Estimation of Dolphin Location by Acoustic Characterization of Click Sequences

Benjamin Masters[1,2] **and Pranav Barot**[1,3]

[1]**University of Waterloo, Department of Systems Design Engineering**
[2]**bpmasters@uwaterloo.ca**
[3]**pbarot@uwaterloo.ca**

INTRODUCTION

This year's international student challenge problem in acoustic signal processing provides hydrophone recordings of dolphins emitting echolocation clicks, and tasks participants with (1) estimating the time of arrival of clicks, and the timing difference between clicks of each dolphin and (2) estimating the position of the dolphins at each click by manipulating the time difference of arrival.

To do so, we apply spectral and temporal signal processing methods on the recordings to identify the time that clicks arrive to each hydrophone, and utilize the geometric relationship between hydrophones to estimate characteristics of the dolphins' position and motion over time.

TASK 1

Methods

Preprocessing

As with any real-world signals, there is a significant amount of noise on the three hydrophones that could mask the stimulus of interest, which are clicks from dolphins. Before attempting any de-noising, the spectral properties of the signals were analyzed in a window that was manually identified to contain a click, and compared with one that did not contain a click. These power spectra are displayed in Figure 1. It's evident that there is significantly higher energy at most frequencies above 25 kHz. However, it seems the energy difference is maximal between 25 kHz and 75 kHz. Therefore, the hydrophone signals were bandpass filtered within this range using a 20th order Butterworth filter. Given that the timing of clicks is important, a zero-phase filtering implementation was chosen to maintain accuracy of time of arrival of clicks.

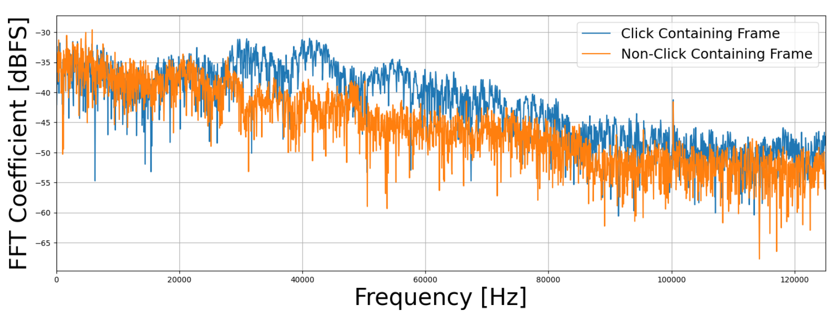

Figure 1. Spectrogram of first 40 seconds of audio from Hydrophone N

From this point, the signals were less noisy, as displayed in Figure 2, and a cut-off point between click trains belonging to each dolphin was selected at 4.4 seconds. The exact same cut-off time was applied to all three hydrophone signals, such that synchrony of the time of arrival between hydrophones was maintained.

Estimating the location of clicks

To estimate the location of clicks at hydrophone 2 (H2), the recording was first buffered into overlapping adjacent frames and the power was computed within each frame. As the time between adjacent dolphin clicks can be as low as 7 ms (Ridgway et al., 2018), it's important that these frame parameters (size and

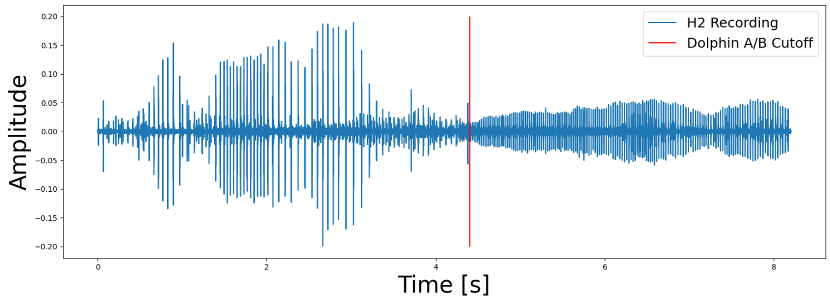

Figure 2. Waveform of hydrophone 2 recordings and a manually identified cutoff point between dolphin click trains.

overlap) are selected carefully. Given the provided sample rate of 250 kHz, this would suggest a maximum frame size of 1750 samples. However, one must also consider the duration of said clicks, as using too-large of a window could result in smearing of a click's energy such that it is no longer distinguishable in any given frame. Ivanov (2004) reported a 'typical' duration of 40 μs, which translates to 10 samples in our signals. Therefore, to maximally capture energy differences between frames containing and not containing clicks, a frame size of 16 samples was used, such that clicks slightly longer than typical would be fully captured. To increase the resolution of the buffered signal, 50% overlap between adjacent frames was used.

From the frame-based power signals, clicks are estimated using a prominence based peak-detection method, as simple thresholding didn't result in robust estimation given the considerable changes in received level over the course of the signal, and the presence of reflections as discussed later. The *find_peaks* function from *scipy*'s signal processing library was used to identify peaks in the frame-by-frame power with a prominence of 35 dB for both dolphins. A further restriction was invoked on the frame-wise distance between adjacent clicks, requiring a minimum distance of 6 ms between frames, a value selected to be just under the previously discussed range of inter-click intervals. These detected peaks are assumed to be clicks, and are plotted alongside the frame-by-frame power for H2 in Figure 3. The inclusion of the minimum distance between frames appears to be important, as Figure 4 illustrates close consecutive peaks. We hypothesize that these consecutive, power-reduced impulses are a result of acoustic reflections, given that the closest peak was found manually to occur $\sim$ 10 μs after the initial impulse. This time difference corresponds to a travel time of $\sim$1.5 m, potentially suggesting that it's the first specular reflection of the initial signal arriving from the sea floor.

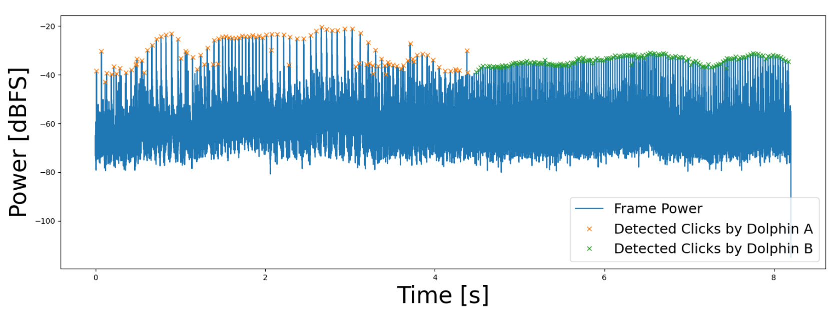

Figure 3. Frame-by-frame power of hydrophone 2 using a window size of 16 samples with 50% overlap. Also plotted are the locations of detected clicks by each dolphin.

Task 1 Deliverables

(1) Uncertainty in timing of click estimation

The uncertainty in our click estimate is related to the resolution of our buffered time domain signal. Given the window size of 16 samples and 50% overlap between adjacent frames, the sample rate of the buffered signal will be $f_b = \frac{f_s}{N(1-O)}$, where f_s is the sample rate of the hydrophone signals (250 kHz), N is the

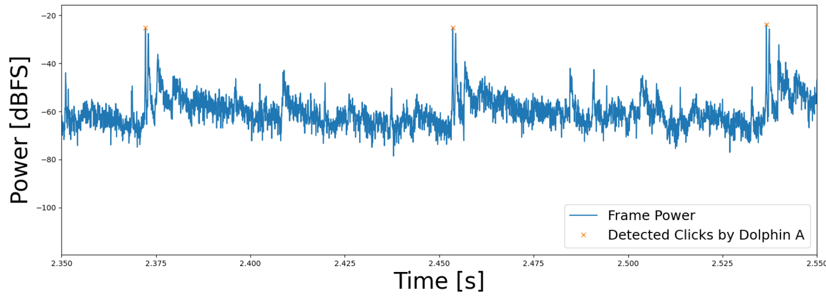

Figure 4. A zoomed in segment of Figure 3, illustrating the locations of clicks along side other extraneous impulses, which are assumed to be arrival of reflections.

window size, and O is the amount of overlap (50% = 0.5). Given that the windows are symmetric around a given point in time, the uncertainty will be equal to half the sampling period. Thus, the uncertainty in our click estimation is as in Equation 1.

$$Uncertainty = \pm \frac{1}{2f_b} = \pm 16\mu s \tag{1}$$

(2) Inter-click intervals

The inter-click interval (ICI) can be estimated by taking the first difference of the locations in time of the previously identified clicks. The results are plotted in Figure 5, and demonstrate that the clicks from Dolphin B are shorter and much less variable than Dolphin A, as verified in the results shown in Table 1.

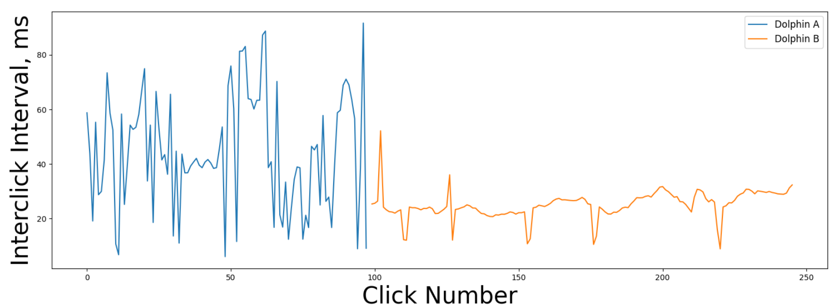

Figure 5. Plot of inter-click intervals by click number, for each dolphin.

Dolphin	Number of Clicks	Mean ICI (ms)	Std. Dev. ICI
A	93	44.68	25.02
B	147	20.83	4.97

Table 1. Task 1 Results

(3) Regularity of echolocation biosonar emissions

It appears from the results that the clicks are not emitted at a constant frequency. Rather, it seems that there are two types of emissions that can be observed in each dolphin. Dolphin B appears to emitting biosonar pulses at a relatively constant frequency, although not programmatically so as would be found in a technological implementation, as observed by a relatively low standard deviation of inter-click interval. However, dolphin A appears to be emitting using a different strategy. In this case, the emission appears to be sort of bimodal, with some rapid successive clicks in a row, and then a longer gap between these 'packets' of clicks.

TASK 2

Methods

Differential time of arrival estimation

One simple option to estimate the differential time of arrival (DTOA) is to associate clicks between hydrophones and manually compute the difference in arrival time of the clicks, using the physical properties of the hydrophone array to determine the direction of arrival of a sound source. However, there are challenges with using this method, mainly that the signals are noisy, and therefore the same number of clicks may not be found on each channel, which introduces difficulty in accurately associating the arrival time of the same click across different microphones. For this reason, we have utilized the phase-transform of the generalized cross correlation (GCC-PHAT) to estimate the time difference of arrival, an approach which estimates the time lag between two signals using the normalized cross-power spectrum of the two signals (Knapp and Carter, 1976). Since the normalization discards magnitude information, the phase information is instead emphasized and therefore it is more robust to local noise and the frequency varying attenuation that can arise from increased propagation distance through the medium. Thus, the cross correlation between two signals, x and y, is estimated as in Equation 2. From this, the lag would be the point of maximum correlation as shown in Equation 3.

$$\hat{R}_{xy}(t) = \int_{-\infty}^{\infty} \frac{G_{xy}(f)}{|G_{xy}(f)|} e^{j2\pi ft} df \tag{2}$$

$$\tau = argmax(\hat{R}_{xy}(t)) \tag{3}$$

Of course, since it's possible that the dolphin's are moving, and therefore that the differential time of arrival between clicks isn't constant, it isn't sufficient to simply apply the GCC-PHAT algorithm to the entire signal duration. For this reason, we utilize the previously detected clicks to define windows around which the cross-correlation should be performed. To determine the size of this window, the maximum plausible differential time of arrival (i.e., that of the end-fire case), was considered. We know that it isn't possible for the same signal to arrive at H2 more than $\frac{14m}{1520ms^{-1}} =\sim 9.2$ ms after it arrived at H1 or H3, and the same is true reciprocally. Therefore, we used a window size equal to just over double the maximum possible time difference. The difference is doubled since the window should be symmetric around each click, given that a click could arrive from either direction.

Rather than implementing the GCC on the frame-by-frame power signals, we instead opt to implement GCC on the non-buffered hydrophone signals. This increases the accuracy and resolution of the DTOA estimation. Since the phase information of the signals is crucial to the performance of the GCC-PHAT algorithm, it's important that the unfiltered hydrophone signals be used for this step, although the clicks that were detected on the filtered and buffered hydrophone signals were still used to determine the center points of the windows within which the GCC was performed.

Since the DTOA estimation is set up such that the current click of interest for H2 will always be at the center point of the H2 signal passed to the GCC algorithm, we opted to apply a Hanning window to this signal (and only this signal) to reduce edge effects and decrease the likelihood that another pair of associated clicks is more highly correlated. Figure 6 illustrates one such implementation of this method, where the H2 signal has had a Hanning window applied, but H1 and H3 have not. The cross-correlation would be found between H1 and H2, and also between H2 and H3, and the red bars indicate the computed time lags. Note, again, that these waveforms are from the raw hydrophone data that have been segmented into uniform windows around the clicks detected on H2.

Bearing estimation

The direction of arrival can be estimated by coordinating the DTOA estimate with the geometry of the system using the relationship expressed in Equation 4, where τ is the lag found in Equation 3, c is the speed of sound in water, and d is the distance between hydrophones. This operation was performed element wise on the estimated lags per peak. Due to some noise in the results the computed lags were first filtered using a 10 point moving average, and then Equation 4 was applied. Results are shown in Figure 7 for dolphin A, and Figure 8 for dolphin B, and the statistics of the results in Table 2. Note that there are

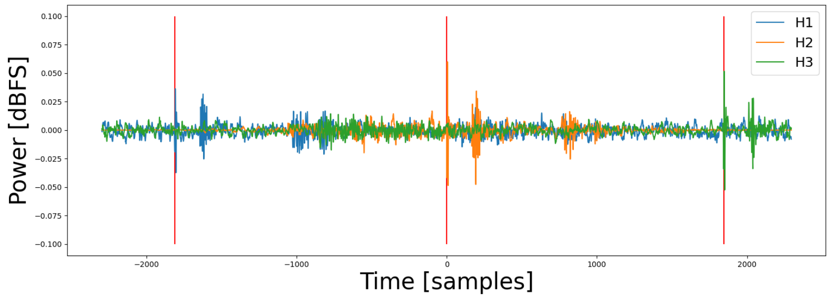

Figure 6. An example of the signals passed to the GCC-PHAT algorithm. The red bar in the middle indicates the location of the click on H2. Here, H1 and H2 would be passed together, and the red bar on the left indicates the point the algorithm identified as the time lag between H1 and H2. Similarly, H2 and H3 would be passed together, and the red bar on the right indicates the computed time lag. Visually, it can be seen that the red bars are associated with the patterns in the waveform that correspond with each other.

less clicks here than earlier, as the previously mentioned click association took place and therefore clicks without correlates in the other channels were removed.

Although the precision of the bearing estimate is related to the uncertainty of the DOA estimate, it is also a function of the bearing value. This is because passing the computed delays as an argument to *arcsin* applies some scaling, which results in a higher resolution for bearing values closer to 0 and a lower resolution for values closer to $\frac{\pi}{2}$. One can find the sensitivity of the bearing to the time lag by differentiating Equation 4 with respect to the time lag, the result of which is shown in Equation 5. Note that this result is not the direct differentiation of Equation 4, but rather the derivative before solving for θ directly (i.e., the derivative of the expression containing $\sin(\theta)$). The uncertainty of this measure would then just be the product of the sensitivity of the bearing value to the time lag and the uncertainty of the time measurement, as found in task 1. Thus, the uncertainty is expressed in Equation 6.

$$\theta = \arcsin\left(\frac{\tau c}{d}\right) \tag{4}$$

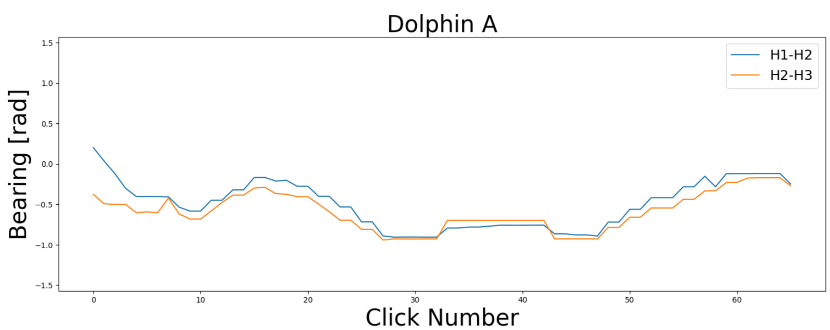

Figure 7. Plot of bearing estimates by click number for dolphin A.

$$\frac{d\theta}{d\tau} = \frac{c}{d\cos\theta} \tag{5}$$

$$Uncertainty = (\pm 16\mu s) * \frac{c}{d\cos\theta} \tag{6}$$

Range estimation

To estimate the range of a source, we utilized a geometric model in which a ray is projected from the center point between each pair of hydrophones, where the slope of the ray from the horizontal is equal to $\frac{1}{\tan(\theta)}$,

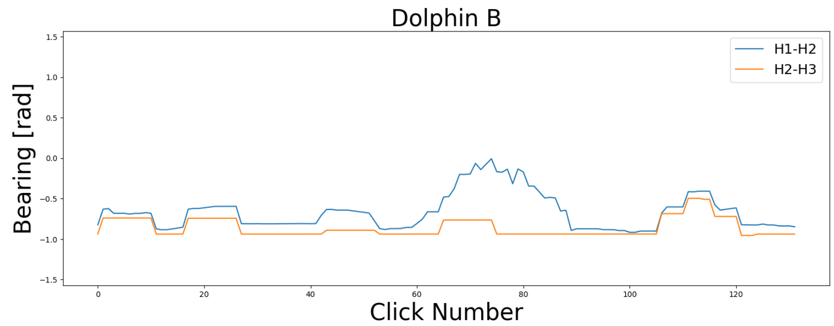

Figure 8. Plot of bearing estimates by click number for dolphin B.

where θ is the bearing estimate, and the intercept of the line is set based on the position of the center point between the two hydrophones. From this point, we can use numerical methods to identify the intersection point and find the magnitude thereof, which is the estimate of the range. Essentially, by obtaining sampled line segments for the ray projected from the center points of the two pairs of hydrophones, we just need to find the position at which the difference between the y-values for the line segments are minimized. This identifies the intersection point of the two lines. To then estimate the range, one just needs to apply the Pythagorean theorem to the intersection point, resulting in the straight line distance from the reference position (i.e., the position of H2), which is the range. The results are plotted in Figures 9 and 10 for dolphins A and B, respectively. Statistical parameters are summarized in Table 2. A demonstration of the ray intersection approach is visualized in Figure 11.

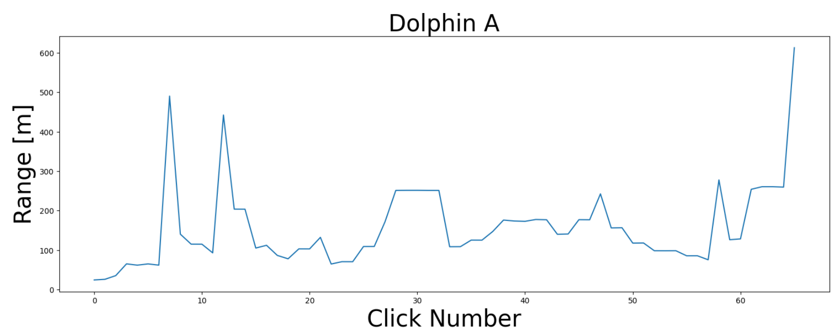

Figure 9. Plot of range estimates by click number for dolphin A.

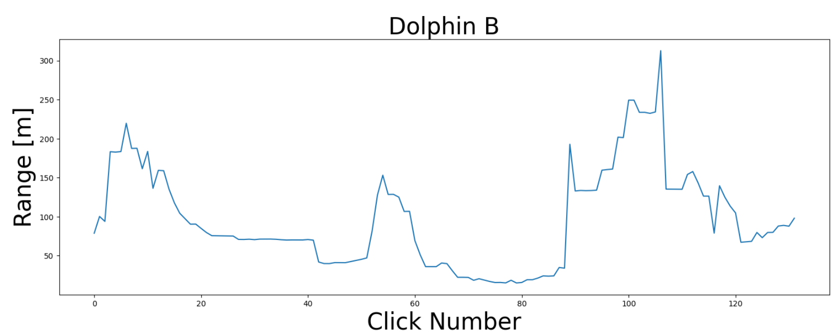

Figure 10. Plot of range estimates by click number for dolphin B.

Discussion

From these results, it is clear that the methods implemented here are not properly identifying the bearing and range, as the intersection points of the projected rays intersect beyond the sea surface, which is

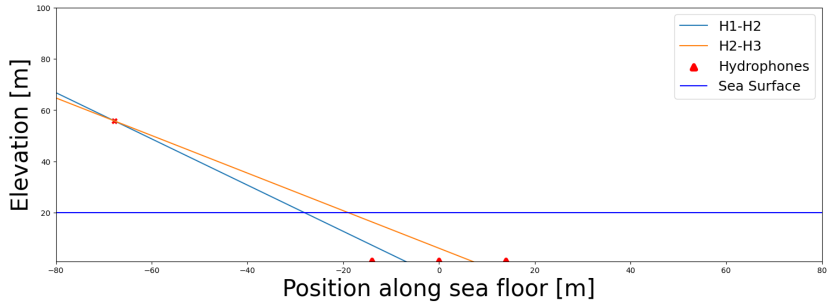

Figure 11. Sample demonstration of the geometric model used to estimate range.

Dolphin	Mean Bearing [rad]	Std. Dev Bearing	Mean Range [m]	Std. Dev. Range
A	-0.5	0.29	157.02	102.83
B	-0.66	0.22	96.92	62.55

Table 2. Task 2 Results

clearly not possible. All results up to Figure 6 seem to be working as expected, though, so it is difficult to determine what went wrong. In general, the methods applied seemed to properly identify the cross correlation between the two signals, but still resulted in implausible findings. This stimulus is relatively difficult to work with given the very short duration of the acoustic signal of interested, and the high amount of noise present on the channels.

All code used in this project can be found at https://github.com/bpmasters/asa-challenge-24

REFERENCES

Ivanov, M. P. (2004). Dolphin's echolocation signals in a complicated acoustic environment. *Acoustical Physics*, 50(4):469–479.

Knapp, C. and Carter, G. (1976). The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(4):320–327.

Ridgway, S. H., Dibble, D. S., and Kennemer, J. A. (2018). Timing and context of dolphin clicks during and after mine simulator detection and marking in the open ocean. *Biology Open*, 7(2):bio031625.