Abstract—In this paper a study of the Cross-Correlation (CC) based, Time Difference of Arrival (TDOA) method for leak source localization on fluid-filled, metallic pipelines is presented. Several, frequency domain CC estimators (General Cross-Correlators or GCC) are considered for the implementation of the localization technique, while real leakage, vibro-acoustic data series, captured from a real workplace building’s tap water pipe installation are utilized within the current analysis. Certain metrics like the rate of acceptable leak positioning results against all localization estimation experiments conducted, as well as the Relative Localization Error (LRE) of the accepted estimations are calculated and illustrated in order to shed light on the performance characteristics of the TDOA approach (in general and also per adopted CC estimator). The above study is pursued within the framework of project ESTHISIS, with end target the development of a smart sensor system for leakage detection in pipes carrying hydrocarbon products.

Index Terms—Leak source localization, TDOA, Generalized Cross-Correlation, PHAT, Hassab-Boucher, Weiner, SCOT.

I. INTRODUCTION

The capability for accurate and reliable localization of leak events on pipelines carrying fluids is of notable importance for the affiliated industries, since such events may lead to significant financial losses, environmental disasters and human health hazards (especially in the case of the hydrocarbon treating industry). Therefore, several approaches have been studied (and even adopted) in the past towards this direction [1]–[5].

A popular approach, that attracts the interest of the research community is one that is based on the estimation of the differential time delay of the leak source signal exhibited when detected by two consecutively placed sensors (of a larger sensor array) via a typical Time Difference of Arrival (TDOA) technique. The differential time delay is usually estimated using Cross-Correlation (CC) based methods as they are better suited for the problem at hand (compared with time-of-flight solutions) [6]–[10].

Exploiting the vibro-acoustic signal emanating from the fault (leak) location and propagating along a rectilinear, pipe-fluid system for leak localization purposes using the aforementioned approach, is an elegant and simple to implement solution. This is due to the fact that acoustic data can be easily captured using accelerometers mounted on the surface of the pipe, which are easy to install, with practically no intrusion, both on new and existing pipe installations, while at the same time are considered low-cost solutions (e.g. compared with ultrasound based applications). Reliable knowledge of the velocity of propagation of the leak signal, as well as clear determination of the usable spectral area of the captured signals (due to multi-modal propagation effects) are essential in order to reach an accurate localization result using the estimated time delay.

In earlier studies on the leakage detection and localization area (under project ESTHISIS [11]) we have (among others) tackled the issues of robust leak vibration velocity estimation, combined with a review of the spectral characteristics of the leak signal. These studies also included initial assessment of the performance of certain CC estimators (mainly on velocity estimation experiments and preliminary evaluation testing using audible sounds in the air interface) [12]–[16].

The findings of the aforementioned work are leveraged in the current study in an effort to additionally explore the performance of actual leak localization implementations through the CC-based, TDOA approach. For the required time delay estimation a series of frequency domain Generalized Cross-Correlators (GCC) [17] is utilized, that also allows for a comparative study of these correlators on the basis of the quality of the results reached (from the use of their output). In terms of the performance metrics adopted, the (success) rate of acceptable leak positioning results (after eliminating outliers and non-sensible results) against all localization estimation experiments conducted is measured, as well as the Relative Localization Error (LRE) of the accepted estimations (using as reference the actual leak position, which is known in
our experimental setups). The relevant analysis is facilitated through the use of leak data captured from a tap water pipe installation, that was utilized in order to simulate leak events in a typical workplace environment (a main, linear, metallic pipeline of a university building in this case).

So, in the sections and paragraphs that follow, an outline of the TDOA method for leak localization is laid out at first, followed by a brief, theoretical presentation of the GCC estimators used in this work for the time delay estimation. Next, the testing and measurement installation and procedure for leak signal acquisition is described in detail, followed by the stepwise procedure of data pre-processing and handling as well as the calculation of the performance criteria. Subsequently, the results of the testing and analysis process are illustrated and discussed and the work concludes with a short summary.

II. TDOA MODEL FOR LEAK LOCALIZATION

A leak localization configuration is diagrammatically depicted in Fig. 1. As shown in this figure, to sensors (accelerometers) $S_1$ and $S_2$ are installed on a pipeline segment, upstream and downstream the leak position $F$ (this corresponds to a so-called ‘in-bracket’ leak). The distance $L$ between the two consecutive sensors (the two sensors may be part of much longer sensor array spanning the total length of a pipeline) is known (aspirational target, $L \approx 100m$). Pinpointing of the leak source location corresponds to the determination of either of the unknown distances $x$ or $L - x$. To this end, an accurate estimation of the (differential) time delay that the leak, vibro-acoustic signal exhibits when arriving at the sensors due to the different travel distance within the pipe-fluid system can be exploited, under the condition that the speed of signal propagation is known beforehand. So, if $c$ denotes the speed of the acoustic wave and $T_D = t_1 - t_2$ denotes the estimated TDOA between the signals captured by sensors $S_1$ and $S_2$, then the sought out distance $x$ can be estimated as

$$x = \frac{1}{2} (L - cT_D).$$

III. CC ASSISTED TIME DELAY ESTIMATION

The mathematical interpretation of the signals $x_1(t)$ and $x_2(t)$ recorded by sensors $S_1$ and $S_2$ respectively may stated as [17]

$$x_1(t) = a_1 s(t - t_1) + n_1(t)$$

$$x_2(t) = a_2 s(t - t_2) + n_2(t)$$

where $s(t)$ is the leakage acoustic signal at location $F$, and $a_1$ and $a_2$ are gain factors at the sensors. The (corresponding) time delays at each sensor are denoted by $t_1$ and $t_2$, while $n_1(t)$ and $n_2(t)$ account for detrimental noise contributions. Under the assumption that signals $s(t)$, $n_1(t)$ and $n_2(t)$ and the corresponding random processes meet the required non-dependency, stationarity and ergodicity criteria, the CC function of $x_1(t)$ and $x_2(t)$ can be defined as

$$R_{x_1x_2}(\tau) \triangleq \mathbb{E} (x_1(t)x_2(t - \tau))$$

where $\mathbb{E} (\cdot)$ denotes the expectation operator. The estimation of the (differential) time delay $T_D$ that corresponds to the TDOA estimate due to unequal travel time of the leak signal towards the sensors, amounts to pinpointing the time lag $\tau$ that maximizes the CC, that is $T_D = \arg\max_{\tau} R_{x_1x_2}(\tau)$.

IV. FREQUENCY DOMAIN CC ESTIMATORS - GCC

Working in the frequency domain, the CC function of (4) may be estimated via the formula that interconnects it with the Cross-Spectral Density (CSD) $S_{x_1x_2}(\omega)$ of the signals at hand (designated FD-CC) as

$$R_{x_1x_2}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{x_1x_2}(\omega)e^{j\omega\tau} d\omega$$

were $\omega$ is the radial frequency. Although the estimate achieved through the FD-CC function is practically equivalent to the estimate reached through its straightforward, time domain (TD-CC) counterpart, the FD-CC allows for further refinement of the estimation process. This is made possible by introducing a weighting function $G(\omega)$ to be applied on the CSD $S_{x_1x_2}$, (this may indicate a pre-filtering process). The resulting formula corresponds to the Generalized Cross-Correlator (GCC) and is defined as

$$\hat{R}_{x_1x_2}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\omega)S_{x_1x_2}(\omega)e^{j\omega\tau} d\omega.$$ 

Different choices for the formulation of the weight function $G(\omega)$, give birth to diverse estimators. For the purposes of the current analysis (apart from the standard FD-CC estimator) the following estimators are considered:

- Smoothed Coherence Transform (SCOT)
- Wiener
- Maximum Likelihood (ML)
- Phase Transform (PHAT)
- Hassab-Boucher (HB).

Comprehensive description of the above GCC estimators can be accessed through extensive bibliography (e.g. [17]–[21]), while initial usage and testing on leak localization related implementations may be accessed in [12], [13].

Finally, it is noted that, for the aforementioned estimators, discrete signal processing on the basis of the available (finite horizon) data-series is employed, while for the estimation of the CSD $S_{x_1x_2}(\omega)$ and Power Spectral Densities (PSD) $S_{x_1}(\omega)$ and $S_{x_2}(\omega)$ (that the weight functions usually incorporate) the non-parametric Welch algorithm is adopted [22]–[24].
V. DATA CAPTURE AND PERFORMANCE ASSESSMENT PROCEDURE

A. Signal acquisition - installation and measurement protocol

The testing and measurement campaign undertaken, in order to collect a sizable amount of experimental leak data (vibro-acoustic data-series), was focused on a rectilinear, metallic pipeline of the water supply system of a university building. The pipeline used, is part of the tap water distribution system and is approximately 30m long, with 1.5/2 inch gauge (48.3mm fixed external diameter, 1.5mm wall thickness) and made of carbon steel. The water pressure in the pipeline was kept relatively stable during the simulated leaks and the measurements (around 5.3bar), facilitated by the existence of an 1500lit pressure vessel.

The sensors used were two PCB J352B, low-noise, high resolution, ground isolated, ceramic shear ICP accelerometers, that fashion an 1000mV/g sensitivity figure and a frequency range (±5%) of 2 – 10000Hz. Dual-rail magnetic adapters, suitable for curved ferromagnetic surfaces, firmly bolted to the accelerometers via a stud, were used for the sensor mounting on the pipe surface (see Fig. 2 (left)) enabling quick relocation. The signal capture was performed by an Analog-to-Digital Converter (ADC) device and more specifically, the PCB 485B39, 2-channel, ICP sensor signal conditioner, which can achieve 24-bit sample resolution with frequency range (±5%) extending from 0.8 to 20700Hz and output protocol that corresponds to USB class 1 audio. The signal recording from the ADC and the succeeding analysis was enabled via typical software packages like MatLab.

Regarding the measurement procedure, the sensors were mounted on diverse positions on the pipeline (several configurations of both the distance between the sensors and their actual position on the pipe), while through the use of existing ‘Tee’ connectors feeding water to taps or test valves (see Fig. 2 (right)), the simulation of a leak was enabled (by opening the relevant tap or valve). Moreover, the simulation of leaks of different severity was possible in terms of measured Signal-to-Noise Ratio (SNR) at the leak location and leak flux. Recordings of 3min were captured for each configuration of leak and sensor position. The ADC was set to 24-bit resolution and 44.1kHz sampling frequency for all measurements. A diverse inventory of measurements was built during this experimentation campaign, with the accumulated signal recording time reaching nearly 6 hours.

B. Performance assessment road-map

For the the evaluation of the performance of the CC-TDOA based leak localization approach 14 different leakage experiments were selected, each one corresponding to a specific configuration in terms of leak and sensors’ position and leak severity. Table I outlines the leak experiments utilized where each leak is designated as Small (S), Medium (M) or Large (L) corresponding to leaks with 2dB, 5dB and 15dB of measured SNR respectively (0.5 up to 12.0 lit/m flux). The distances L and m indicated for each experiment are in alignment with a configuration as the one depicted in Fig. 1.

The procedure followed for the evaluation of the method and the estimators, starting with initial processing actions and culminating in the performance assessment via the adopted criteria is laid out subsequently, in the form of consecutive steps (performance metrics definitions and intermediate results and observations are also included).

1) Pre-processing: The initial data handling considerations take into account that the accelerometer mounting technique (magnetic adapter) is known to reduce the resonant frequency of the sensor considerably (the manufacturer, indicates an upper bound of 2kHz for flat frequency response). Additionally, the fundamental mode of the vibro-acoustic wave which propagates in the pipe-fluid system at longer distances (and with lower attenuation) compared to higher modes and is non-dispersive (and therefore can be assigned with a frequency independent propagation speed) appears in low frequency regime [25]–[27], with usable spectral area up to around 1100Hz (for the specifications of the monitored pipeline). Moreover, at very low frequencies (up to a few tens of cycles) detrimental noise contributions appear (e.g. power grid frequency interference, rotary pump harmonics and other), that degrade the integrity of the recordings.

Based on the above remarks, the captured data are initially down-sampled (after appropriate anti-aliasing filtering)
to a lower rate of around 5.5kHz, bringing the upper usable frequency to around 2.75kHz. Further filtering is performed in order to isolate the spectral content within the desired 200 – 1100Hz area.

2) Segmentation: Every leak recording (full) segment that typically has a 3min duration is segmented into 180, non-overlapping subsegments of 1sec duration. In this way, a large pool of leak ‘micro-experiments’ is built for each of the 14 leak configurations examined (180 for each recording).

3) CC estimation: A GCC estimator is applied to each individual data subsegment (for each experimental configuration) and a set of time index lags that correspond to the maximum value of the CC obtained by the GCC estimator are computed as \( N_i/F_s \), where \( N_i \) represents the corresponding discrete lag indexes, \( F_s \) is the current sample rate and \( i = 1, 2, \ldots, 180 \). Obviously, due to the physical constrains of the problem at hand, \( N_i \leq N_{max} = \min(Abs(F_s \frac{c}{L})) \), where \( c \) is an estimate of the sound propagation velocity. It is noted that this step and the ones that follow, are executed for all six of the GCC estimators considered in this work.

4) Success rate calculation - outlier elimination: From the previously estimated lag values, outlier values that correspond to \( N_i > N_{max} \) are discarded, as well as time index lags that correspond to values of \( N_i \) of negative polarity compared to the assumed position of the leak with respect to the middle of the pipe, established either by the use of the rule of the majority, or by comparison of the relative power of the sound propagating signals by the two sensors. Next, a performance indicator called hereafter the Success Rate (SR) is defined as

\[
SR = \frac{S_{nominal}}{180} \times 100
\]

where \( S_{nominal} \) is the cardinality of the set of nominal values of \( N_i \), after outliers are eliminated. This is illustrated in Fig. 3, for the first leak of Table I, for the FD-CC estimator, where \( SR = 65\% \), meaning that 117 out of 180 estimated values of the time index lag are accepted using the aforementioned selection criteria.

5) Leak source localization: Based on the set of accepted time index lag values, a mean value of it (\( \bar{N} \)) is computed and the localization estimation is obtained by the use of (1) as \( \hat{x} = \frac{1}{L} \left( L - c \frac{\bar{N}}{F_s} \right) \) where \( L \) is the pipe length and \( c \) is an estimation of the vibro-acoustic signal propagation velocity, which - based on our previous work [12] - is set equal to \( c = 1045m/sec \).

6) Localization performance evaluation: The performance of the various GCC methods used is evaluated via the Relative Localization Error (RLE) defined as

\[
RLE = \frac{x - \hat{x}}{L} \times 100
\]

where \( x \) is the true leak location and \( \hat{x} \) is the estimated leak location (from the previous step of the process). Alternatively, the absolute value of RLE (ARLE) can be used for the same purpose.

VI. RESULTS AND DISCUSSION

A first insight into the performance of the GCC estimators for leak source localization (utilizing the road-map described earlier) is obtained through the localization results summarized in the compound Table II, where RLE and SR figures are also included, for the baseline FD-CC and the more refined HB estimators. It is evident that even though the FD-CC comes with acceptably low estimation error, the HB estimator is superior in terms of low RLE and high SR as well as robustness in all measurement conditions.
Further illustration of the performance of the examined estimators is available in Fig. 4 where the SR for all methods is presented (bar chart) along with the localization results (blue, box-whisker type plot indicates the average estimation of the leak location along with the the standard deviation margins) and the actual leak location as reference (red encircled cross). A quite similar performance for the SCOT, PHAT and HB estimators can be discerned from this graphical representation.

Corroborating conclusions are drawn from the depictions of Fig. 5 where the RLE for all methods and all leak configurations is presented. The coherent performance of the three aforementioned estimators is evident, with all three yielding results with somewhat lower RLE compared to the rest of the GCC methods and the baseline FD-CC.

Fig. 4. Localization results (average estimation and std margins) and success rate for six different GCC estimators.

VII. CONCLUSIONS

In this paper, an experimental testing of the CC based, TDOA approach for leak source localization was presented, though the utilization of a set of CC estimators, that adhere to the generalized CC approach. Real leakage data recordings were used for the evaluation, while metrics like the success rate and the relative localization error were adopted for the quantification of the results and the derivation of comparative conclusions. As is the case for similar applications like leak signal speed of propagation estimation, the evaluation process highlighted the validity of the TDOA approach in general and the superior performance of the weighted, frequency domain, CC estimators. Further study with diverse leak configurations in terms of utilized pipelines, distances and fluids is to follow.
Fig. 6. Absolute relative localization error bar chart, for six different GCC estimators and all experimental configurations (along with mean ARLE and SR values).

REFERENCES


