A NEW NOISE ANNOYANCE MEASUREMENT METRIC FOR URBAN NOISE SENSING AND EVALUATION

Huan Zhou, Rongshan Yu, Ying Song and Haiyan Shu

Smart Energy and Environment Cluster
Institute for Infocomm Research, A*STAR
Singapore

ABSTRACT

This paper investigates the problem of noise-induced annoyance level evaluation, and proposes a novel annoyance measurement metric for more efficient and accurate evaluation of annoyance level of different types of noises. Results from a large-scale subjective listening test using 90 different noise clips and 96 subjects show that the proposed method can produce more consistent and reliable annoyance ratings than the widely adopted ISO method. Based on the subjective test results, we further develop an objective noise annoyance level measurement model based on the selected psychoacoustic features extracted from the noise samples. Our evaluation results show that the objective model produces satisfactory prediction accuracy on noise annoyance level.

Index Terms— annoyance measurement, multiple regression, subjective listening tests, annoyance modeling

1. INTRODUCTION

Urban noise pollution is a major concern worldwide as a cost of rapid economic growth and development. Among health effects from noises, annoyance is one of the most salient effects on humans [1]. Although there are more and more smart cities are using IoT with large scale WSN (Wireless Sensor Networks) to provide ubiquitous sensing and measuring urban noise environments, most of the existing systems only provide limited information about the environment noises such as averaged sound pressure level, which in many cases only provides a poor measurement on impacts of noise to human. Therefore, there is an emerging need to assess the quality of the urban environment by exploring the relationship between urban noise exposure and annoyance [2, 3].

From methodology point of view, there are two kinds of noise-induced annoyance evaluation approaches. One approach is via social surveys, where questionnaires or interviews are conducted in situ by asking people who have been exposed to the noise environment for a long time to respond their experienced annoyance [4]. The other approach focuses more on short-term effect of noise and the annoyance level information of noise is collected via subjective jury tests where subjects report their perceived annoyances of short-term noise exposures under a controlled laboratory environment (e.g. an anechoic room) [5].

In either annoyance evaluation methods, the degree of noise annoyance is usually quantified using two indices, namely, a 5-point verbal rating scale (not at all/slightly/moderately/very/extremely) [6] and a 11-point numerical rating scale (0-to-10 with even distribution) [4]. However, as pointed out in previous studies [7, 8]), there lacks a systematic mechanism for the listener to associate either verbal rating scale or numerical rating scale with their perceived level of annoyance of the listening test samples, which possibly generates confusion between ratings. Also, the new psychological evidence [9] shows that human prefers to compare candidate with benchmark rather than assess with numerical scores.

In this research, we investigate the subjective jury test based annoyance evaluation approach. A new annoyance measurement metric is used in the jury test and its advantages are demonstrated through results analysis. The new metric introduces a reference stimulus with adjustable sound level as a benchmark, and the noise annoyance is evaluated by comparison between the reference and noise, and rated by numerical adjusted level value. The subjective jury test results show the new metric outperforms the traditional method with more consistent and reliable annoyance ratings.

In practical IoT applications such as urban noise monitoring or environment sensing, it is preferable to have a objective *psychoacoustic annoyance* (PA) model such that the noise monitoring system can compute the annoyance of environment noise directly from captured audio signal. For example, the well-known dose-response curve in [10] relates the overall noise annoyance to noise exposure if the noise exposure information (like day-night average sound level L_{dn}) is available. More sophisticated PA models for noise annoyance level prediction based on psycho-acoustical features of noise stimuli such as loudness, roughness and sharpness were also developed in [5, 11, 12, 13, 14, 15]. In this paper, we follow the idea and proposes a PA model for typical urban noises based on results from the large-scale listening tests using multiple regression and dimension reduction.

The rest of the paper is organized as follows. The new annoyance measurement method is firstly proposed in Section 2. Then corresponding subjective jury tests are detailed in Section 3. PA modeling is then discussed in Section 4. Lastly, Section 5 concludes the paper.

2. NEW NOISE-INDUCED ANNOYANCE METRIC

Although there is considerable diversity of opinions on how subjective annoyance should be measured, existing methods can be typically categorized into either absolute rating scale (verbal or numeric) based method, or discrimination scale (e.g. rank rating) based method. Pros and cons of these existing methods are summarized in Table 1.

To overcome those respective limitations, a novel noise annoyance measurement method was proposed in our previous work [17, 18]. The fundamental idea of the method is to introduce a refer-

Table 1. Comparison of Existing Annoyance Metrics

Annoyance metric	Pros	Cons
absolute rating scale (e.g. [6, 4])	absolute index convenient for analysis and model	confusion between scales and less accuracy
discrimination scale (e.g. [7, 16])	less ambiguous for untrained subjects	no quantitative index and time consuming

ence stimulus with adjustable sound level (in dB) as a benchmark for noise annoyance assessment. During a subjective jury test, the subject firstly experiences short-term noise exposure from a sound pair: the reference stimulus and the target noise. Next, the subject is required to adjust sound level of the reference stimulus to reduce the annoyance difference until the final sound pair brings the same or similar sensation to the subject, or the adjusted sound level reaches its range limits. In the end, the final bipolar value of sound level adjustment on the reference stimulus is recorded as an annoyance index to reflect the annoyance degree of the target noise.

In the proposed method, the difference of annoyance between the reference stimulus and the target noise is not evaluated. Instead, the sound level adjustment is recorded. This overcomes the scales confusion in the absolute rating scale method and no quantitative issue in the discrimination scale method. The proposed method theoretically combines the merits of both two types of existing annoyance measurement methods to ease the listening test process and produces less ambiguous results.

Clearly, the effectiveness of the method will by and large determined by the reference stimulus being used in the jury test. Ideally, the reference stimulus should possess the following properties:

- It should be a reproducible synthetic signal for the method to be practically applicable to different labs
- The reference noise should hold a monotonic increasing relationship between its sound level (in dB) and subjective annoyance response.
- The reference noise should be complex and contains "rich" psychoacoustic features that are commonly recognized as annoyance contributors of noise.

Under the above designing guidelines, we construct a reference stimulus signal as

$$s(n) = \sum_{i=1}^{k} \omega_i \cdot \zeta_i(n)$$

where ω_i denotes weighting, k is the number of synthetic signals, and ζ_i are synthetic signals with very different psychoacoustic characters. For different application scenarios, the number (k) and the types of synthetic signal are varied. For our urban noise evaluation approach, we choose the five unique synthetic signals as given in Table 2 which cover the rich psychoacoustic features for evaluation.

By varying the weighting ω_i, the constructed signal $s(n)$ presents different annoyance level over different sound levels. In our approach, we choose the ω_i set as given in Table 2 to generate the reference stimulus, where its annoyance level versus sound level performance is presented in Fig. 1 (note that herein, the adjust sound of 0 dB corresponds to a sound level of 60dB). Obviously, this generated signal fulfills the requirement of reference stimulus properties as listed above. In the following, this generated signal is used as the reference stimulus for our proposed subjective jury test.

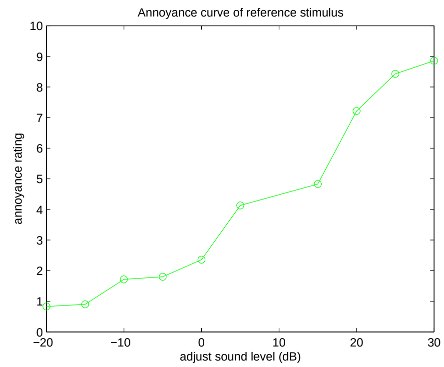

Fig. 1. Annoyance curve of reference stimulus

3. SUBJECTIVE NOISE ANNOYANCE LEVEL EVALUATION EXPERIMENTS

The above proposed annoyance measurement method was used in a large-scale urban noise evaluation experiments for formal assessment of annoyance level of common urban noises. For comparison, the widely adopted ISO 11-point annoyance measurement method is also applied in the evaluation experiments as a benchmark method.

3.1. Noise Recording

We first collected 128 typical urban noises samples at different timing and locations in Singapore for the evaluation experiments. The samples are recorded in mono, with 44.1K Hz sampling rate and 16 bit resolution using Tascam DR-100 recorder and Neuman microphone. These samples are representative of typical Singapore urban noise types, such as highway noise, metro noise, construction noise and communication noise. Most samples have a duration of 3-5 minutes.

3.2. Noise Stimuli preparation

In general subjective jury test, the testing stimuli are typically recommended less than 15 seconds to avoid possible masking or memory errors [19]. Thus, it is necessary to extract short noise stimuli clips from the raw recording samples. To this end, we divided each recording sample into a few 10-second segments based on an event detection algorithm. After that, we calculated different multiple psychoacoustic features of each segment, and selected representative segments based on feature distributions. The resulting short segments are further filtered using subjective evaluation to ensure the content diversity and balance of noise stimuli number with each urban noise type.

Table 2. Synthetic signals and their weighting factors

index	ω_i	signal
1	2	board band noise
2	1	sine tone at 1kHz (tonality)
3	1	sine tone modulated with envelope variation at 70 Hz (roughness)
4	1	sine tone with modulation frequency at 4 Hz (fluctuation strength)
5	2	narrowband noise centered at 5kHz (sharpness)

Finally, total 90 noise clips, with duration of 10 seconds each, are selected and used as noise stimuli for the large-scale subjective jury test.

3.3. Experimental Setting and Procedure

A total of 96 subjects are invited to conduct the subjective jury test. They include 36 males and 60 females, with age of 20 to 60, having a wide range of socioeconomic backgrounds (e.g. education level, residential area). They are paid for their successful participation.

To trade-off the expected quantity and quality of the test results, both noise stimuli and subjects are randomly and equally divided into 3 batches (i.e. 30 noise stimulus per subject and 32 subjects per noise stimuli). It turned out that with this division, the average time consumption of each test session is ideally controlled less than 45 minutes. Other detailed experiment setting and procedure descriptions are similar to [17].

3.4. Experiment Results Analysis

Following the above experiment procedures, two groups of annoyance ratings are obtained from all the 90 noise stimuli. Post-screening is further applied on the raw results so that more reliable results can be obtained by omitting some outliers judgments from a particular subject. In this study, the Box-and-Whisker technique is used and about 15% subjects from the raw results are discarded by this rejection process.

The remaining valid data are normalized to [0, 10] to further analyze their statistical means and confidence intervals (CIs) at 95% CI. For clarity, we name the rating results using the proposed method as Group 1, and results using the ISO 11-point annoyance measurement method Group 2. The analysis of their means shows that the overall mean annoyance ratings from Group 1 is similar that from Group 2. However, almost all CIs of Group 1 are distinctively narrower than corresponding that of Group 2. In Fig. 2, the results of 30 noise stimuli from one batch are illustrated for clarity. The average CI of all 90 noise stimuli is 0.942 for Group 1, and 1.329 for Group 2. More than 30% reduction of CI shows that the proposed method can provide less ambiguous results, that is, more consistent annoyance rating with less variation.

Furthermore, the above raw rating results are re-sampled to a subset with randomly sampled results with less number of subjects. Such a subset is further statistically re-analyzed. To reduce the sampling bias, such a procedure is repeated with 10 iterations. The final averaged CIs are compared in Table 3. It shows that, with the same number of subjects, the proposed method presents narrow average CI than that of the ISO method. And the proposed method can use 12 subjects to achieve similar average CI whereas ISO method may need 20 subjects. Less subjects are needed with the proposed method for a subjective jury test to produce comparable annoyance rating results which means it is easier to obtain statistically significant listening test results by using the proposed method.

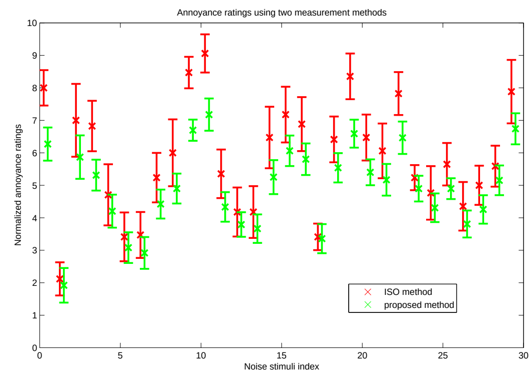

Fig. 2. Comparison of annoyance level evaluation between the proposed method and ISO method.

Table 3. Comparison of CIs with different number of subjects

# subject / method	20	16	12	8
Proposed	1.184	1.433	1.686	2.460
ISO method	1.700	2.042	2.344	5.396

4. OBJECTIVE NOISE ANNOYANCE LEVEL MODELING

In urban noise monitoring, most existing systems only report sound pressure level of the recorded noise, which are insufficient since it is well-known that sound pressure level does not provide a full picture of the human perceptual quality of a sound. Thus, it would be good for an objective predictive model to deploy not only sound pressure level information but also other sound psychoacoustic features that incur annoyance to human perception. And this model can be integrated into noise monitoring systems to provide more actionable information on the environment under monitoring. To this end, we develop a multiple regression based model to predict the annoyance level based on psychoacoustic features extracted from noise samples based on the data from our subjective jury test.

4.1. Psychoacoustic Feature Extraction

In our model, we use five psychoacoustic features including loudness, RMS (root-mean-square) energy, roughness, brightness and tonality. All of them are frequently used as sound quality metrics [20]. Except for loudness, each of the other four features is extended to 7 parameters including its mean, standard deviation, maximum, minimum, 1st quartile, 3rd quartile and median over the duration of the testing sequence in order to capture both the mean and statistical dynamics of the given psychoacoustic features of the testing sequence. For loudness, we calculate Zwicker's Loudness [21] that is standardized and generally used for time-varying sound signals. Three parameters, N_5 (used for environmental sounds), maximum(N_{max}) and mean(N_{mean}) are extracted [14, 22, 11]. Therefore, the psychoacoustic features used in our model have $3 + 4 \times 7 = 31$ dimensions in total.

Table 4. Predictive feature set

Psychoacoustic features	Dimension
loudness	3
RMS energy	7
Roughness	7
Brightness	7
Tonality	7

4.2. Regression Modeling

To deal with the potential multicollinearity problem (i.e., with very high variance inflation factors) of the selected prediction parameters, we use several dimension reduction methods including stepwise selection [12, 23, 24], PCA (Principle Component Analysis) [25, 26], LASSO (Least Absolute Shrinkage and Selection Operation, penalizing the sum of absolute regression coefficients) [27] and PLS (Partial Least Square, fitted with the orthogonal scores algorithm) [28]. The regression model is then trained by splitting the dataset randomly into a training dataset (with 72 objects) and a test dataset (with 18 objects). The goodness-of-fit of regression model is evaluated using 4 performance tests: *df* (degree of freedom), *mean square error* (MSE), Pearson's product-moment correlation and Spearman's rank correlation.

The obtained prediction performances are listed in Table 5. It can be seen that PLS based model achieves the best performance in terms of fitting the annoyance rating to the subjective test results. The performance of LASSO based model is slight worse than that of the PLS based model. The prediction results by PLS on the test dataset (with 18 objects) are visualized in Fig. 3. The overall prediction results are reasonably good which shows the possibility for evaluating annoyance ratings of urban noises by using only objective parameters at low complexity, with the primary advantage of avoiding time-consuming subjective jury listening test.

Table 5. Comparison of prediction performances on test set

Method	df	MSE	Pearson	Spearman
stepwise	16	7.312	0.930	0.932
PCA	3	16.428	0.860	0.837
LASSO	7	3.857	0.964	0.934
PLS	**3**	**2.886**	**0.967**	**0.953**

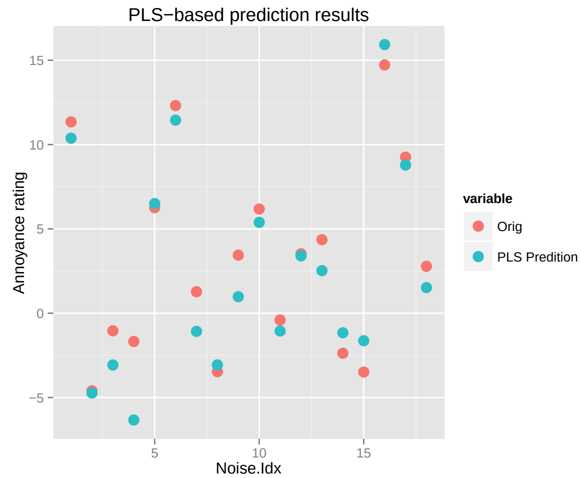

Fig. 3. Prediction results using PLS regression method.

5. CONCLUSIONS

In this paper, we describe a new annoyance measurement method based on pair-wise comparison and artificial generated reference noise signal. The effectiveness of the proposed method is validated using a large scale subjective test with 90 testing stimuli extracted from a database of common urban noise recorded in Singapore, and 96 subjects. To apply the listening test results in practical applications such as environment sensing or urban noise control, we further develop a objective noise annoyance level model based on multiple regression with selected psychoacoustic features extracted from the audio signal and dimension reduction. Evaluation results shows that the developed objective model achieves satisfactory prediction performance with regarding to subjective annoyance level scores.

Acknowledgement

This material is based on research/work supported by the Singapore Ministry of National Development and National Research Foundation under L2 NIC Award No. L2NICCFP1-2013-7.

6. REFERENCES

[1] Lin Fritschi, A. Lex Brown, and etc., "Burden of disease from enviromental noise: Quantification of healthy life years lost in europe," 2011.

[2] Karina Mary de Paiva Vianna, Maria Regina Alves Cardoso, and Rui Manuel Calejo Rodrigues, "Noise pollution and annoyance: an urban soundscapes study," *Noise & Health*, vol. 17, pp. 125–133, 2015.

[3] Lei Yu, Jian Kang, Liang Hong, and Charles Xie, "Soundscape identification in noise annoyance evaluation," *Proceedings of Inter-noise 2014*, 2014.

[4] "Acoustics - assessment of noise annoyance by means of social and socio-acoustic surveys," *ISO/TS 15666*, 2003.

[5] F.C. Marquis, Eric Premat, and D. Aubree, "Noise and its effects - a review on qualitative aspects of sound. part ii: Noise and annoyance," *Acta Acustica united with Acustica*, vol. 91, pp. 626–642, Jul 2005.

[6] J.M. Fields, R.G. De Jong, and T. Gjestland etc., "Standardized noise reaction questions for community noise surveys: research and a recommendation," *Journal of Sound & Vibration*, vol. 242, no. 4, pp. 641–679, 2001.

[7] M.S. Khan, O. Johansson, and U. Sunback, "Subjective annoyance response to diesel engine sound during idling conditions," *Int. Journal of Occupational Safety and Ergonomics*, vol. 2, no. 1, pp. 16–26, 1996.

[8] Daniel L. Steele and Song Hui Chon, "A perceptual study of sound annoyance," *AudioMostly*, pp. 19–24, Sep 2007.

[9] Mark H. Waugh, "Mentors, muses, and memories: Personal narratives from psychological assessment," *Journal of Constructivist Psychology*, vol. 29, pp. 80–99, 2015.

[10] T.J. Schultz, "Synthesis of social surveys on noise annoyance," *J. Acoust Soc Am.*, vol. 64, no. 2, pp. 377–405, Aug 1978.

[11] E. Zwicker and H. Fastl, "Psychoacoustics: Facts and models," *Springer*, 2007.

[12] Eleonora Carletti, Francesca Pedrielli, and Camilla Casazza, "Annoyance prediction model for assessing the acoustic comfort at the operator station of compact loaders," *The 16th International Congress on Sound and Vibration*, 2009.

[13] J. Morel, C. Marquis-Favre, and L.-A Gille, "Relationship between road and railway noise annoyance and overall indoor sound exposure," *Applied acoustics*, vol. 101, pp. 47–57, 2016.

[14] A. Klein, C. Marquis-Favre, R. Weber, and A. Trolle, "Spectral and modulation indices for annoyance-relevant features of urban road single-vehicle pass-by noises," *J. Acoust. Soc. Am.*, vol. 137, no. 3, pp. 1238–1250, March 2015.

[15] Hussein M Elmehdi, "Using mathematical models to predict annoyance from combined noise sources in the city of dubai," *Inter-noise 2014*, Nov. 2014.

[16] Antonio J. Torija, Ian H. Flindell, and Rod Self, "frequency weightings based on subjectively dominant spectral ranges," *EuroNoise*, pp. 1645–1649, June 2015.

[17] Huan Zhou, Rongshan Yu, and Ying Song, "A novel noise-induced annoyance measurement method," *TENCON 2016*, Nov. 2016.

[18] Ying Song, Rongshan Yu, Huan Zhou, and Haiyan Shu, "Annoyance measurement of singapore urban environmental noise," *TENCON 2016*, Nov. 2016.

[19] "General methods for the subjective assessment of sound quality," *ITU-R BS.1284-1*, 2003.

[20] Fastl H., "Psychoacoustics and sound quality metrics," *Sound Quality Symposium*, 1998.

[21] Eberhard Zwicker and Hugo Fastl etc., "Program for calculating loudness according to DIN 45631 (ISO 532B)," *J. Acoust. Soc. Jpn*, vol. 12, no. 1, pp. 39–42, 1991.

[22] J.M. Barrigón Morillas etc., "Sound quality in urban environments and its relationship with some acoustics parameters," *Acústica*, 2008.

[23] Eleonora Carletti, Francesca Pedrielli, and Camilla Casazza, "Development and validation of a numerical prediction model to estimate the annoyance condition at the operation station of compact loaders," *International Journal of Occupational Safety and Ergonimics*, vol. 17, no. 3, pp. 233–240, 2011.

[24] Michael Bauer and Dominique Collin etc., "Cosma- a european approach on aircraft noise annoyance research," *Inter-noise 2014*, 2014.

[25] Francesca Pedrielli and Eleonora Carletti, "Application of psychoacoustics to the noise signals generated by gear pumps," *The 22nd International Congress on Sound and Vibration*, July 2015.

[26] Antonio J. Torija and Diego P. Ruiz etc., "Relationship between road and railway noise annoyance and overall indoor sound exposure," *Transportation research part D: transport and environment*, vol. 16, no. 1, pp. 15–22, Jan. 2011.

[27] Robert Tibshirani, "Sound quality in urban environments and its relationship with some acoustics parameters," *Journal of the royal statistical society, Series B (methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[28] Johan Trygg and Svante Wold, "Orthogonal projections to latent structures (O-PLS)," *Journal of chemometrics*, vol. 16, no. 3, pp. 119–128, 2002.