Enhancing TDE-based Drone DoA Estimation with Genetic Algorithms and Zero Cyclic Sum

Rigel P. Fernandes
Program of Defense Engineering
Military Institute of Engineering (IME)
Rio de Janeiro - RJ, Brazil
rigelfernandes@gmail.com

José A. Apolinário Jr.
Program of Electrical Electrical
Military Institute of Engineering (IME)
Rio de Janeiro - RJ, Brazil
apolin@ime.eb.br

José M. de Seixas
Signal Processing Lab, COPPE.
Universidade Federal do Rio de Janeiro
Rio de Janeiro - RJ, Brazil
seixas@lps.ufrj.br

***Abstract*—This paper discusses a way to enhance an acoustic-based approach to obtaining the direction of arrival (DoA) of a drone's ego noise using a microphone array. We focus on obtaining better time delay estimations (TDE) from a set of possible candidates. Recently, a large number of works have been put forward to detect and classify drones with different techniques. However, more investigation is required to tackle the drone DoA estimation problem using the time difference of arrival between pairs of microphones for the case of strongly corrupted audio signals, possibly by noise and multipath. The main problem in a complex acoustic environment is accurately estimating the time difference of arrival. With a traditional approach, this task becomes nearly impossible without the line of sight assumption, that is, whenever the highest cross-correlation peak between signals does not correspond to the delay between them. This paper uses genetic algorithms to search for the correct delays between pairs of microphones among a set of possible delays (primary and secondary delays). We define a fitness function based on the concept of zero cyclic sum of closed loops, i.e., when forming a closed loop, the sum of all theoretical delays should equal zero. A drawback of closed loops is that incorrect delays may result in a zero-sum; we thus created a fitness function that considers all possible closed loops of a given array. We exploited different approaches to estimate the direction of arrival using the combination of genetic algorithms and zero cyclic sum. In our experiments, the method successfully found all correct delays in simulations, providing strong evidence of its effectiveness when a correct delay exists among multiple possible delays. Furthermore, in experimental trials, it significantly enhanced the number of correct delays detected, further validating its utility and potential in practical scenarios.**

***Index Terms*—small drones, DoA estimation, genetic algorithms, zero cyclic sum**

I. Introduction

Unmanned aerial vehicles (UAVs) [1], also referred to as drones, have gained immense popularity in civil [2], industrial, and military applications [3]. These devices can perform a myriad of tasks, e.g., delivery [4], surveillance [5], mapping [6], photography [7], and agriculture monitoring [8]. In recent years, small UAVs have improved greatly, e.g., increased range, speed, and payload capacity [9]. The dual-use capabilities of small UAVs [10] enable them to be utilized by various actors for different purposes. Defense forces can leverage the use of small UAVs to gain strategic advantage over their enemies by conducting surveillance and reconnaissance missions and launching strikes over the enemies [11], [12]. On the other hand, terrorists can easily adopt low-cost drones to carry out attacks on targets [9], [11]. Also, criminals can use them to commit crimes. It is, therefore, critical for both public and private sectors to develop and implement effective countermeasures against the malicious use of small UAVs [3].

A counter-drone system comprises two stages, e.g., threat evaluation and weapon assignment [13]. A significant challenge is the drone threat evaluation step, i.e., drone detection [14]–[16] and parameter estimation such as DoA [17]–[19], localization [20], [21], model [22], and payload additional weight [23].

Using different signals, a drone threat evaluation system can detect and estimate drone parameters; for instance, radar [24], radio frequency [25], optical [26]–[29], and acoustic signals [22], [30], [31]. Other works employ two or more sensors to detect drones, e.g., optical and acoustic [32], acoustic and RF [33], video and acoustics collected from a drone [34]. Extensive literature about drone detection and parameter estimation can be found in [35]–[37].

Among all parameters that can be estimated from a drone, the DoA is the most important to designate a weapon to neutralize the possible threat. In this work, we contribute to the acoustic-based drone DoA estimation problem. We develop a method to estimate the direction of arrival in the event of corrupted audio signals, i.e., when the signal-to-noise ratio (SNR) is low and/or in the event of strong multipath. This method uses primary and secondary peaks of the cross-correlation, i.e., it chooses from a set of possible delays.

DoA estimation using acoustics has been extensively studied in audio signal processing, and various methods have been proposed for estimating the DoA of sound sources. Most of the available approaches try to estimate the DoA from a subset of delays (obtained from the major cross-correlation peaks) [38], [39]. However, in some cases, almost all delays from the primary peaks of the cross-correlation are incorrect, making it impossible to estimate the DoA or enhance the results using a subset of delays obtained from the largest peaks. Another approach that searches not only for a subset

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoa de Nível Superior – Brazil (CAPES) – Finance Code 001. It was also supported by CNPq, FAPERJ, and the Brazilian Navy.

of microphone pairs but also for the correct peaks of cross-correlation functions was introduced in [40].

The quest for a consistent fundamental loop, or zero cyclic sum (ZCS) [41] as seen in [42], usually employs a limited number of secondary loops and/or a few cross-correlation pairs, as in [40]. Therefore, the main contribution of our method to the drone DoA estimation is the incorporation of secondary peaks of the cross-correlations to decide the best set of all possible $N = M(M-1)/2$ delays, M being the number of microphones, to estimate the drone DoA. This simple yet effective approach positively impacts minimizing the error of the estimated DoA due to corrupted signals while maintaining the expected good results for the case of uncorrupted signals.

Our method unleashes the TDE-based algorithms to estimate drone DoA with an enhanced set of delays that includes the secondary peaks of the cross-correlation with a heuristic search based on genetic algorithms (GA) and a fitness function based on ZCS. The use of GA comes from the fact that an exhaustive search would not be feasible. Hence, we have named our method GA-ZCS to reflect its unique approach to the problem.

The rest of this paper is organized as follows. Section II describes the problem and assumptions of the drone estimation problem, data collection, drone acoustic signal, DoA estimation, and time delay estimation problems. Section III describes the method proposed herein to tackle the DoA estimation problem. Section IV presents simulation and experimental results, while Section V concludes the paper.

II. The drone DoA estimation problem

A. Problem statement and assumptions

In this work, our focus is on addressing the challenging task of estimating the direction of arrival (DoA) using delay estimates obtained from audio signals emitted by a quadcopter drone. These signals are characterized by a low SNR, indicating the presence of intense background noise and potential multipath effects.

To conduct our simulations, we utilize the geometry of a compact acoustic array consisting of seven microphones, specifically the MiniDSP UMA-8 model [43]. By leveraging this array, we calculate the theoretical delays of acoustic front waves emitted by the drone, considering both zenith and azimuth angles denoted as θ and ϕ, respectively. These theoretical delays represent the ideal estimates in the absence of background noise and multipath effects. Additionally, we acknowledge that estimated delays might manifest as secondary peaks within a cross-correlation analysis.

To simulate real-world conditions, we construct a data matrix denoted as $\mathbf{V}_{N\times C}$. Each row of the matrix contains a combination of theoretical delay and other random delays that can occur due to factors such as multipath interference or other acoustic events in the surrounding environment.

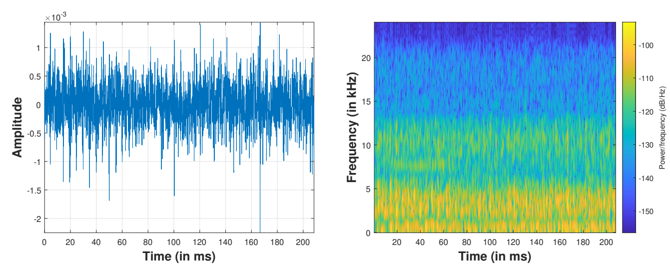
Moving on to the experimental phase, we gather two sets of acoustic drone signals, each lasting 20 seconds, employing the UMA-8 microphone array consisting of $M = 7$ microphones. From these signals, we estimate potential delays and construct the data matrix $\mathbf{V}_{N\times C}$. Each row of this matrix contains candidate delays, which are identified as peaks observed in the cross-correlations of microphone pairs.

Consequently, our primary challenge lies in identifying the correct delays among various cross-correlation peaks arising from different microphone pairs. The simulations we conduct serve as a proof of concept for the proposed method, while the experimental trials provide evidence of the benefits this method offers in addressing the drone DoA estimation problem. Notably, the actual drone signals employed in this study were obtained during the hovering of a DJI Phantom 4 quadcopter.

B. Drone acoustic signal

Figure 1 provides a visual representation of the drone hovering signal, showcasing 10,000 samples in the time domain. Additionally, a spectrogram computed with a sample rate of 48 kHz is presented. It is worth noting that the drone ego noise is primarily concentrated in the frequency range below 5 kHz. However, under favorable conditions and when the drone is in close proximity to the microphone array, it becomes possible to capture drone noise in higher frequencies, reaching up to 13.5 kHz.

For a more comprehensive exploration of the acoustic characteristics of drone noise, interested readers are encouraged to refer to the extensive study presented in [44]–[46]. These references delve into the intricacies of drone noise analysis and provide valuable insights into the subject matter.



(a) (b)

Fig. 1. Acoustic drone signal emitted by a DJI Phantom 4. (a) Time domain signal; and (b) Time-frequency representation.

C. DoA estimation with the GCC method

The Generalized Cross-Correlation (GCC) method is based on the cross-correlation function, which measures the similarity between two signals as a function of the time delay between them [47]. In the case of multiple microphones, the GCC method estimates the cross-correlation function between pairs of microphone signals to obtain the TDE between the signals. The basic principle is to find the time delay that maximizes the cross-correlation function, $r_{x_ix_j}(\tau)$ defined as:

$$r_{x_ix_j}(\tau) = \mathbb{E}[x_i(k)x_j(k-\tau)], \tag{1}$$

where $\mathbb{E}[\cdot]$ is the expectation operator and τ is the delay between two given sensors, x_i and x_j. From the peaks in the cross-correlation function, we find candidates for the time delay between the signals, which in turn can be used to determine the direction or location of the sound source.

The GCC is usually obtained as

$$\hat{r}^{\text{GCC}}_{x_i x_j}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \psi(\omega) X_i(\mathrm{e}^{\mathrm{j}\omega}) X_j(\mathrm{e}^{\mathrm{j}\omega})\, d\omega\,, \tag{2}$$

The Generalized Cross Correlation with Phase Transform (GCC-PHAT) normalizes the magnitude spectrum of the cross-correlation function as follows:

$$\psi^{\text{PHAT}}(\omega) = \frac{1}{|X_i(e^{\mathrm{j}\omega}) X_j(e^{-\mathrm{j}\omega})|}, \tag{3}$$

such that

$$\hat{r}^{\text{PHAT}}_{x_i x_j}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{X_i(\mathrm{e}^{\mathrm{j}\omega}) X_j(\mathrm{e}^{\mathrm{j}\omega})}{|X_i(\mathrm{e}^{\mathrm{j}\omega}) X_j(\mathrm{e}^{\mathrm{j}\omega})|}\, d\omega\,. \tag{4}$$

Finally, the TDE is obtained as follows:

$$\hat{\tau}_{ij} = \underset{|\tau|\leq\tau_{\max}}{\arg\max}\, |\hat{r}^{\text{PHAT}}_{x_i x_j}(\tau)|, \tag{5}$$

where $\tau_{\max}$ is the maximum possible delay (in number of samples) imposed by the distance between microphones i and j.

This normalization process effectively enhances the phase information while suppressing the amplitude differences between the signals, resulting in improved time delay estimation. By incorporating phase information, the GCC-PHAT method achieves enhanced robustness to reverberation and noise compared to the original GCC method [47].

After estimating the delays, we can minimize a least squares (LS) cost function to obtain a closed-form solution for the unit norm vector in the direction of the sound propagation. From this vector, the direction of arrival, azimuth (horizontal angle) and zenith (vertical angle, complement of the elevation), are readily available, as seen, for instance, in [40].

Interpolation plays another crucial role in the GCC-PHAT method, enhancing the accuracy and resolution of time delay estimation. By utilizing interpolation techniques, the GCC-PHAT method can estimate time delays with higher precision and handle sub-sample time delay resolutions. However, this technique is only effective if we are dealing with the correct peak.

D. Time delay estimation problems

One common source of error in cross-correlation-based time delay estimation is the presence of noise. When the signals being correlated are contaminated by noise, it can introduce spurious correlations and lead to incorrect time delay estimates. The noise can distort the shape of the cross-correlation function, resulting in erroneous peak positions or false peaks that do not correspond to the true time delay.

Another factor that can cause inaccurate time delay estimation is the presence of reverberation or multipath in the signals. Reverberation can significantly affect the shape and amplitude of the cross-correlation function, making it difficult to accurately identify the true peak representing the time delay. The reflections and multiple paths of sound propagation can create additional peaks or distort the main peak, leading to incorrect estimates.

Figure 2 illustrates several pertinent problems associated with TDE. In Figure 2 (a), we observe an accurate TDE both with and without interpolation, even in the presence of low levels of background noise. Figure 2 (b) showcases the benefits of interpolation, where a cross-correlation with fractional delays in samples yields a precise estimation. Figure 2 (c) presents a distorted function with a false main peak. Finally, Figure 2 (d) highlights a scenario with significant noise, wherein a secondary true peak emerges.

The abundance of different peaks of the cross-correlations involving drone noise instigates intriguing possibilities in the context of experimental trials. These possibilities include exploring secondary peaks, leveraging the peaks of interpolation, utilizing classical cross-correlation peaks, and considering samples before and after the main peak. By delving into these aspects, we can enhance our understanding and refine the TDE methodology in practical scenarios.

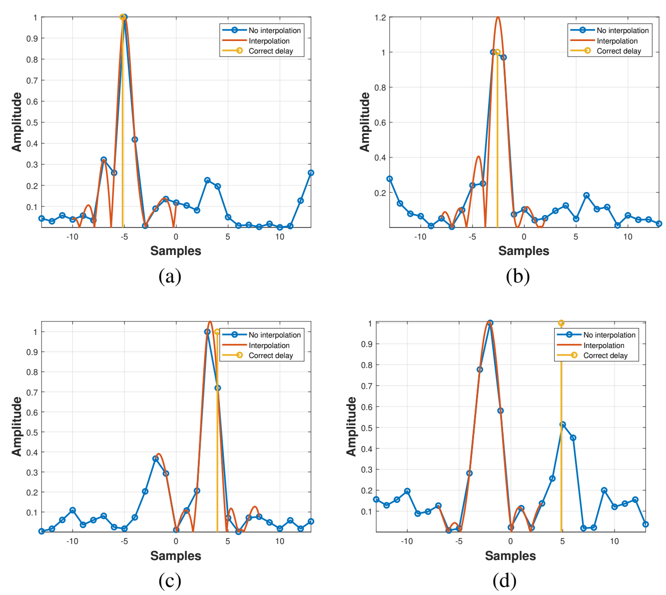


Fig. 2. Different time delay estimation problems. (a) Accurate time delay estimation with and without interpolation (b) More accurate time delay estimation with interpolation (c) distorted function with a false peak and (d) secondary true peak.

III. The Proposed Method

A. Genetic algorithms and zero cyclic sum

GA play a crucial role as a powerful heuristic search technique in solving complex problems [48]. It is inspired by the principles of natural selection and evolution, mimicking the process of survival of the fittest to find global or local optimal solutions.

One of the key advantages of genetic algorithms is their ability to handle large solution spaces and navigate through complex landscapes of possibilities. Unlike traditional search methods, GAs do not rely on explicit problem domain knowledge or constraints. Instead, they explore the solutions space by iteratively generating and evaluating a population of candidate solutions.

The size of the solutions space in this work denoted as S, is determined by the formula $S = C^N$, where C represents the number of candidate delays and $N = 21$ for $M = 7$ microphones. For instance, if $C = 2$ the solutions space is $C^2 = 2,097,152$. As we increase the value of C, the solutions space grows exponentially. When $C = 3$, the solutions space expands to a massive 10,460,353,203 potential solutions. If $C = 4$ the solution space would have 4,398,046,511,104 potential solutions. This number represents trillions of unique combinations of four candidate delays that need to be considered.

These enormous solution spaces pose significant challenges for traditional search methods, as exhaustively evaluating each possible solution becomes computationally infeasible. This is where genetic algorithms excel. By employing a heuristic search approach, genetic algorithms can efficiently explore these expansive solutions spaces and navigate toward local or global optimal solutions without having to evaluate every single possibility.

GAs leverage the concept of individuals represented as chromosomes, where each chromosome encodes a potential solution to the problem. These solutions are evaluated based on the ZCS fitness function that quantifies their proximity to a zero-sum. Through the use of selection, crossover, and mutation operators, GAs promote the exchange and recombination of genetic material between individuals, mimicking the genetic diversity and variation found in natural evolution.

The method employs GA to efficiently identify the correct delays from a multitude of incorrect delays, particularly in situations with low SNR. The C candidate delays for each cross-correlation function $r_{x_i x_j}$ are the elements of each row of data matrix $\mathbf{V}$ denoted as $\{\tau_{ij,1} \quad \tau_{ij,2} \quad \dots \quad \tau_{ij,C}\}$. For $M = 7$, which implies in $N = 21$, the matrix $\mathbf{V}_{N\times C}$ with all candidate delays, is defined as

$$\mathbf{V} = \begin{bmatrix} \tau_{12,1} & \tau_{12,2} & \tau_{12,3} & \cdots & \tau_{12,C} \\ \tau_{13,1} & \tau_{13,2} & \tau_{13,3} & \cdots & \tau_{13,C} \\ \tau_{14,1} & \tau_{14,2} & \tau_{14,3} & \cdots & \tau_{14,C} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \tau_{67,1} & \tau_{67,2} & \tau_{67,3} & \cdots & \tau_{67,C} \end{bmatrix}. \tag{6}$$

The population, consisting of I individuals (or chromosomes), represents the collection of potential solutions to the problem. Each chromosome consists of genes, denoted as g, which can take values from 1 to C, according to the number of candidate delays C. These genes allow for the exploration of all possible delay candidates in the matrix $\mathbf{V}_{N\times C}$. The chromosome is represented as a column of the matrix $\mathbf{P}_{N\times I}$:

$$\mathbf{P} = \begin{bmatrix} g_{11} & g_{12} & g_{13} & \cdots & g_{1I} \\ g_{21} & g_{22} & g_{23} & \cdots & g_{2I} \\ g_{31} & g_{32} & g_{33} & \cdots & g_{3I} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{N1} & g_{N2} & g_{N3} & \cdots & g_{NI} \end{bmatrix}, \tag{7}$$

the matrix $\mathbf{P}_{N\times I}$ is a set of possible solutions, thus we create a vector of delays using each column of $\mathbf{P}_{N\times I}$, for instance if $\mathbf{P}_{(:,1)} = \{1, \ 3, \ 2, \ \dots, \ 9\}$ the corresponding vector of delays, $\mathbf{v}$, corresponds to

$$\mathbf{v} = [\tau_{12,1} \ \ \tau_{13,3} \ \ \tau_{14,2} \ \ \dots \ \ \tau_{67,9}]^{\mathrm{T}}.$$

B. Fitness function

The fitness function stands as a crucial component within the method, with the ZCS being employed for this purpose. This function evaluates the sum of delays within specific subsets that form closed loops, aiming to minimize the occurrence of erroneous zero-sum outcomes. By considering all possible subsets that create closed loops and summing their results, we significantly reduce the likelihood of encountering a zero-sum result without the correct delays.

To facilitate the computational calculation of the fitness function, we have devised a method that involves the identification and enumeration of closed loops based on the number of delays. More specifically, when employing a $M = 7$ microphone array, we find that, with 3 delays, there are 35 closed loops. Similarly, with 4 delays, we observe 35 closed loops. Moving on to 5 delays, we encounter 21 closed loops, while 6 delays give rise to 7 closed loops. Finally, when utilizing 7 delays, a single closed loop is formed. The total number of closed loops, denoted as L, corresponds to 99 in this context; $L = 35+35+21+7+1 = 99$. It is important to note that the delay τ_{31}, which closes the loop, can be determined by taking the negative value of τ_{13}. Similarly, τ_{75} can be expressed as $-\tau_{57}$, and in general, any delay τ_{ji} that closes the loop can be written as $\tau_{ji} = -\tau_{ij}$. By utilizing this property, we can compute all possible delays once and then manipulate them to identify the correct value of τ that closes the loop. This approach saves computational resources by avoiding redundant calculations and facilitates the determination of the correct delay for loop closure. The complete listing of all possible closed loops for 3, 4, 5, 6, and 7 delays can be found in Table I.

We can create a matrix $\mathbf{D}_{L\times N}$ based on Table I to be able to sum all delays

$$\mathbf{D} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & \dots & 0 & 1 \end{bmatrix}; \tag{8}$$

TABLE I
ALL POSSIBLE CYCLIC PATHS IN A SEVEN MICROPHONE ARRAY

# delays	Closed loops						
3	τ_{12}	τ_{23}	τ_{31}				
	$\vdots$	$\vdots$	$\vdots$				
	τ_{56}	τ_{67}	τ_{75}				
4	τ_{12}	τ_{23}	τ_{34}	τ_{41}			
	$\vdots$	$\vdots$	$\vdots$	$\vdots$			
	τ_{45}	τ_{56}	τ_{67}	τ_{74}			
5	τ_{12}	τ_{23}	τ_{34}	τ_{45}	τ_{51}		
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		
	τ_{34}	τ_{45}	τ_{56}	τ_{67}	τ_{73}		
6	τ_{12}	τ_{23}	τ_{34}	τ_{45}	τ_{56}	τ_{61}	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
	τ_{23}	τ_{34}	τ_{45}	τ_{56}	τ_{67}	τ_{72}	
7	τ_{12}	τ_{23}	τ_{34}	τ_{45}	τ_{56}	τ_{67}	τ_{71}

each element of the resulting vector $\mathbf{f} = \mathbf{D}\mathbf{v}$ is the sum of all subsets such that the fitness function, denoted as f, is then calculated as

$$f = \mathbf{f}^{\mathbf{T}}\mathbf{f} = \|\mathbf{f}\|^2.$$

This fitness function captures the squared norm of the resulting vector, encompassing the contributions from all subsets and providing a measure of the fitness or quality of the estimation. The GA-ZCS method is detailed in Algorithm 1.

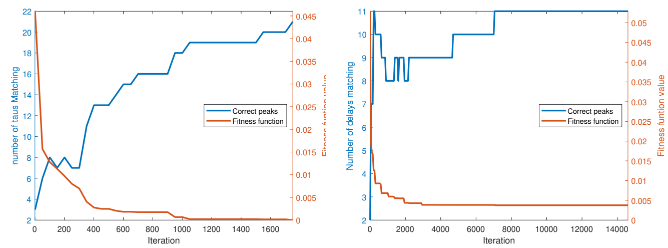
Algorithm 1 Heuristic search using genetic algorithms with zero cyclic sum fitness function (GA-ZCS)

for $i = 1 : N$ **do**
  Compute $r_{x_i x_j}, ij = 12$ to 67
  Obtain C candidate delays (larger peaks of $r_{x_i x_j}$)
  $\mathbf{V}_{i,:} \Leftarrow [\tau_{ij,1} \;\; \tau_{ij,2} \;\; \dots \;\; \tau_{ij,C}]$
end for
Create population $\mathbf{P}_{N\times I}$ of random integers $[1, C]$
for $i = 1 : N$ **do**
  $\mathbf{v}_i \Leftarrow$ map $\tau_{ij,x}$ in $\mathbf{V}_{i,:}$ according to $\mathbf{P}_{:,k}$
  First evaluation of individuals $\mathbf{P}_{N\times I}$
end for
while $f > 10^{-15}$ OR $k < 2000$ **do**
  $\mathbf{v}_i \Leftarrow$ map $\tau_{ij,x}$ in $\mathbf{V}_{i,:}$ according to $\mathbf{P}_{:,k}$
  Crossover neighbor individuals
  Mutate
  $f \leftarrow$ Evaluate individuals $\mathbf{P}_{N\times I}$
  $\mathbf{P}_{N\times I} \Leftarrow$ Select the best I individuals
  Increment k
end while

IV. RESULTS

A. Simulation results

The initial evaluation of the proposed approach for drone DoA estimation involved simulated delays, which approximate the potential delays based on the utilized array geometry. Figure 3 illustrates the evolution of the fitness function and the number of matching delays with the theoretical 21 delays. It is evident that the ZCS fitness function serves as a guiding measure for the algorithm, leading to the enhancement of individuals within the genetic algorithm. By using this fitness function, the algorithm is directed towards improving the accuracy of delay estimates by identifying the correct peaks among the $C = 10$ possible delays for each cross-correlation function. This facilitates the overall improvement of the algorithm's performance in estimating the delays more effectively. However, it is important to note that the GA may converge to a local minimum, and reaching the global minimum could be a time-consuming process.



(a) (b)

Fig. 3. GA search progress with fitness function and number of correct delays evolution. (a) 1600 iterations and the search stopped according to the fitness function criteria (b) GA and the convergence to a local minimum.

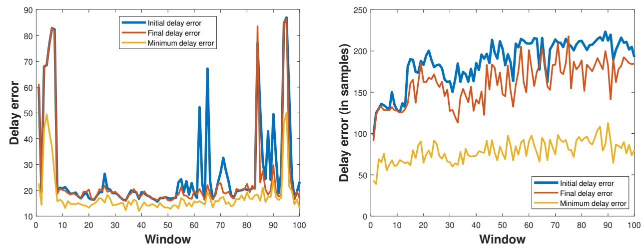
B. Experimental results

After conducting simulation tests, we proceeded to test the proposed method with actual drone noise signals. The signals we evaluated were captured while the drone was hovering at a distance of 30 and 280 meters from the microphone array. We divided each signal into 100 segments of 200ms each, allowing us to estimate the delays between sensor pairs. We collected 6 candidate delays for each $r_{x_i x_j}$ forming a matrix $\mathbf{V}_{N\times 6}$.

Contrasting to simulations, where we had theoretical delays among other delays, it is highly unlikely for the peaks of an actual cross-correlation to correspond perfectly with the theoretical delay. Therefore, we monitored the progress of the GA by assessing the delay error (in samples), which is calculated as $\sum_{i=1}^{N} |\hat{\mathbf{v}}_i - \mathbf{v}_i|$, $\hat{\mathbf{v}}$ represents the vector of estimated delays mapped by the GA, and $\mathbf{v}$ consists of the theoretical delays.

Figure 4 illustrates the delay error for each analyzed window. Specifically, Figure 4 (a) showcases the delay error progression for each window of 200 ms analyzed. Figure 4 (b) highlights the benefits of utilizing the GA algorithm in mitigating large delay errors.

The results depicted in Figure 4 denote the minimum delay error, i.e., the minimum possible error if we choose the element of matrix $\mathbf{V}$ that minimizes the error of the TDE. The simulation results prove that GA-ZCS can find all theoretical delays, so one challenge is to develop a method that finds an accurate delay for each pair of microphones.



(a) Drone 30 m from the array. (b) Drone 280 m from the array.

Fig. 4. Progress of GA search with fitness function, minimum delay error in samples, and the error reduction highlighting the benefits of the heuristic search. (a) Analysis of 100 windows using GA search progress while the drone hovers in close proximity to the microphone array (30 m); and (b) Analysis of 100 windows with the drone hovering at a distance of 280 m from the microphone array.

Introducing a signal enhancement technique as a step in this method may enhance the signal of interest, specifically the drone signal, in each channel. This enhancement would lead to a peak in the cross-correlation function associated with the drone. By applying signal enhancement techniques, we can potentially improve the detectability and accuracy of the drone signal, thereby enhancing the performance of the time delay estimation method. Exploring such signal enhancement techniques holds promise for further improving the robustness and effectiveness of the overall approach.

V. Conclusions

We utilized genetic algorithms with a zero cyclic sum fitness function to tackle the time delay estimation problem. Our results revealed that obtaining a single accurate time delay estimate out of the ten candidate delays enabled the method to achieve the optimal solution. This emphasizes the effectiveness of genetic algorithms in mitigating the time delay estimation problem in the context of drone signals. We are currently investigating a more advanced approach to enhance the cross-correlation function of drone signals. Furthermore, we provided an explanation for the experimental results falling short of reaching the optimal solution. By comparing the delay error between the traditional delay estimation method (peak of the cross-correlation) and GA-ZCS, we observed that GA-ZCS has the potential to make a significant contribution to the field of delay estimation. This highlights the promising nature of exploiting a set of candidate delays to minimize the error in time delay estimation of drone acoustic signals. As the next step, we plan to collect more acoustic drone signals with different background noise levels and distinct multipath effects. This will allow us to further validate and refine our method.

References

[1] S. A. H. Mohsan, N. Q. H. Othman, Y. Li, M. H. Alsharif, and M. A. Khan, "Unmanned aerial vehicles (UAVs): practical aspects, applications, open challenges, security issues, and future trends," *Intelligent Service Robotics*, pp. 1–29, 2023. [Online]. Available: https://doi.org/10.1007/s11370-022-00452-4

[2] K. Al-Dosari, Z. Hunaiti, and W. Balachandran, "A review of civilian drones systems, applications, benefits, safety, and security challenges," *The Effect of Information Technology on Business and Marketing Intelligence Systems*, pp. 793–812, 2023. [Online]. Available: https://doi.org/10.1007/978-3-031-12382-5_43

[3] D. Kunertova, "The war in ukraine shows the game-changing effect of drones depends on the game," *Bulletin of the Atomic Scientists*, vol. 79, no. 2, pp. 95–102, 2023. [Online]. Available: https://doi.org/10.1080/00963402.2023.2178180

[4] H. Eskandaripour and E. Boldsaikhan, "Last-mile drone delivery: Past, present, and future," *Drones*, vol. 7, no. 2, p. 77, 2023. [Online]. Available: https://doi.org/10.3390/drones7020077

[5] F. G. Serrenho, J. A. Apolinário Jr., A. L. L. Ramos, and R. P. Fernandes, "Gunshot airborne surveillance with rotary wing UAV-embedded microphone array," *Sensors*, vol. 19, no. 19, p. 4271, 2019. [Online]. Available: https://doi.org/10.3390/s19194271

[6] N. Afshar-Mohajer and C.-Y. Wu, "Use of a drone-based sensor as a field-ready technique for short-term concentration mapping of air pollutants: A modeling study," *Atmospheric Environment*, vol. 294, p. 119476, 2023. [Online]. Available: https://doi.org/10.1016/j.atmosenv.2022.119476

[7] M. Shawky, A. Alsobky, A. Al Sobky, and A. Hassan, "Traffic safety assessment for roundabout intersections using drone photography and conflict technique," *Ain Shams Engineering Journal*, p. 102115, 2023. [Online]. Available: https://doi.org/10.1016/j.asej.2023.102115

[8] P. Goodrich, O. Betancourt, A. C. Arias, and T. Zohdi, "Placement and drone flight path mapping of agricultural soil sensors using machine learning," *Computers and Electronics in Agriculture*, vol. 205, p. 107591, 2023. [Online]. Available: https://doi.org/10.1016/j.compag.2022.107591

[9] T. Pledger, "The role of drones in future terrorist attacks," *Association of the United States army*, vol. 26, 2021. [Online]. Available: https://www.ausa.org/publications/role-drones-future-terrorist-attacks

[10] M. Suojanen, V. Kuikka, J.-P. Nikkarila, and J. Nurmi, "An example of scenario-based evaluation of military capability areas. An impact assessment of alternative systems on operations," in *Annual IEEE Systems Conference (SysCon) Proceedings*. IEEE, 2015, pp. 601–607. [Online]. Available: https://doi.org/10.1109/SYSCON.2015.7116817

[11] J. Watling and N. Waters, "Achieving lethal effects by small unmanned aerial vehicles: Opportunities and limitations," *The RUSI Journal*, vol. 164, no. 1, pp. 40–51, 2019. [Online]. Available: https://doi.org/10.1080/03071847.2019.1605017

[12] S. Kreps and P. Lushenko, "Drones in modern war: evolutionary, revolutionary, or both?" *Defense & Security Analysis*, pp. 1–4, 2023. [Online]. Available: https://doi.org/10.1080/14751798.2023.2178599

[13] A. Naseem, S. T. H. Shah, S. A. Khan, and A. W. Malik, "Decision support system for optimum decision making process in threat evaluation and weapon assignment: Current status, challenges and future directions," *Annual reviews in control*, vol. 43, pp. 169–187, 2017. [Online]. Available: https://doi.org/10.1016/j.arcontrol.2017.03.003

[14] M. Z. Anwar, Z. Kaleem, and A. Jamalipour, "Machine learning inspired sound-based amateur drone detection for public safety applications," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2526–2534, 2019. [Online]. Available: https://doi.org/10.1109/TVT.2019.2893615

[15] E. Akbal, A. Akbal, S. Dogan, and T. Tuncer, "An automated accurate sound-based amateur drone detection method based on skinny pattern," *Digital Signal Processing*, p. 104012, 2023. [Online]. Available: https://doi.org/10.1016/j.dsp.2023.104012

[16] S. Al-Emadi, A. Al-Ali, and A. Al-Ali, "Audio-based drone detection and identification using deep learning techniques with dataset enhancement through generative adversarial networks," *Sensors*, vol. 21, no. 15, p. 4953, 2021. [Online]. Available: https://doi.org/10.1016/j.dsp.2023.104012

[17] A. Kyritsis, R. Makri, and N. Uzunoglu, "Small UAS online audio DOA estimation and real-time identification using machine learning," *Sensors*, vol. 22, no. 22, p. 8659, 2022. [Online]. Available: https://doi.org/10.3390/s22228659

[18] Z. Shi, X. Chang, C. Yang, Z. Wu, and J. Wu, "An acoustic-based surveillance system for amateur drones detection and localization," *IEEE transactions on vehicular technology*, vol. 69, no. 3, pp. 2731–2739, 2020. [Online]. Available: https://doi.org/10.1109/TVT.2020.2964110

[19] C. Yang, Z. Wu, X. Chang, X. Shi, J. Wo, and Z. Shi, "DOA estimation using amateur drones harmonic acoustic signals," in

2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM). IEEE, 2018, pp. 587–591. [Online]. Available: https://doi.org/10.1109/SAM.2018.8448797

[20] W. Qiu, G. Wang, and W. Zhang, "Acoustic slam based on the direction-of-arrival and the direct-to-reverberant energy ratio," *Drones*, vol. 7, no. 2, p. 120, 2023. [Online]. Available: https://doi.org/10.3390/drones7020120

[21] Z. Wu, C. Yang, X. Chang, J. Wu, and Z. Shi, "A drone localization system based on multiple time delays fusion," in *2019 IEEE Radar Conference (RadarConf)*. IEEE, 2019, pp. 1–6. [Online]. Available: https://doi.org/10.1109/RADAR.2019.8835531

[22] R. P. Fernandes, J. A. Apolinário Jr., A. L. L. Ramos, and J. M. de Seixas, "Applying the majority voting rule in acoustic detection and classification of drones," in *XXXIX Brazilian Symposium on Telecommunications and Signal Processing*. Sociedade Brasileira de Telecomunicações, 2021. [Online]. Available: https://doi.org/10.14209/sbrt.2021.1570726624

[23] O. A. Ibrahim, S. Sciancalepore, and R. Di Pietro, "Noise2Weight: On detecting payload weight from drones acoustic emissions," *Future Generation Computer Systems*, vol. 134, pp. 319–333, 2022. [Online]. Available: https://doi.org/10.1016/j.future.2022.03.041

[24] H. Dale, C. Baker, M. Antoniou, M. Jahangir, G. Atkinson, and S. Harman, "SNR-dependent drone classification using convolutional neural networks," *IET Radar, Sonar & Navigation*, vol. 16, no. 1, pp. 22–33, 2022. [Online]. Available: https://doi.org/10.1049/rsn2.12161

[25] M. F. Al-Sa'd, A. Al-Ali, A. Mohamed, T. Khattab, and A. Erbad, "RF-based drone detection and identification using deep learning approaches: An initiative towards a large open source drone database," *Future Generation Computer Systems*, vol. 100, pp. 86–97, 2019. [Online]. Available: https://doi.org/10.1016/j.future.2019.05.007

[26] U. Seidaliyeva, M. Alduraibi, L. Ilipbayeva, and N. Smailov, "Deep residual neural network-based classification of loaded and unloaded UAV images," in *2020 Fourth IEEE International Conference on Robotic Computing (IRC)*. IEEE, 2020, pp. 465–469. [Online]. Available: https://doi.org/10.1109/IRC.2020.00088

[27] D. Lee, W. G. La, and H. Kim, "Drone detection and identification system using artificial intelligence," in *2018 International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 2018, pp. 1131–1133. [Online]. Available: https://doi.org/10.1109/ICTC.2018.8539442

[28] Y.-C. Lai and Z.-Y. Huang, "Detection of a moving UAV based on deep learning-based distance estimation," *Remote Sensing*, vol. 12, no. 18, p. 3035, 2020. [Online]. Available: https://doi.org/10.3390/rs12183035

[29] J. Xie, J. Yu, J. Wu, Z. Shi, and J. Chen, "Adaptive switching spatial-temporal fusion detection for remote flying drones," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 6964–6976, 2020. [Online]. Available: https://doi.org/10.1109/TVT.2020.2993863

[30] A. Sedunov, H. Salloum, A. Sutin, N. Sedunov, and S. Tsyuryupa, "UAV passive acoustic detection," in *2018 IEEE International Symposium on Technologies for Homeland Security (HST)*. IEEE, 2018, pp. 1–6. [Online]. Available: https://doi.org/10.1109/THS.2018.8574129

[31] H. Liu, K. Fan, and B. He, "Acoustic source localization for anti-uav based on machine learning in wireless sensor networks," in *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 2020, pp. 1142–1147. [Online]. Available: https://doi.org/10.1109/ICIEA48937.2020.9248139

[32] F. Christnacher, S. Hengy, M. Laurenzis, A. Matwyschuk, P. Naz, S. Schertzer, and G. Schmitt, "Optical and acoustical UAV detection," in *Electro-Optical Remote Sensing X*, vol. 9988. SPIE, 2016, pp. 83–95. [Online]. Available: https://doi.org/10.1117/12.2240752

[33] A. Toma, N. Cecchinato, C. Drioli, G. L. Foresti, and G. Ferrin, "CNN-based processing of radio frequency signals for augmenting acoustic source localization and enhancement in uav security applications," in *2021 International Conference on Military Communication and Information Systems (ICMCIS)*. IEEE, 2021, pp. 1–5. [Online]. Available: https://doi.org/10.1109/ICMCIS52405.2021.9486424

[34] A. Toma, N. Cecchinato, C. Drioli, G. Oliva, G. Ferrin, G. Sechi, and G. L. Foresti, "Onboard audio and video processing for secure detection, localization, and tracking in counter-uav applications," *Procedia Computer Science*, vol. 205, pp. 20–27, 2022. [Online]. Available: https://doi.org/10.1016/j.procs.2022.09.003

[35] B. Taha and A. Shoufan, "Machine learning-based drone detection and classification: State-of-the-art in research," *IEEE Access*, vol. 7, pp. 138 669–138 682, 2019. [Online]. Available: https://doi.org/10.1109/ACCESS.2019.2942944

[36] G. Lykou, D. Moustakas, and D. Gritzalis, "Defending airports from UAS: A survey on cyber-attacks and counter-drone sensing technologies," *Sensors*, vol. 20, no. 12, p. 3537, 2020. [Online]. Available: https://doi.org/10.3390/s20123537

[37] D. Utebayeva, L. Ilipbayeva, and E. T. Matson, "Practical study of recurrent neural networks for efficient real-time drone sound detection: A review," *Drones*, vol. 7, no. 1, p. 26, 2022. [Online]. Available: https://doi.org/10.3390/drones7010026

[38] A. M. Borzino, J. A. Apolinário, and M. L. de Campos, "Robust DOA estimation of heavily noisy gunshot signals," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015, pp. 449–453. [Online]. Available: https://doi.org/10.1109/ICASSP.2015.7178009

[39] I. L. Freire, P. C. Prandel, and J. A. Apolinário Jr., "Sobre a escolha de sinais em arranjos de microfones estimando DOA com gcc-phat (in portuguese)," *Simpósio Brasileiro de Telecomunicações. SBrT*, 2012. [Online]. Available: http://dx.doi.org/10.14209/sbrt.2012.76

[40] Â. M. C. R. Borzino, J. A. Apolinario Jr, and M. L. R. de Campos, "Consistent DOA estimation of heavily noisy gunshot signals using a microphone array," *IET Radar, Sonar & Navigation*, vol. 10, no. 9, pp. 1519–1527, 2016. [Online]. Available: https://doi.org/10.1049/iet-rsn.2016.0015

[41] P. Yuan, "On the index of minimal zero-sum sequences over finite cyclic groups," *Journal of Combinatorial Theory, Series A*, vol. 114, no. 8, pp. 1545–1551, 2007. [Online]. Available: https://doi.org/10.1016/j.jcta.2007.03.003

[42] J. Scheuing and B. Yang, "Disambiguation of TDOA estimation for multiple sources in reverberant environments," *IEEE transactions on audio, speech, and language processing*, vol. 16, no. 8, pp. 1479–1489, 2008. [Online]. Available: https://doi.org/10.1109/TASL.2008.2004533

[43] MiniDSP, "UMA-8 USB mic array - V2.0," https://www.minidsp.com/products/usb-audio-interface/uma-8-microphone-array/, 2023, [Online; accessed 05-june-2023].

[44] L. Wang and A. Cavallaro, "Ear in the sky: Ego-noise reduction for auditory micro aerial vehicles," in *2016 13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE, 2016, pp. 152–158. [Online]. Available: https://doi.org/10.1109/AVSS.2016.7738063

[45] ——, "Acoustic sensing from a multi-rotor drone," *IEEE Sensors Journal*, vol. 18, no. 11, pp. 4570–4582, 2018. [Online]. Available: https://doi.org/10.1109/JSEN.2018.2825879

[46] D. Mukhutdinov, A. Alex, A. Cavallaro, and L. Wang, "Deep learning models for single-channel speech enhancement on drones," *IEEE Access*, 2023. [Online]. Available: https://doi.org/10.1109/ACCESS.2023.3253719

[47] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE transactions on acoustics, speech, and signal processing*, vol. 24, no. 4, pp. 320–327, 1976. [Online]. Available: https://doi.org/10.1109/TASSP.1976.1162830

[48] M. Srinivas and L. M. Patnaik, "Genetic algorithms: A survey," *computer*, vol. 27, no. 6, pp. 17–26, 1994. [Online]. Available: https://doi.org/10.1109/2.294849