

Convolutional Filtering for Accurate Signal Timing from Noisy Streaming Data

Jonathan Wang^{*†}, Kesheng Wu^{*}, Alex Sim^{*}, Seongwook Hwangbo[‡]

^{*}Lawrence Berkeley National Laboratory

[†]Rice University

[‡]Hyundai Electric & Energy Systems Co., Ltd.

Emails: [†]jonathanwang017@berkeley.edu, ^{*}{kwu, asim}@lbl.gov, [‡]smhwangbo@hyundai-electric.com

Abstract—Our society depends heavily on the electric power infrastructure. To ensure its reliability, key power grid components such as transformers are extensively monitored for signs of failures and errors. This work concentrates on a type of event known as the partial discharge (PD) because it is the symptom of insulation weakness, the most common cause of transformer failures. More specifically, our work is to locate the position of a partial discharge to provide information for preventive maintenance. Our method utilizes the information from a set of ultra-high frequency (UHF) sensors inside the transformer, and proceeds in two steps: first determine the signal arrival time and then locate the position based on time differences. To determine the arrival time, we develop a convolutional filtering method based on the Savitzky-Golay filter. To provide accurate locations, we simulate the electromagnetic wave propagation using finite-difference time-domain (FDTD) to generate a reference table of the travel time from each FDTD mesh point to the sensors.

We exercise our method using two sets of UHF measurements with different signal to noise ratios. In both cases, our method provides more accurate locations than other methods. The difference is particularly prominent when the signal is weak. With weak signals, the best existing method, the cumulative energy method, was only able to predict the PD location within 300 mm of the known sources in 13% of the test cases, while our method is correct in 48% of the test cases.

I. INTRODUCTION

Power transformers are the most important and expensive elements of a nation's electric power infrastructure. Although the reliability of transformers has increased over time, when a transformer does fail, it can have severe consequence for the power grid users and operators. For example, users may suffer days or even weeks of power outage, while the grid operator faces significant financial penalties. Extensive online diagnostic systems have been developed to detect and prevent such failures [1]. This work focuses on analyzing data produced from one of these diagnostic systems to locate an early warning sign of a common fault.

Transformer failure is usually triggered by external events such as lightning strikes, switching transients, and short-circuits. However, susceptibility to these events is usually due to some incipient weakness in the insulation system. The transformer insulation strength can degrade to the point that it cannot withstand system events such as short-circuit faults or transient over-voltages [2], which leads to a type of internal arcing event known as partial discharge (PD). Figure 1 illustrates the causes of failure in transformers in Korea over the

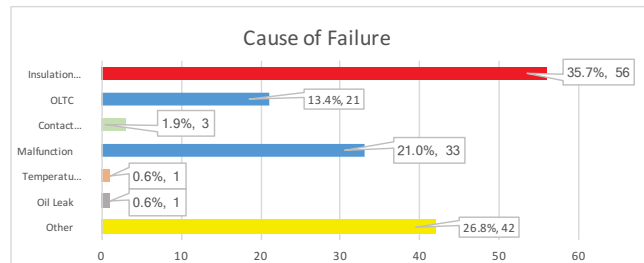


Fig. 1. Failure Occurrence of Transformers Over Last 10 years

last 10 years [3]. Insulation failure is the most frequent cause at 35.7%. Detecting and locating PD events would allow owners and operators of transformers to take appropriate actions to avoid unexpected failures. Such preventive measures could also protect other equipment connected to the transformers, such as Gas Insulation Switchgear (GIS) and switchboards in the substation, which are also expensive components of the electric power grid.

PD occurs within transformers where the electric field exceeds the local dielectric strength of the insulation. Although the PD may initially be quite small, if left unchecked, the damaged area can grow, eventually leading to electrical breakdown. This has motivated the development of many preventive diagnosis and monitoring tools to assess the internal condition of a transformer [4]. One of the best known monitoring techniques is Dissolved Gas Analysis (DGA), which is routinely employed to detect internal electrical discharge in power transformers. DGA can provide some information about the nature and severity of the PD, but not the location needed for remedial actions [5].

Several methods have been developed for PD localization, including transformer winding modeling, acoustic methods, and ultra-high frequency (UHF) sensors. The transformer winding models are sensitive to weak PD activities, but they can only predict the turn number in which the PD occurs and perform 1D localization [6], [7]. Using the acoustic method, 3D localization is possible; however, this method has less sensitivity to weak PDs and those that occur inside the winding [8]. The UHF sensors are sensitive to weak PDs both inside and outside of the winding. They can also predict the PD

location in 3D coordinates [9]. These sensors can detect UHF electromagnetic (EM) waves of varying frequencies (0.3 GHz - 3.0 GHz). Additionally, these UHF sensors can be installed inside transformer walls, which act as a Faraday cage to protect the UHF sensors from external noise [10]. Thus, UHF sensors are considered the most effective at locating PD positions.

Due to the different characteristics of the various materials used in the transformer, such as the dielectric constant and losses, EM waves will arrive at the sensors with different time delays and levels of attenuation. This creates significant challenges in locating the source of a PD event. Like most previous approaches, we use the time difference of arrival (TDOA) to locate the PD source. However, we have a more effective way to determine the signal arrival time and have a location method that takes into account the details of the transformer involved.

When a signal first arrives at a sensor, it is buried among the background noise. There are a number of different approaches that have been published [11]; however, we believe they are not robust enough. To more accurately determine the arrival time, we studied a number of different approaches and developed a new filter-based approach that works much better than others.

Because a transformer has a large amount of metal inside, the electromagnetic waves travels in complex paths. To accurately locate the source of a PD event, we decide to use a reference table of travel times generated from detailed finite-difference time-domain (FDTD) simulations. We match observed time differences with the simulated values in the reference table to determine the PD location. This approach requires a reference table for each model of transformer. However, since there are relatively small number of different models, the effort required to generate these reference tables should be modest compared to the potential savings from prevention of transformer failures.

We exercise the methods using a set of sensor measurements from lab-generated PD events, where the general locations of the PD sources are known. This allows us to test the algorithms on real measurement data, whereas most of the previous studies used only synthetic data. Tests confirm that our method for determining the arrival time is indeed more accurate, and the resulting PD locations are also more accurate.

In the remainder of this paper, we present related works in Section II, and describe methods used for this study in Section III. We go over experimental results in Section IV, and discuss parameters and improvements for these methods in Section V. The conclusion and future work are in Section VI.

II. RELATED WORKS

This work focuses on locating the position of a PD source using UHF sensor data. We briefly describe the UHF sensors, the work on determining the arrival time from sensor data, and an FDTD method for generating the reference table of arrival times used to determine the PD locations. We also use this opportunity to accentuate the distinctiveness of our approach.

The UHF PD detection technique has better accuracy due to lower signal damping as well as lower possibility of

external interference since the sensors are shielded by the transformer walls. UHF sensors have also been shown to be capable of detecting PDs in transformer oil under DC voltage conditions [12] and locating PD sources inside transformer windings [13]. However, the challenge of PD detection circuits is the design of signal interpretation circuits and proper sensors (antennas) capable of signal detection in the broadband range. There are many types of UHF sensors that can detect electromagnetic waves radiated from PD sources. The general types of UHF PD sensors are disc and spiral, monopole and conical types [14].

Accurately locating the source of PDs using UHF sensors is an active research topic [15]. A key step in this work is to calculate the arrival time of signal at each sensor. The state of art in determining the arrival time include the cumulative energy method [16] and the energy criterion [11]. In this work, we develop a method using the Savitzky-Golay filter to accentuate the signal and a threshold to detect the signal arrival. Tests show that our method is able to better preserve features.

The Savitzky-Golay filter has been used for saturation detection [17]. In our case, we use it to handle noisy signals to improve detection robustness. In addition, we only use the Savitzky-Golay filter to find a rough signal arrival, which we then refine using a moving average since the resulting window is not as sensitive to smoothing. Our study presents an improved signal timing method coupled with a refined lookup method using FDTD in order to localize PDs.

As in previous methods, we use the time difference of arrival (TDOA) to locate the PD source [15], [18]. However, our PD localization method takes into account the actual structure of the transformer by using a reference table generated from FDTD simulations. We select the FDTD mesh point corresponding to the time differences with the minimum Euclidean distance from the time differences of our signal timings.

The FDTD technique is the state of the art method for modeling the propagation of electromagnetic waves inside transformers. In 1966, Yee [19] proposed a technique to solve Maxwell's curl equations using the FDTD technique. It is specifically designed to model the propagation of electromagnetic transients and their interactions with complex metallic structures inside transformers. The FDTD formulation is a convenient method for solving electromagnetic field problems. FDTD, which is widely applied to the field of electromagnetic computation, can be used to simulate the electric and magnetic fields that are measured by sensors. For example, the modeling of PDs is outlined in [20], while the modeling of a UHF spiral sensor using the FDTD method was proposed in [21]. The intensity of electric fields generated by PD sources in dielectric materials was investigated using FDTD modeling in [22]. FDTD simulation has also been applied to study the propagation characteristics of electromagnetic waves caused by PD in GIS [23]. FDTD computes the voltage at each sensor at each time step. The arrival time at a sensor is calculated using energy criterion [11], the best known timing method. In this study, FDTD simulations are used to generate a table

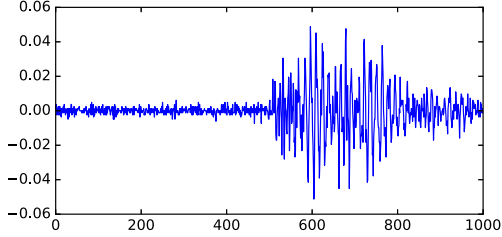


Fig. 2. Raw Data from PD2 Channel 1

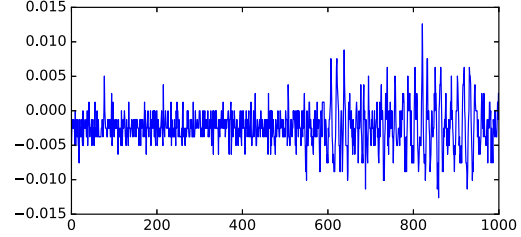


Fig. 3. Raw Data from PD2 Channel 4

of arrival time at the sensors, where each row of the table corresponds to a PD event at a given mesh point. We use the differences in arrival time at the sensors to look up the source position of a PD event. Note that we could improve both spatial and temporal resolutions of FDTD simulations to create a reference table with more accurate arrival time values. In this work, due to runtime constraints, the FDTD mesh was calculated with 300 mm between neighboring points. The timestep sizes used by FDTD are about 0.006 - 0.009 ns, which are small enough to model 3GHz UHF waves.

III. METHODS

The goal of this study is to use the signals from the four sensors in the transformer to determine the source of the partial discharge (PD). This process is broken up into two steps:

- Identify the arrival time of the signals in the sensor data
- Localize the partial discharge using the signal timings

We use sample data collected from a test transformer with two possible PD locations, designated PD1 and PD2. Within our test transformer, partial discharges are triggered in laboratory conditions and signals were collected from these triggered events. At each PD location, there are 26 possible specific failure points. There are 22 signal samples for PD1 and 23 samples for PD2. Each sample includes four sets of 1000 data points, with one set from each of four sensors. A **channel** is the data for one sensor, and a **signal** refers to the set of four **channels**. Each data point represents the voltage detected by the sensor in increments of 0.4 ns. Using these two sets of data, we try to locate the coordinates of PD1 and PD2.

In addition to the emulated PD data, we also validate our method with simulated FDTD data based on two other transformers, designated 3MVA and TL1965. This data simulates a PD occurring at each of the FDTD indices and records the voltage detected at each sensor over a time period of 50 ns.

Signal Timing

We examine various methods to pinpoint the arrival time of the signal in the sensor data. This is the point where the PD signal is stronger than background noise. The primary challenge is the low signal to noise ratios (SNR) present in certain sensors, since the noise obscures the arrival of the signal.

Figures 2 and 3 show the sensor data for which we identify the arrival of the signal. Figure 2 is an example of data with a high SNR, making it simple for any timing method to find the signal. This data, designated channel 1 (CH1), is obtained from one of the four sensors, which is fairly close to the PD source, resulting in the stronger signal. However, in Figure 3, the signal is obscured by high levels of noise. This data is channel 4 (CH4), which is obtained from a sensor further away from the PD source and likely blocked by internal structures.

In our initial experiments, we tested several methods, including a simple threshold, an envelope, analyzing the data spread, noise cancellation, moving average threshold, and frequency comparison, all of which were not consistently successful due to the effect of the high levels of noise. We also tried filtering the data first with a lowpass filter to remove noise, but it was not sufficient to improve these methods. We compared these results against the results of the existing cumulative energy method [16], which is a standard method for determining signal arrival times. We find the knee point in the cumulative energy curve using the Kneedle method [24]. We also compared these methods with the energy criterion method [11], which had almost identical results to those of the cumulative energy method.

To improve the signal timing, we use the **Savitzky-Golay filter** [25] to improve the SNR, allowing for more robust detection in noisy data [17]. The Savitzky-Golay filter (SG filter) smooths data through convolution, by fitting windows of the signal to a low degree polynomial using least squares. The solutions to the least squares equations are applied to the data subsets to estimate the smoothed signal. Compared to other filters such as a moving average, this filter preserves features better, making it easier to find the signal arrival with threshold methods. We use the Savitzky-Golay filter with a threshold to define a window for the arrival time. This threshold is determined based on the mean and standard deviation of the noisy data, which is the first half of the window.

Once we identify this initial window, which usually includes only one significant peak, we can use a moving average to further eliminate noise and refine our initial estimate to find the signal arrival time. In the original signal, the moving average would reduce the amplitude of peaks too much, making it difficult for a threshold to find the signal. However, this is not an issue within the window since the threshold only has one peak to identify.

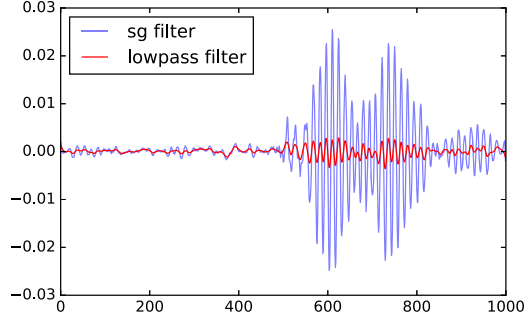


Fig. 4. SG Filtered Data from PD2 Channel 1

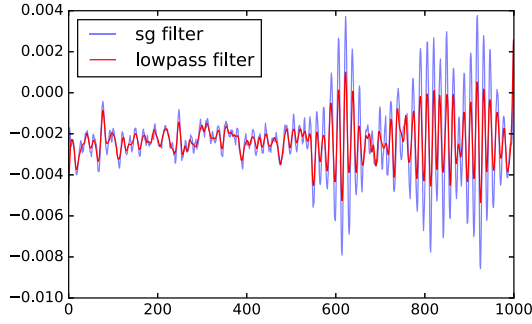


Fig. 5. SG Filtered Data from PD2 Channel 4

Figures 4 and 5 show the improved noise reduction effect of the SG filter over the lowpass filter. Even in the worst case (channel 4), the signal has double the amplitude of the noise when using the SG filter method.

Partial Discharge Localization

We explore two methods of localizing the partial discharge. The first method is multilateration, solving four quadratic equations for four variables - the x , y , and z coordinates as well as the origin time. The second method is finite-difference time-domain (FDTD), which simulates the behavior of electromagnetic signals in the transformer.

Multilateration: We define 4 quadratic equations

$$(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2 - (c_e(t - t_i))^2 = 0$$

$$c_e = c/\sqrt{1.7}$$

x , y , and z are the coordinates of the PD source, and t is the time the signal originates. c_e is the effective speed of light (c) traveling through the oil in the transformer. This value is selected based on the relative permittivity of the oil medium. For $i = 1, 2, 3, 4$, x_i , y_i , and z_i are the coordinates of the sensors, and t_i is the time that the signal reaches sensor i . t_i is found using the signal timing method, and the sensor coordinates are known. From these equations, we solve for x , y , z , and t . In our experiments, we solve for the PD source

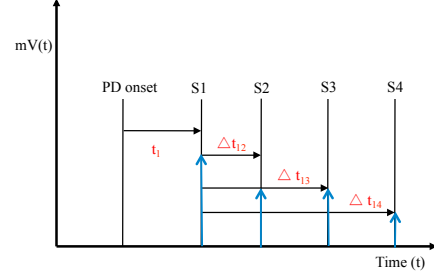


Fig. 6. Signal arrival times in reference to unknown PD onset

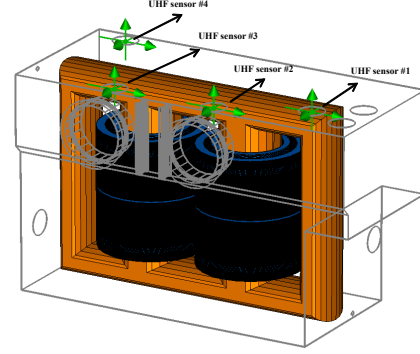


Fig. 7. Model of power transformer and position of UHF sensors

using a derivative-free spectral residual method [26]. While the localized PD is able to differentiate between PD1 and PD2, small errors in the timing greatly affect the identified coordinates. In addition, the signals do not actually travel in straight lines due to the internal structure of the transformer.

FDTD Reference Table: To account for the path of the signals, we use FDTD simulation. The transformer is divided into 300 mm mesh points, and a PD signal is simulated from each point. The time that it takes to travel from each mesh point to each of the four sensors is recorded, and using these four times, we calculate three time differences to compare our signal timings against. These time differences are the differences between three of the times and one base time as shown in Figure 6.

To simulate the time difference, we use the CST Microwave Studio software on a one-phase 750MVA power transformer with a 7000 mm x 2000 mm x 4500 mm tank. Figure 7 shows the simulation model and sensing locations for calculating the time differences. As shown in Figure 7, since the installation locations of the UHF sensors are restricted, the sensors were forced into a close arrangement. As a PD source, a dipole antenna with a standardized voltage source (IEC 60270) was placed in pre-defined coordinates.

Since sensor 2 consistently provides low noise data in our experiments, we use the sensor 2 timing as the base for the time differences. We look up time differences between sensors from the FDTD simulation, and compare our signal timings to determine the closest mesh point in the transformer to the PD

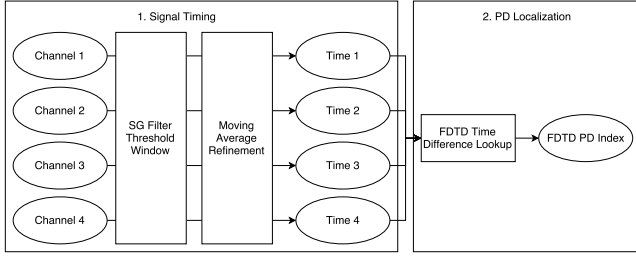


Fig. 8. PD Localization Procedure Flowchart

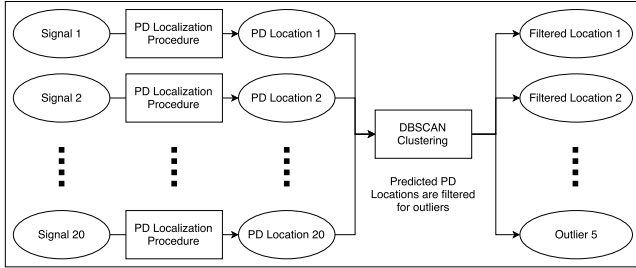


Fig. 9. Finding Outliers Procedure Flowchart

source. The FDTD mesh point whose time differences have the lowest root mean squared error (RMSE) from the time differences of the signal timings is selected as the PD source.

Figure 8 shows the entirety of the PD localization procedure, which consists of finding the timing for each channel of the signal and then using those timings to look up the nearest FDTD mesh point based on RMSE. This procedure is then repeated on each of the signals.

DBSCAN Outliers

Since we have a collection of about 20 samples for each PD, we use density-based spatial clustering of applications with noise (DBSCAN) [27] to detect outlier samples that may have some variance. DBSCAN labels points as either part of a cluster or noise based on a parameter of the number of points to define a cluster and the distance of the cluster. This algorithm clusters points without needing the number of clusters as an input. Using DBSCAN, we cluster the localized PD points, and remove any that are farther than 1000 mm from the largest cluster. This distance is selected to provide a light filter to remove extreme cases and not significantly affect the results of our predictions.

Figure 9 shows how multiple signals are handled. Each signal is used to find a PD location using its four channels. The resulting locations are filtered using DBSCAN to determine which signals are outliers. The outliers are not included in our accuracy score calculation.

IV. EXPERIMENTAL RESULTS

In our experiments, we have two partial discharge (PD) locations each with 26 possible coordinates. To quantify our PD location methods, we measure the number of those coordinates that the identified PD encompasses within a fixed

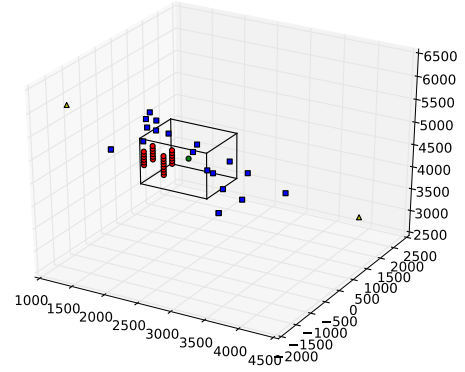


Fig. 10. Localization Results of Multilateration on PD1

error range. The error range is a cube around the designated point. We first remove **outlier** points using DBSCAN, and then find the **accuracy**, which is the percent of localized origins which contain a PD coordinate, as well as the **total hits**, which is the total number of coordinates encompassed by all of the localized origins, including outliers. A good result would have few outliers, high accuracy and total hits. This method of evaluation is selected since there is no single point to measure our prediction against.

Partial Discharge Localization Method

Our first goal is to select a localization method to find the PD source once we have the signal times. Using the standard cumulative energy method to find the signal times, we test multilateration and FDTD methods using the PD1 and PD2 samples. With an error range of 500 mm (1000 mm diameter cube around point), we compare the accuracy and total hits using the two localization methods.

The 3D plots in Figures 10 and 11 show the coordinates of the actual and predicted PDs. The yellow triangles indicate outliers, which are not included in the accuracy calculations. The blue squares indicate predicted coordinates that are within the main cluster (not outliers). The green point is the center of the blue points, and used as a reference point for the cube, which is the error range around the center point. The red circles are the actual possible PD coordinates. As Figure 10 shows, the points localized by multilateration method stretch to pretty far from the PD coordinates. About half of the points are not even within the error range.

On the other hand, using FDTD on the same dataset, there are no outliers, and all of the localized points are within the 500 mm error range, most encompassing multiple possible PD coordinates as seen in Figures 11. There are very few blue squares in these figures since many of the localized points overlap at the identified locations.

Table I shows that FDTD method has a much higher accuracy in identifying close PD sources than multilateration method. These results were generated with the cumulative energy signal timing method, which we use as a basis for

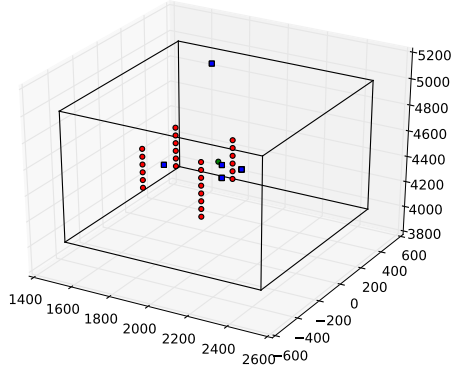


Fig. 11. Localization Results of FDTD on PD1

TABLE I
LOCALIZATION RESULTS OF FDTD AND MULTILATERATION WITH 500 MM ERROR

| Localization Method | PD | Outliers | Accuracy | Total Hits |
|---------------------|-----|----------|----------|------------|
| Multilateration | PD1 | 2 | 0.4 | 69 |
| | PD2 | 9 | 0.79 | 196 |
| FDTD | PD1 | 0 | 1 | 536 |
| | PD2 | 0 | 1 | 238 |

a comparison. The primary downside to this method is the time that it takes to generate the FDTD simulation as well as the inherent error given by the mesh size of FDTD.

Signal Timing Method

The more challenging task is identifying the signal arrival time in the sensor data with a low SNR. Comparing our method using the Savitzky-Golay filter (SG filter) with the existing cumulative energy method, we examine the localization results using FDTD with a 500 mm error range.

As seen in Table II, most of the methods have a large number of outliers and low accuracy. While some of the localized points are fairly close to the PD coordinates, they

TABLE II
LOCALIZATION RESULTS OF SIGNAL TIMING METHODS WITH 500 MM ERROR

| Timing Method | PD | Outliers | Accuracy | Total Hits |
|-----------------------|-----|----------|----------|------------|
| Threshold | PD1 | 10 | 0.42 | 84 |
| | PD2 | 2 | 1 | 376 |
| Envelope | PD1 | 10 | 0.92 | 286 |
| | PD2 | 0 | 0 | 0 |
| Data Spread | PD1 | 4 | 0.22 | 104 |
| | PD2 | 13 | 1 | 260 |
| Noise Cancellation | PD1 | 15 | 0 | 52 |
| | PD2 | 15 | 0 | 30 |
| Wavelength Comparison | PD1 | 13 | 0 | 0 |
| | PD2 | 12 | 1 | 142 |
| Moving Average | PD1 | 13 | 1 | 206 |
| | PD2 | 0 | 1 | 216 |
| Cumulative Energy | PD1 | 0 | 1 | 536 |
| | PD2 | 0 | 1 | 238 |
| SG Filter | PD1 | 0 | 1 | 554 |
| | PD2 | 0 | 1 | 382 |

TABLE III
LOCALIZATION RESULTS OF SIGNAL TIMING METHODS WITH 300 MM ERROR

| Timing Method | PD | Outliers | Accuracy | Total Hits |
|-------------------|-----|----------|----------|------------|
| Cumulative Energy | PD1 | 0 | 0.91 | 292 |
| | PD2 | 0 | 0.13 | 42 |
| SG Filter | PD1 | 0 | 0.95 | 298 |
| | PD2 | 0 | 0.48 | 154 |

TABLE IV
LOCALIZATION RESULTS OF MULTILATERATION WITH 500 MM ERROR

| Timing Method | PD | Outliers | Accuracy | Total Hits |
|-------------------|-----|----------|----------|------------|
| Cumulative Energy | PD1 | 2 | 0.4 | 69 |
| | PD2 | 9 | 0.79 | 196 |
| SG Filter | PD1 | 4 | 0.61 | 148 |
| | PD2 | 6 | 0.41 | 94 |

are deemed outliers since most of the points are farther. For instance, the noise cancellation method has a fair number of hits, but it has low accuracy because most of the points close to the actual PDs are outliers. Conversely, some methods show high accuracy since all of the points far from the PDs are outliers. Examples of this are the envelope and wavelength comparison methods, which have a case of high accuracy but a large number of outliers. All of the points from the wavelength comparison method aside from a small accurate cluster are deemed outliers, falsely improving the accuracy of the method.

To determine an appropriate method for our data, we want few outliers, high accuracy and high total hits. This can be seen in the simple threshold method and the moving average method for PD2. However, only the cumulative energy and SG filter methods achieve all three criteria for both PDs. While both of these timing methods are successful, the SG filter method captures more PD coordinates than the cumulative energy method. This means that the points selected by the SG filter method are more centralized around the possible PD points, indicating a better prediction.

To further quantify the performance of these methods, we test PD localization with a smaller error range of 300 mm. With a tighter error range, the results in Table III indicate that the SG filter method has better localization accuracy than the cumulative energy method. The accuracy results for PD2 are much lower since the noisiest channel in the PD2 data has a much lower SNR than the noisiest channel in PD1. Since our sample size only includes about 20 data samples for each PD, we need to test our methods against more sample data.

We also test our method of the SG filter with multilateration to determine the error. As Table IV shows, with multilateration, neither timing method is consistently better in terms of either outliers or accuracy. Consequently, either the cumulative energy or SG filter method can be used to approximate the PD source, since the error from multilateration is greater than the error from either signal timing method.

Validation Tests

Aside from testing our methods on emulated PD data, as additional validation, we use the simulated FDTD data to

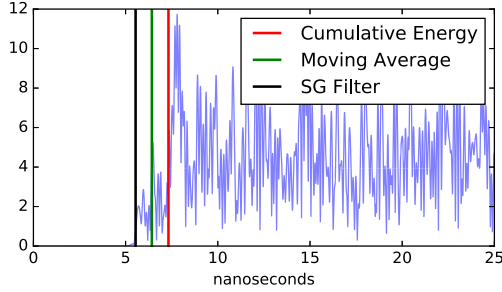


Fig. 12. Signal Timing on Simulated Data Channel 1

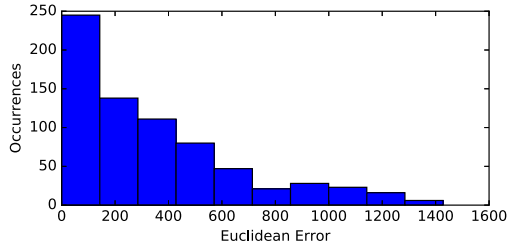


Fig. 13. Distribution of Error from SG Filter Method

check the signal timings without using the predicted PD source as the quantifier. To simulate a real signal, we add gaussian noise to the FDTD data, and test our timing methods on the noisy data. Figure 12 compares the identified signal timings in the simulated data given by three different methods. The blue lines represent the signal, which begins at around 5 nanoseconds. The SG filter method consistently provides the timing closest to the actual signal arrival, so our method indeed offers the best timing results.

In addition, using the validation data, the average error from all of the possible FDTD indices is about **340 mm** in terms of Euclidean distance, which is fairly close to the minimum 300 mm error given by the size of the FDTD indices. Also, this validation test has a single sample for each mesh point, so in real cases some outliers would be removed.

The distributions of the FDTD mesh point errors from the 3MVA transformer are shown in Figures 13 and 14. The prediction errors from the SG filter method are a lot lower

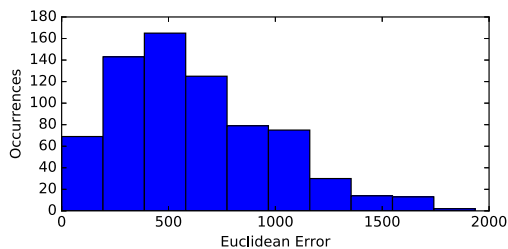


Fig. 14. Distribution of Error from Cumulative Energy Method

than those of the cumulative energy method. This indicates that there are some indices in the transformer that are harder to model than others, but on a whole, the SG filter method is able to predict the behavior of these indices more accurately. We achieved similar results from the TL1965 transformer, although with less consistency. This is attributed to the greater complexity of the TL1965 transformer.

V. DISCUSSION

Our method for locating the source of a partial discharge involves a number of different tools and configuration parameters. Here is a brief summary of our choices used in Section IV and possible alternatives for additional work.

An important tool we use in our work is the FDTD simulation reference table, which contains the signal arrival times computed from the mesh point voltages. We used the energy criterion method, which was the best available method at the start of the work, to compute the reference table. In tests, we see that this method works quite well; however, through our work we developed a better method, with which we plan to update the reference table.

We use four channels (sensors) to locate the PD, since that is the minimum number of signals required, and because that is the number of signals that can be captured at once with the current monitoring hardware. However, there exist more signals that could be used. Using more signals should improve the location accuracy of our method.

The specific parameters for the SG filter method were determined empirically in our experiment. These parameters include the window size and the order of the Savitzky-Golay filter as well as the size of the window identified by the initial threshold and the range of data that is considered noisy to be used for the threshold. These parameters can be changed to improve signal timing performance in other datasets, which may not have the same frequency or size as the sensor data that we used.

As seen in our validation tests, the effectiveness of our method varies on different transformers. This can be attributed to the different complexities of various transformers, including metal obstacles or the shape of the transformer. Potentially, we could adjust the parameters to better fit the properties of the individual transformer.

VI. CONCLUSION

The goal of our study is to accurately locate the source of partial discharges in a transformer. We break this task in two steps: compute the arrival time at each sensor and then locate the source based on the time differences. We studied a number of methods to determine the arrival time from sensor measurements and developed a convolutional filter based method that works better than all others. We determine the PD location by performing a reference table lookup for the arrival time differences for the four sensors. This reference table is constructed using FDTD simulations on the detailed structure of the transformer.

Our convolutional filtering method with the SG filter significantly improves our ability to determine arrival time from the sensor data. In tests with simulated events where the source locations are known, our convolutional filter produces the best timing results among all the methods tested. Given the arrival time values, looking up the source location using the reference table also produces much more precise locations than multilateration. Using our method, we are able to locate the PD to within 500 mm and 300 mm much more frequently than with other methods. With 500 mm tolerance, our method consistently achieved more total hits than the existing cumulative energy method. In the PD1 data, the SG filter method attained 554 hits compared to the 536 hits with the cumulative energy method, and in the PD2 data, 382 hits versus 238 hits. With 300 mm tolerance, in the PD1 dataset where the signal-to-noise ratios are high, the accuracy with cumulative energy method is 91% and our method is 95%. However, in the high noise case of the PD2 data, we had a more drastic improvement from 13% accuracy to 48% with our method.

In future work, we would like to test this method on other real transformer sensor measurements. With those data, we can better understand the impact of the signal selection and the parameters of our SG filter method on the accuracy of the signal timing procedure.

ACKNOWLEDGMENT

This work was supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

REFERENCES

- [1] A. Haddad and D. Warne, *Advances in High Voltage Engineering*. The Institution of Engineering and Technology, 2009.
- [2] G. J. Paoletti and A. Golubev, "Partial discharge theory and technologies related to medium voltage electrical equipment," *IEEE Transactions on Industry Applications*, vol. 37, pp. 456–460, Jan. 2001.
- [3] C. S. Park, J. B. Lee, and C. H. Cho, "A study on integrated preventive diagnostic system for gis," in *GCC Power on GIGRE, 2015*, 2015.
- [4] M. Wang, A. J. Vandermaar, and K. D. Srivastava, "Review of condition assessment of power transformers in service," *IEEE Electrical Insulation Magazine*, vol. 18, pp. 12–25, Nov. 2002.
- [5] M. Duval, "A review of faults detectable by gas-in-oil analysis in transformers," *IEEE Electrical Insulation Magazine*, vol. 18, pp. 8–17, May 2002.
- [6] A. Akbari, P. Werle, H. Borsi, and E. Gockenbach, "Transfer function-based partial discharge localization in power transformers: a feasibility study," *IEEE Electrical Insulation Magazine*, vol. 18, pp. 22–32, Sep. 2002.
- [7] S. M. H. Hosseini and P. R. Baravati, "Transformer winding modeling based on multi-conductor transmission line model for partial discharge study," *Journal of Electrical Engineering and Technology*, vol. 9, pp. 154–161, 2014.
- [8] K. Raja and T. Floribert, "Comparative investigations on uhf and acoustic pd detection sensitivity in transformers," in *Conference Record of the 2002 IEEE International Symposium on Electrical Insulation*. IEEE, Apr. 2002.
- [9] S. Tenbohlen, D. Denisov, S. M. Hoek, and S. M. Markalous, "Partial discharge measurement in the ultra high frequency (uhf) range," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 15, pp. 1544–1552, Dec. 2008.
- [10] A. M. Ishak, M. D. Judd, and W. H. Siew, "A study of uhf partial discharge signal propagation in power transformers using ftd modelling," in *2010 45th International Universities Power Engineering Conference*. IEEE, Aug. 2010.
- [11] S. M. Markalous, S. Tenbohlen, and K. Feser, "Detection and location of partial discharges in power transformers using acoustic and electromagnetic signals," *IEEE Transactions on Dielectrics and Electrical Insulation*, Dec. 2008.
- [12] M. D. Judd, L. Yang, and I. B. B. Hunter, "Partial discharge monitoring for power transformers using uhf sensors part 1: sensors and signal interpretation," *IEEE Transactions on Dielectrics and Electrical Insulation*, pp. 5–14, Mar. 2005.
- [13] S. Coenen, S. Tenbohlen, S. Markalous, and T. Strehl, "Sensitivity limits of uhf pd measurements on power transformers," in *Proceedings of the International Symposium on High Voltage Engineering*, Aug. 2009.
- [14] T. Hoshino, K. Nojima, and M. Hanai, "Real-time pd identification in diagnosis of gis using symmetric and asymmetric uhf sensors," *IEEE Transactions on Power Delivery*, pp. 1072–1077, Jul. 2004.
- [15] H. R. Mirzaei, A. Akbari, E. Gockenbach, M. Zanjani, and K. Miralikhani, "A novel method for ultra-high-frequency partial discharge localization in power transformers using the particle swarm optimization algorithm," *IEEE Electrical Insulation Magazine*, vol. 29, no. 2, pp. 26–39, Mar. 2013.
- [16] R. Sarathi, I. P. M. Sheema, and V. Subramanian, "Propagation of partial discharge signals and the location of partial discharge occurrences," in *2013 8th IEEE International Conference on Industrial and Information Systems*. IEEE, Dec. 2013.
- [17] B. M. Schettino, C. A. Duque, and P. M. Silveira, "Current-transformer saturation detection using savitzky-golay filter," *IEEE Transactions on Power Delivery*, pp. 1400–1401, Jun. 2016.
- [18] C. Abraham and S. V. Kulkarni, "Ftd simulated propagation of electromagnetic pulses due to pd for transformer diagnostics," in *Proceedings of the IEEE Region 10 Conference*. IEEE, Nov. 2008.
- [19] K. S. Yee, "Numerical solution of initial boundary value problem involving maxwell's equation in isotropic media," *IEEE Transactions on Antennas and Propagation*, pp. 302–307, Jan. 1966.
- [20] M. D. Judd, "Using finite difference time domain techniques to model electrical discharge phenomena," in *2000 Annual Report Conference on Electrical Insulation and Dielectric Phenomena*. IEEE, Oct. 2000.
- [21] A. J. Reid, M. Stewart, and M. D. Judd, "Ftd modeling of uhf partial discharge sensor response," in *International Conference on Sustainable Power Generation and Supply*. IEEE, Apr. 2009.
- [22] A. Bojovschi, W. S. T. Rowe, and A. K. L. Wong, "Electromagnetic field intensity generated by partial discharge in high voltage insulating materials," *Progress in Electromagnetics Research*, pp. 167–182, 2010.
- [23] T. Hoshino, S. Maruyama, and T. Sakakibara, "Simulation of propagating electromagnetic wave due to partial discharge in gis using ftd," in *International Conference on Condition Monitoring and Diagnosis*. IEEE, Apr. 2008.
- [24] V. Satopaa, J. Albrecht, D. Irwin, and B. Raghavan, "Finding a 'kneedle' in a haystack: Detecting knee points in system behavior," in *2011 31st International Conference on Distributed Computing Systems Workshops*. IEEE, 2011.
- [25] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964.
- [26] W. L. Cruz, J. M. Martinez, and M. Raydan, "Spectral residual method without gradient information for solving large-scale nonlinear systems of equations," *Mathematics of Computation*, vol. 75, no. 255, pp. 1429–1448, Jul. 2006.
- [27] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise," in *KDD'96 Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. AAAI, 1996.