Pak. J. Engg. Appl. Sci. Vol. 26 January, 2020 (*p.* 102–108)

Automatic Mosquito Surveillance and Visualisation using Acoustic Signals

Ayesha Hakim

Department of Computer Science MNS-University of Agriculture, Multan
* **Corresponding Author:** Email: ayesha.hakim@mnsuam.edu.pk

Abstract

Mosquitoes are considered to be among the biggest disease spreading flying insects causing worldwide health hazards. In Pakistan, Malaria and Dengue fever are among the most dangerous infectious viral diseases transmitted through the bite of infected Anopheles and Aedes mosquitoes. Manual identification of these mosquitoes is hard and dangerous leading to severe health risks. This paper presents an automated system based on audio analysis and IoT sensors to identify and classify mosquito species including Aedes (transmits dengue), Anopheles (transmits Malaria), and Culex (transmits viral diseases including avian malaria) species using the acoustic recordings of their wingbeat frequency. Computational analysis of these audio signals leads to inter-species classification and behavioral study of different mosquito species under varying environmental conditions. We demonstrate the proposed approach on a standard dataset and compare it to human annotations, with promising results.

Key Words: aedes, anopheles, culex, acoustic signals, wingbeat frequency, visualisation, machine learning, d3js

1. Introduction

Mosquitoes are considered to be among the biggest disease spreading flying insects causing worldwide health hazards. They are dangerous due to their capability to transmit viruses and parasites causing devastating diseases. According to a recent research survey [1], mosquitoes kill at least 725,000 people per year around the world. More specifically, dengue, transmitted through *Aedes* mosquito, kills 50 to 100 million people per year; malaria, transmitted through *Anopheles* mosquito, kills 400,000 people; while *culex* mosquito (aka common house mosquito [C.pipens]), is responsible for transmitting viral diseases including, West Nile encephalitis, Japanese encephalitis, Usutu virus, and St. Louis encephalitis), causing more than 10,000 deaths per year worldwide. In Pakistan, Malaria and Dengue fever are among the most dangerous infectious viral diseases transmitted through the bite of infected *Anopheles* and *Aedes* mosquitoes.

Insects produce sounds to communicate, finding mates, or isolate species [2]. Theses sounds are produced by various mechanisms including wing-beating, chewing, or flying. This study demonstrates the use of acoustic signals, caused by wing-beat frequency, to identify and classify three of the most important disease-transmitting mosquito's species in Pakistan: *Aedes*, *Anopheles, and Culex*. The research reveals that there is a difference in the acoustic signals produced by male and female mosquito species. This will enable the proposed system to differentiate between male and female mosquito sounds that will not only lead to early detection of disease vectors but also open up new opportunities of research on inter-species classification, as well as behavioural study of different mosquito species under varying environmental conditions and changing climate.

In this paper, we present a smart mosquito surveillance device comprising of mobile phone as acoustic sensor to record mosquito wingbeat sounds and analyse the relationship between their population and behaviour in varying climate conditions using IoT sensors to measure temperature, humidity, and air quality. For this, we developed an unsupervised machine learning API that takes mosquitoes' wing-beats audio files as input, normalize, and transform the data into lower dimension with high variability and calculates the pair-wise similarity matrix depicting the relationship between mosquitoes' wing-beat sounds. The similarity matrix was used into the web services API to generate the online interactive graphical visualisations in order to observe the relationship between different mosquito sounds that revealed interesting results. We observed between-species and within-species sound variation produced by different mosquitoes' species.

Timely identification and classification of disease-carrying mosquitoes is a vital step towards taking actions to eradicate them. This system might be used as a step towards developing remote, non-destructive, and smart automated acoustic traps for eliminating disease transmitting mosquitoes. It might be used for continuous long-term unattended monitoring and automated logging for estimating population density of mosquitoes at a specific location. In case of mosquito detection, an alert may be generated to inform relevant authorities to take further action before the outbreak of any disease.

2. Relevant Research

One of the most crucial input for appropriate and timely control of mosquito-borne diseases is the direct monitoring of mosquito populations in the field settings. Manual monitoring of mosquito populations is cost-intensive, time consuming, requires extensive labor, and needs an expert to accurately identify the mosquito species that can raise safety issues for humans. In the literature, there are frequent attempts of automatic surveillance of crop insect-pests, but very less work has been done on that of mosquito species. Most of the existing work on automated mosquito surveillance used visual stimuli, and a very few used acoustic signals as input.

Fuchida et al. classified mosquito species based on vision-based stimuli by extracting their morphological features such as colour, length of body and legs. They used SVM to recognise mosquito species among other bugs such as flies [3]. Favret and Sieracki, used sparse processing technique and support vector machine for classifying large number of closely related mosquito species, on a dataset containing images of 72 species of fruit flies and 76 species of mosquitoes [4]. Similarly, Wang et al. used Artificial Neural Networks and SVM on mosquito images and achieved 93% accuracy on the dataset of very clear images that might not be possible to capture in field settings [5]. Rose et al. used machine learning techniques to classify male and female mosquitos' species, and they pointed out that the acoustic recordings nearby mosquito traps (such as BG-Counter [6]) are not long enough to be used for automated surveillance [7]. BG-Counter is an insect trap that measures population density of insects but could not differentiate between insect species.

As compared to visual stimuli, acoustic tracking has the potential to be used as a relatively cheap method to monitor behavioural activity of mosquitoes. Sound sensors are generally cheaper than image sensors and there is no additional lighting equipment required. Also, the amount of incoming data is smaller compared to image tracking [8]. Differences in wingbeat frequencies can be used to classify individual mosquitoes to species complex level; however, overlap in frequencies does occur [9]. Mukundarajan et al. in [10] used commercially available mobile phones as powerful sensors to acquire acoustic data of mosquito wingbeat sounds. They attached sensitive microphones with mobile phones and adjust signal-to-noise ratio depending on the noise level at a public place. Similarly, Rama et al. detected insect fight sounds in the field and found that by combining information from the first four harmonics improved the detection rate and reducing the false positives [11]. Ouyang et al., used a recording device consisting of a set of infrared emitters and receivers to count wing-beat of mosquitoes that enter the device. They used expectation-maximization algorithm (EM-GMM), and compared it with ANN and the nearest neighbor model, on living male and female *Aedes albopictus*, *Aedes aegypti* and *Culex quinquefasciatus* [12].

A few papers studied the effect of climate change on mosquito population density using manual observations in laboratory settings, such as [13,14,15]. The evaluations were done in the lab settings at varying temperatures that might be different from field settings.

To the best of our knowledge, none of the existing solutions records location and environmental conditions to study the relationship between mosquito behavior and population density with environmental and climate conditions in the region. Also, there has been no attempt to automatically create and update real time infographics to monitor and track the presence of disease vector mosquito anywhere in the country. This would be useful tool for early prediction of a potential outbreak of mosquito-borne diseases including dengue and malaria.

For an experimental evaluation, we identified a lack of comprehensive dataset of mosquito sounds specially in relation with varying environmental conditions to be used as a standard for automated systems. The audio recordings in most of the existing dataset are too short to be used for mosquito acoustic tracking and it is hard to compare the results that has been tested on different datasets most of those are recorded in laboratory settings with controlled background noise.

3. Datasets

For testing and evaluation of our proposed methods, we have used a dataset containing mosquito wing-beat frequency sounds publicly available at Dryad Digital Repository [10]. This dataset contains wing-beat frequency sounds of medically important mosquitoes including *Aedes aegypti, Aedes albopictus, Aedes mediovittatus, Aedes sierrensis, Anopheles albimanus, Anopheles arabiensis, Anopheles atroparvus, Anopheles dirus, Anopheles farauti, Anopheles freeborni, Anopheles gambiae, Anopheles merus, Anopheles minimus, Anopheles quadriannulatus, Anopheles quadrimaculatus, Anopheles stephensi, Culex pipiens, Culex quinquefasciatus, Culex tarsalis, Culiseta incidens.* The live mosquito specimens were collected from the field and recorded in different labs and facilities to create a curated database of species. Later, field acoustic data was collected from different locations, including homes and gardens, using mobile phone recorders during free flight or captured in a ziplog bag. The field data was annotated by comparing with lab recordings of the sound of specific species. The wingbeat frequency data from each species was isolated for each 20 ms sample window using a peak finding routine on the spectrogram. For complete description of the dataset, please refer to [10].

4. Methodology

The whole process of mosquito species identification is divided into two sub processes. The first subprocess is focused on data cleaning, normalisation, feature extraction, and dimensionality reduction. In the second process, the web services used for visualisation have been developed. Both processes are combined as a standalone aggregated service that is able to cluster mosquito species and visualise relationship within and between-species mosquito audio recordings.

4.1 Feature Extraction

We extracted the short-term and mid-term features of each audio recording using the approach similar to [10]. First, each audio signal is split into time windows (frames). In the literature, the most widely accepted short-term frame (time window) size is 20 to 100msecs. We used a frame size of 50msec and a frame step of 25msec using overlapping framing. For each frame, 34 short-term features (same set of features as used in [10]) are extracted. As a result, each frame is represented as a feature vector of 34 elements each.

Once the short-term features are extracted, the mid-term features are computed by calculating the two statistics of each short-term feature. The following statistics are computed: (a) the average value (μ), (b) the standard deviation (σ^2). As a result, each frame is represented as a 68 dimensional feature vectors, where the first half of the values (in each frame) corresponds to the average value, while the second half to the standard deviation of the respective short-term feature. A long-term average is calculated with respect to all frames resulting in a one feature vector for each audio recording. Each of these long-term averages of mid-term feature vectors is fed into the dimensionality reduction technique to extract the most varying features out of the whole audio signal.

4.2 Dimensionality Reduction

A 68-dimensional feature vector has been computed for each audio recording. The number of feature vectors depend on the number of audio recordings. Each feature vector is normalised to 0-mean and 1-standard deviation. To improve the efficiency of the algorithms, the first step is to find a simplified representation of high-dimensional data in order to visualise and understand the relationships among multiple variables. Generally, in a multivariate dataset there is more than one variable measuring the same kind of behaviour. The problem may be simplified by replacing such redundant groups of variables by a single new variable.

A standard technique to model data variation and analyse sets of datapoints in high dimensional spaces is Principal Component Analysis (PCA) [16]. PCA finds a new set of variables, called principal components (PCs), by identifying a linear transformation (translation, rotation, and scaling) of the original variables in the dataset. All principal components are mutually orthogonal, such that ideally there is no redundant information. In this case, no redundancy means that the principal components are uncorrelated with each other. Each component accounts for a maximal amount of variance in the observed variables that was not accounted for by the preceding components and is therefore uncorrelated with all of the preceding components. The principal components are statistically independent to each other only for normal (Gaussian) random variables [16]. As a whole, the set of principal components form an orthogonal basis for the space of the original dataset. The resultant basis has maximum variance of the dataset along the first basis vector, and successively less variance amongst the following basis vectors. A scree graph was generated to select the suitable number of principal components in order to transform the data into low dimensional space while preserving the variation in the data.

4.3 Calculating Similarity Matrix

Cosine similarity has been used as a measure of similarity between two non-zero vectors of an inner product space normalised by the product of their magnitudes that measures the cosine of the angle between them [17]. For any pair of real-valued vectors x and y, t is calculated as,

$$\mathrm{SM}(x, y) = 1 - \frac{x.y}{\parallel x \parallel \parallel y \parallel}$$

In the past, cosine similarity has been used successfully for speaker clustering and verification [18-22]. Unlike Euclidean distance, cosine distance regards only to the "shape" of the pattern but not to its magnitude and gives a fair measure to the frames with relatively low power [23, 24]. We computed the pair wise cosine similarity between the transformed feature vectors to get a square-form similarity matrix for each audio recording. This similarity matrix was used in web services for online visualisation.

5. Online Interactive Visualisation

Extracting meaningful visualisation based on the relationships between data variables is useful, especially in large datasets. After transforming the audio signals to a lower-dimensional space, a similarity matrix is calculated based on the pairwise cosine distances of feature vectors in the training set. This similarity matrix was converted into JavaScript Object Notation (json) format [25] to be used in web services for online visualisation. Json is a light-weight, text-based data interchange format that makes the similarity matrix language-independent. Based on the similarity matrix, an interactive chord diagram is generated using the powerful data-driven-document (D3) approach to visualise similarity among birdsongs in the browsers. D3 provides efficient scene transformation thus providing flexible animation, interaction, complex, and dynamic visualisations for the web [26].

A chord diagram is a graphical method of displaying the inter-relationships between entities based on a matrix of size $n * n$, where the matrix m represents a directed flow amongst a network of n nodes. Each element of the matrix, $m[i][j]$, represents the flow of the i^{th} node to the j^{th} node. $m[i][j]$ must be nonnegative, though it can be zero if there is no flow from node i to node j. In our case, $m[i][j]$represents similarity of birdsong content in the audio recording i to the audio recording j.

The matrix is passed to *d3.chord*, that returns an array of chords. Each element of chord array represents bidirectional flow between two nodes i and j, and returns zero if there is no flow. Each cord is an object with two sub objects: source and the target. The source and target has the following properties: startAngle - the start angle in radians, endAngle - the end angle in radians, value - the flow value $m[i][j]$, , index - the node index i , and subindex - the node index j. The chords are then passed to *d3.ribbon* to display the network relationships using coloured ribbons. The returned array includes only the unique chords and the chord objects for which the value $m[i][j]$ is non-zero [27].

6. Experimental Results

We implemented our methods in Python with PyCharm IDE using the pyAudioAnalysis toolbox [28]. For visualization of mosquito species clustering based on their sounds similarities, d3js chord diagram representation has been used. Each node in the graph represents a single audio recording. Fig. 1 presents an interactive chordial graph, drawn on the basis of similarity measures of mosquito wingbeat frequency (Aegypti, albopictus, mediovittatus, anopheles atroparvus, anopheles merus), presenting identification of mosquito species. The sound visualization uses the open source AudioContext APIs to play an audio file, and AnalyzerNode to retrieve the frequency values [29]. The colours of the ribbons are based on the file names in the dataset. Similar file names generate similar colours. In the dataset, the name of file is based on the contents of the audio recording, for instance, aegypti.wav contains the audio recording of a aegypti wingbeat sounds. Each ribbon represents a relationship between the audio content at both ends. By aurally examining the overlapping ribbons between two different mosquito species, we found the reasons such as, some matching notes in the syllable, silence, or some kind of noise in both audio recordings.

Fig. 2 shows Zero crossing rate and energy of a sample audio recording of (a) Aedes aegypti, (b) Anopheles astroparvus, and (c) Culex pipiens. These results demonstrates that the mosquito wingbeat signals are weak and varies for each species. The difference between the wingbeat frequency of two mosquito species that has been used for surveillance purposes.

We compared the results of clustering and visualisation with the ground truth data by aurally examining each audio recording. In most occurrences, the similarity in the mosquito sounds from two different species is caused by the environmental noise. Since, the cosine similarity is computed on the mean-centered feature vectors, it is reduced to the measure of Pearson correlation coefficient ρ_{AB} [17]. Therefore, in our experiments, cosine similarly performs similarly to Pearson correlation and braycurtis distance measures and perform better than sqEuclidean and Mahalanobis distance for birdsong clustering.

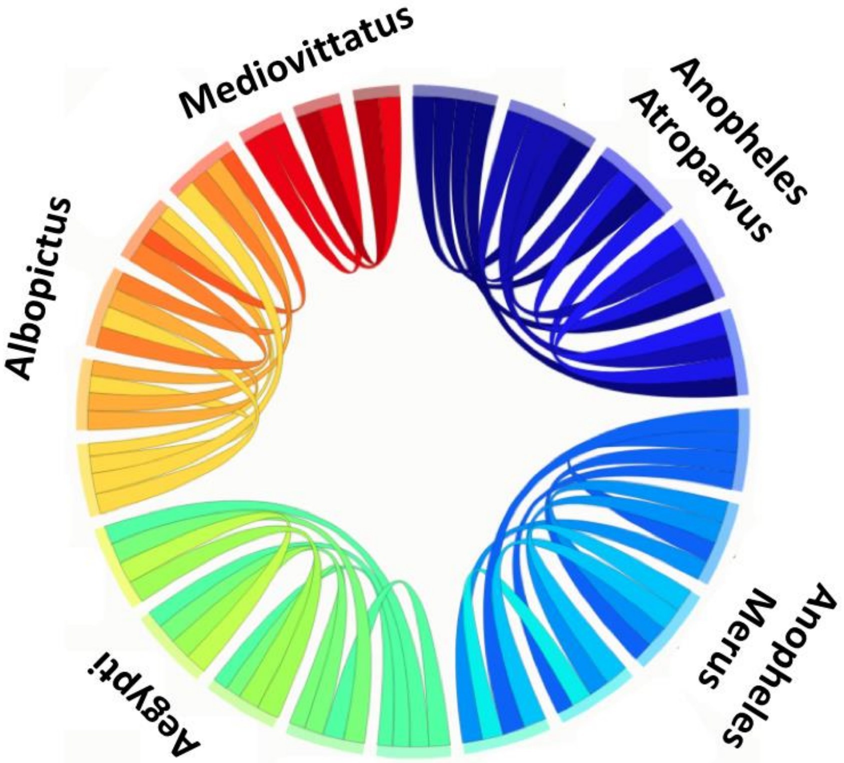
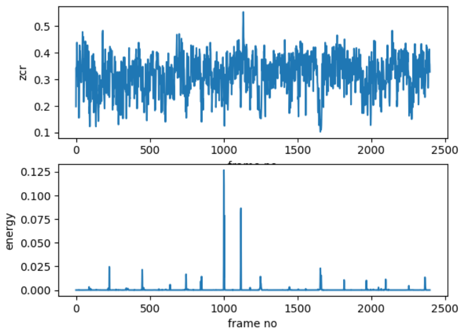
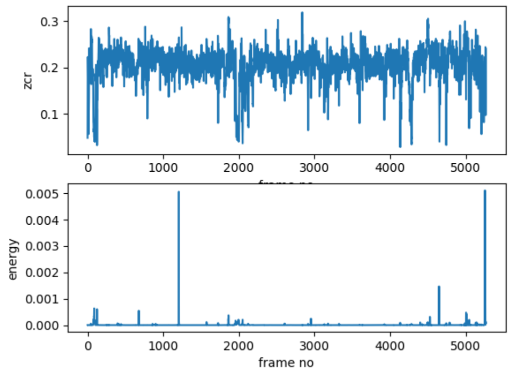


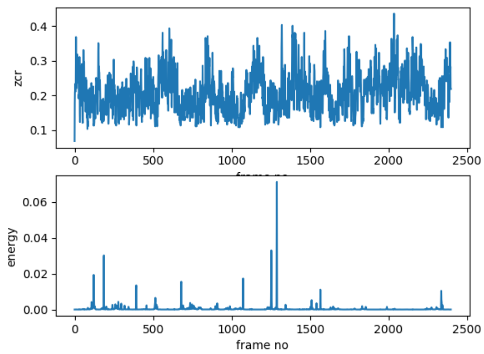
Fig 1: An interactive chordial graph, drawn on the basis of similarity measures of mosquito wingbeat frequency (Aegypti, albopictus, mediovittatus, anopheles atroparvus, anopheles merus), presenting identification of mosquito species



(a) Aedes aegypti



(c) Culex pipiens



(b) Anopheles astroparvus

The system can also generate infographics by indicating identification of specific mosquito species detected at a particular location on the map. Timely identification and classification of disease-carrying mosquitoes is a vital step towards taking actions to eradicate them. This system might be used as a step towards developing remote, non-destructive, and smart automated acoustic traps for eliminating disease transmitting mosquitoes. It might be used for continuous long-term unattended monitoring and automated logging for estimating population density of mosquitoes at a specific location.

7. Conclusions and Future Work

This paper presents an online tool that cluster and classify mosquito species, and produce an interactive online visualisation tool on a geographical map. This tool is produced by aggregating the results obtained by an unsupervised machine learning technique with the web services APIs that takes mosquito wingbeat audio recordings as input and cluster them on the basis of the similarity among the audio features. By automatic analysis of these similarity patterns, different mosquito species are clustered together that was used for the classification of mosquito species. The standard machine learning techniques are combined with powerful data-driven-document visualisation technique based on JavaScript to visualise similarity between different sound recordings and their mutual relationship with each other.

The variation in mosquito wingbeat frequency between and among species is common, but the visualisation graphs reveal some very interesting results showing some variations within-species songs. By 'understanding' and differentiating between within-species and between-species song and call variation, we may be able to detect and recognise their specific behaviour. An online tool detecting within-species song type variation opens up an exciting area of investigation for evolutionary and behavioural analysis of mosquito in varying environmental conditions.

This paper describes the first step of an ongoing project aiming at developing smart machines for mosquito surveillance in the public in the presence of noises such as wind, leaves, rain, thunderstorm, and other animals; classify a mosquito species based on its wingbeat frequency; and take appropriate action if a disease carrying mosquito is detected.

One of the biggest challenges in this area is the lack of annotated mosquito audio data, specially the data of disease carrying mosquitoes. Another problem is the external noise when the device is installed at a public place. We are working on developing techniques for removing background noise and improving the quality of input signals by various signal processing and machine learning algorithms. Human voices would be detected and removed before storing data into the database to avoid ethical issues. We are also working on developing an easy to use app that can be used by the community to record sound of mosquitoes in their regions. Entomologists from different organizations have been requested to annotate individual recording in the dataset. The dataset generated from this app would be stored on an online central repository that would be used as a standard for automatic mosquito surveillance systems.

8. References

[1] Barcelona Institute of Global Health (https://www.isglobal.org).

[2] J. E., & Rechcigl, N. A. (2016). Insect pest management : techniques for environmental protection. CRC Press.

[3] Fuchida, M.; Pathmakumar, T.; Mohan, R.E.; Tan, N.; Nakamura, A. Vision-Based Perception and Classification of Mosquitoes Using Support Vector Machine. *Appl. Sci.* 2017, *7*, 51.

[4] Favret C, Sieracki JM. Machine vision automated species identification scaled towards production levels. Syst Entomol. 2016;41(1):133–143. doi: 10.1111/syen.12146.

[5] Wang J, Lin C, Ji L, Liang A. A new automatic identification system of insect images at the order level. Knowl-Based Syst. 2012;33:102–110. doi:10.1016/j.knosys.2012.03.014.

[6] Pruszynski C. The BG-Counter: A new surveillance trap that remotely measures mosquito density in real-time. In: Sickerman SL, editor. Wing Beats. 27(1) Montgomery: Publications Press, Inc.; 2016.

[7] A. Rose, M. Weber, I. Potamitis, P. Villalonga, M. Geismar, M. Williams, J. Ecarnação, and M. Geier (2017). Automatically counting and identifying mosquitoes in real time: Latest update on the EU-funded REMOSIS project.

[8] Spitzen, J., & Takken, W. (2018). Keeping track of mosquitoes: a review of tools to track, record and analyse mosquito flight. *Parasites & vectors*, *11*(1), 123.

[9] Brogdon, W. G. (1994). Measurement of flight tone differences between female Aedes aegypti and A. albopictus (Diptera: Culicidae). *Journal of medical entomology*, *31*(5), 700-703.

[10] H. Mukundarajan, FJH. Hol, EA Castillo, C. Newby, M. Prakash (2017) Data from: Using mobile phones as acoustic sensors for high-throughput mosquito surveillance. Dryad Digital Repository. https://doi.org/10.5061/dryad.98d7s.

[11] Raman, D. R., Gerhardt, R. R., & Wilkerson, J. B. (2007). Detecting insect flight sounds in the field: Implications for acoustical counting of mosquitoes. *Transactions of the ASABE*, *50*(4), 1481-1485.

[12] Ouyang, T. H., Yang, E. C., Jiang, J. A., & Lin, T. T. (2015). Mosquito vector monitoring system based on optical wingbeat classification. *Computers and Electronics in Agriculture*, *118*, 47-55.

[13] Tramonte, A. R., & Christofferson, R. C. (2019). Investigating the probability of establishment of Zika virus and detection through mosquito surveillance under different temperature conditions. *PloS one*, *14*(3), e0214306.

[14] Li, R., Xu, L., Bjørnstad, O. N., Liu, K., Song, T., Chen, A., ... & Stenseth, N. C. (2019). Climate-driven variation in mosquito density predicts the spatiotemporal dynamics of dengue. *Proceedings of the National Academy of Sciences*, *116*(9), 3624-3629.

[15] Caminade, C., Kovats, S., Rocklov, J., Tompkins, A. M., Morse, A. P., Colón-González, F. J., ... & Lloyd, S. J. (2014). Impact of climate change on global malaria distribution. *Proceedings of the National Academy of Sciences*, *111*(9), 3286-3291.

[16] Jolliffe, I. (2011). Principal component analysis. In International encyclopedia of statistical science (pp. 1094-1096). Springer, Berlin, Heidelberg.

[17] Dangeti, P. (2017). Statistics for Machine Learning. Packt Publishing Ltd.

[18] Senoussaoui M, Kenny P, Stafylakis T, Dumouchel P (2014) A Study of the cosine distance-based mean shift for telephone speech diarization, audio, speech, and language processing. IEEE/ACM Trans 22(1):217–227.

[19] N. Dehak, P. Kenny, R. Dehak, P. Dumouchel, and P. Ouellet, "Front-end factor analysis for speaker verification," IEEE Transactions on Audio, Speech, and Language Processing, Vol. 19, No. 4, May 2011, pp. 788-798.

[20] N. Dehak, R. Dehak, J. Glass, D. Reynolds, and P. Kenny, "Cosine Similarity Scoring without Score Normalization Techniques," Proc. IEEE Odyssey Workshop, Brno, Czech Republic, June 2010.

[21] N. Dehak, Z. Karam, D. Reynolds, R. Dehak, W. Campbell, and J. Glass, "A Channel-Blind System for Speaker Verification," Proc. ICASSP, pp. 4536-4539, Prague, Czech Republic, May 2011.

[22] M. Senoussaoui, P. Kenny, N. Dehak and P. Dumouchel, "An i-vector Extractor Suitable for Speaker Recognition with both Microphone and Telephone Speech," in Proc Odyssey Speaker and Language Recognition Workshop, Brno, Czech Republic, June 2010.

[23] Vaizman, Y., McFee, B., & Lanckriet, G. (2014). Codebook-based audio feature representation for music information retrieval. IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), 22(10), 1483-1493.

[24] Van Dongen, S., & Enright, A. J. (2012). Metric distances derived from cosine similarity and Pearson and Spearman correlations. arXiv preprint arXiv:1208.3145.

[25] Crockford, D. (2006). The application/json media type for javascript object notation (json) (No. RFC 4627).

[26] Bostock, M., Ogievetsky, V., & Heer, J. (2011). D³ data-driven documents. IEEE Transactions on Visualization & Computer Graphics, (12), 2301-2309.

[27] d3 chordial graph documentation, https://github.com/d3/d3-chord/.

[28] Giannakopoulos T. (2015). pyaudioanalysis: An open-source python library for audio signal analysis. PloS one, 10(12), e0144610.

[29] MDN web docs https://developer.mozilla.org/en-US/docs/Web/API/Audio Context, last updated on Aug 28, 2018.