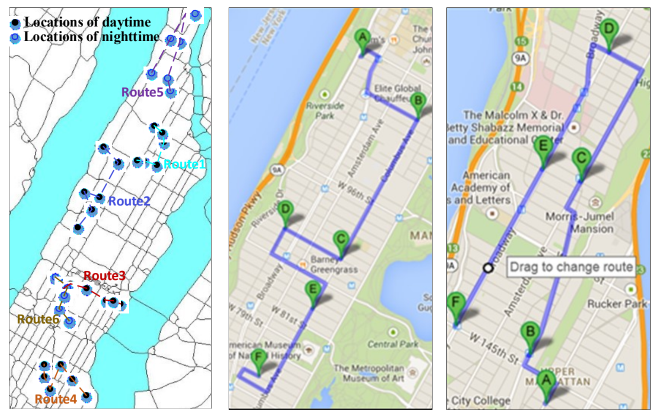
Location Selection and Route Design

Figure 3 A) shows the 36 places in Manhattan, where we measured noise levels in the experiment. Specifically, 24 locations (labeled by black solid circles) were measured during the daytime of a weekday; the other 12 locations (labeled by blue empty circles) were measured during the night of a weekday. We select the locations, considering the following aspects.

†: *311 is NYC's governmental non-emergency services, which allows people in the city to complain everything that is not urgent by making a phone call, or texting, or using a mobile app. [1]*

First, we try to prove that the more 311 complaints are made in a location, the higher the real noise level the location suffers, when the number of 311 is big enough. So, we selected 30 locations, which have enough 311 complaints but with a significant difference between each other's complaints. *Second*, we want to verify that a location without (or having few) 311 complaints may not really be a quiet place. To this end, we selected 6 places having less than three 311 complaints. *Third*, we want to validate that people's tolerance to noise levels changes over time, e.g. people's tolerance to noise is much lower in the night than the daytime. So, we measured the locations at different time of day. *Fourth*, to facilitate the commute of the user who helps us collect the noise data, we designed several routes, each of which is comprised of six locations to measure, e.g., as illustrated in Figure 3 B) and C). Each route is designed to allow a user to travel through six locations in one hour (i.e. a time slot) and measure each location for at least 5 minutes.

The first two aspects motivate us to fill the missing entries of the tensor, and the third one justifies the necessity of modeling the noises of different time slots separately.



A) All locations B) Route2 C) Route5

Figure 3. Locations selected to measure noise

As shown in Figure 3 B) and C), in each route, the user departed from the location marked 'A' and traveled through A→B→C→D→E→F. The time slots of the six routes being measured are listed in Table 1. The coordinates of each location are listed in Table 2.

Table 1. Measuring time slots of six routes

Route1	Route2	Route3	Route4	Route5	Route6
9am ~ 10am	10am ~ 11am	2pm ~ 3pm	3pm ~ 4pm	10pm ~ 11pm	11pm ~ 12pm

Hardware and Settings

In the experiment, a mobile phone (Samsung GALAXY Note I) was employed to measure the ambient sound powers, running an application on the Android operating system. A user carried the same mobile phone to measure the 36 locations by traveling through the six routes shown in Table 1. The interfaces of our application are shown in Figure 4. To convert the voltage sensed by the microphone to a noise level, we employ the *A-weighting* noise level, which is defined in the international standard IEC 61672:2003. The conversing function is defined in Equation 2.

$$L_{AT} = 10log_{10}(\frac{1}{T}\int_0^T \frac{v_A(t)^2}{v_0^2}dt), \qquad (2)$$

where $v_A(t)$ is A-weighting voltage, v_0 is reference voltage, T is the time interval, and L_{AT} is the equivalent noise level over T. In the experiment, T is set as 1 second, which means a noise level value is computed every 1 second. After measuring the noise level in a place for five minutes, all the computed values L_{AT} are stored in a file and sent to a server.

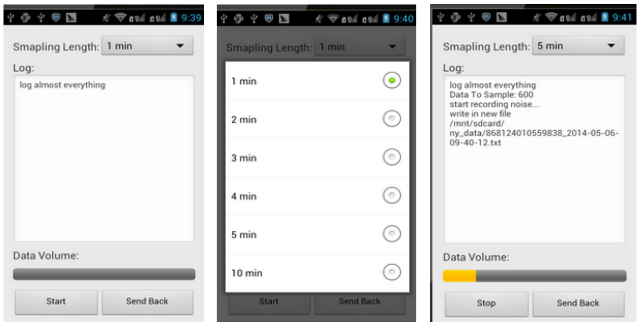


Figure 4. User interface of the mobile client

In the experiment, it is actually not necessary to calibrate the mobile phone measurements, as we only want to verify whether the inferred noise indicators can differentiate between the noise situations of different places. As long as the ranking among different locations in terms of the noise levels can be ensured, the goal of the experiments is achieved.

Table 2. The coordinates of the 36 selected locations

Daytime					
	Route1			Route2	
	latitude	longitude		latitude	longitude
A	40.80936	-73.94927	A	40.80356	-73.96914
B	40.80804	-73.94612	B	40.79923	-73.96286
C	40.80514	-73.94524	C	40.78842	-73.97077
D	40.79815	-73.94832	D	40.78454	-73.97360
E	40.79941	-73.95338	E	40.77924	-73.97904
F	40.79922	-73.95567	F	40.76078	-73.98446
	Route3			Route4	
	latitude	longitude		latitude	longitude
A	40.76233	-73.98234	A	40.73326	-73.98118
B	40.76041	-73.97582	B	40.73763	-73.98611
C	40.75717	-73.96816	C	40.73035	-73.98998
D	40.75686	-73.96545	D	40.73352	-73.98985
E	40.75590	-73.96319	E	40.73453	-73.99225
F	40.72843	-73.97569	F	40.73734	-73.99265

Night					
	Route5			Route6	
	latitude	longitude		latitude	longitude
A	40.82155	-73.94287	A	40.76410	-73.98848
B	40.82470	-73.94427	B	40.76228	-73.98611
C	40.83584	-73.93991	C	40.76141	-73.98406
D	40.84414	-73.93754	D	40.75983	-73.98417
E	40.83659	-73.94314	E	40.75504	-73.98663
F	40.82647	-73.95053	F	40.75096	-73.98272

Noise Level Computation

The mobile client produces a noise level every second, resulting in 300 measurements in 5 minutes for each location. As illustrated in Figure 5, we computed the average of the top 10% big measurements as the real noise level of a place in a given time span, instead of the average of all 300 measurements. A loud sound is more likely to be considered as an annoying noise by people. Figure 5 plots the measurements of a location in Route1, whose average noise level is 62.57 dBs and the top 10% average is 71.94 dBs.

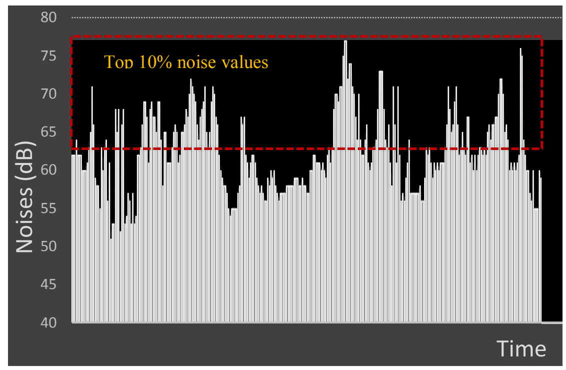
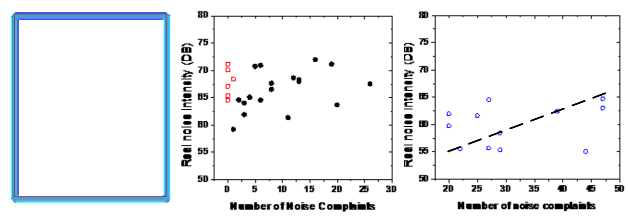


Figure 5. Calculating the noise level for a location

Correlation with Noise Complaints

Figure 6 correlates the real noise levels measured by the mobile phone and the number of 311 complaints around the location. As illustrated in Figure 6 A), the complaints within a circle distance of 200 meters to a location is counted for the location. In Figure 6 B) and C), each point denotes a location, coordinated by its real noise level and 311 complaints.



A) 311 complaints of a location B) Correlation in the daytime C) Correlation in the night

Figure 6. Correlation between 311 complaints and real noise levels

On the one hand, given *the same time span* in a day, the more 311 calls made in a location, the louder real noises in the location. We see the same trend in Figure 6 B) and C). If given sufficient 311 complaints, we can recover the noise situation throughout the city by doing some simple statistics on the complaint data. On the other hand, there are some locations (marked by the red circles shown in Figure 6 B)) have very few 311 complaints but still with considerable real noise. This is caused by the sparsity of 311 complaint data, i.e., having no complaint records does not mean no noise. As shown in Figure 6 C), the real noise level in 6am-6pm is actually higher than 7pm-11pm; however, more complaints were made in the latter time span, as people's tolerance to noises is much lower in the night.

We ranked the 24 locations measured in daytime and the 12 locations measured in the night separately. For each set of locations, two ranks were obtained, one by the inferred noise indicators, the other by the real noise levels collected by the mobile phone. We then used *nDCG* to measure the distance between the two ranks, where the rank by the real measurement collected by the mobile phone is regarded as the ground truth. *nDCG* computes the relative-to-the-ideal performance of information retrieval techniques. The discounted cumulative gain of *G* is computed as follows: (here, b = 3.)

$$CG[i] = \begin{cases} G[1], & if\ i = 1 \\ DCG[i-1] + G[i], & if\ i < b \\ DCG[i-1] + \frac{G[i]}{log_b i}, & if\ i \geq b \end{cases} \quad (3)$$

Given the ideal discounted cumulative gain *DCG'*, then *nDCG* at *i*-th position can be computed as $NDCG[i] = DCG[i]/DCG'[i]$.

Figure 7 shows the ranking performance for the daytime and nights, where nDCG@2 to nDCG@10 are presented. Overall, both ranks achieved a high performance with nDCGs around 0.8. Specifically, the rank of the night has a better performance than the daytime. As more 311 complaints were created in the night, the accuracy of the inference in [10] becomes higher, leading to a closer rank to its ground truth. The results shows that the inferred noise indicator can reveal the true noise situation of a location.

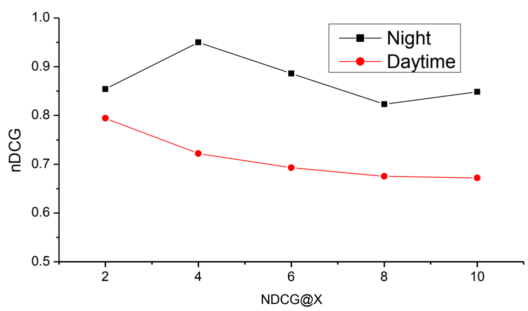


Figure 7. Ranking performance of the inferred noise indicator

Conclusion

In this paper, we report on two categories of noise-measuring methods: one is based on professional sound level meters; the other is to use general mobile phones running a specific application. We discussed the advantages and disadvantages of different methods in different scenarios. We then applied the single mobile phone-based method to a study in NYC. Comparing the real-measured noise levels with the corresponding noise complaints, we verified the correlation between them. The raw noise level data we collected in the experiment and the mobile application can be download from [2].

REFERENCES

1. 311. http://nycopendata.socrata.com/Social-Services/311-Service-Requests-from-2010-to-Present/erm2-nwe9
2. Data release: http://research.microsoft.com/apps/pubs/?id=217236
3. http://en.wikipedia.org/wiki/Sound_level_meter
4. Bulter, D., Noise management: Sound and vision. *Nature*, 5 February 2004, 280-481.
5. D'Hondt, E., and Stevens, M. Participatory noise mapping. Demo Proceedings of the 9th International Conference on *Pervasive 2011*, pp. 33–36.
6. Filipponi, L., Santini, S., and Vitaletti, A. Data collection in wireless sensor networks for noise pollution monitoring. In *Proc. DCOSS 2008*, IEEE Press (2008), 492-497.
7. Liu, L., and Zhu, Y., iCal: Intervention-free calibration for measuring noise with smartphones. Submitted to *Globecom 2014*.
8. Maisonneuve, N., Stevens, M., Niessen, M. E., and Steels, L. NoiseTube: Measuring and mapping noise pollution with mobile phones. In *Proc. ITEE 2009*, Springer Berlin Heidelberg, 215-228.
9. Rana, R. K., Chou, C. T., Kanhere, S. S., Bulusu, N., and Hu, W. Ear-phone: an end-to-end participatory urban noise mapping system. In *Proc. IPSN 2010*, ACM Press (2010), 105-116.
10. Zheng, Y., Liu, T., Wang, Y., Liu, Y., Zhu, Y., and Chang, E., Diagnosing New York City's Noises with Ubiquitous Data. In *Proc. UbiComp 2014*.
11. Zheng, Y., Capra, L., Wolfson, O., Yang, H. Urban Computing: concepts, methodologies, and applications. ACM transactions on Intelligent Systems and Technology, 2014.