ARC FAULT DETECTION IN

DC PHOTOVOLTAIC SYSTEMS

A Dissertation

by

ZHAN WANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Robert S. Balog
Committee Members,	Prasad Enjeti
	Tie Liu
	Sivakumar Rathinam
Head of Department,	Miroslav M. Begovic

December 2016

Major Subject: Electrical Engineering

ABSTRACT

Arc faults have always been a concern for electrical systems as they can cause fires, personnel shock hazard, and system failure. In photovoltaic (PV) systems, a large number of electrical connectors and long wire runs are expected. Combined with the high DC voltage, deterioration of the wire insulation due to aging or other circumstances such as rodent bites and abrasion due to chaffing with trees, building walls, or conduit during installation can cause electric arcs to occur. These dc arcs may result in shock hazards, fires, and system failures or faults in the PV systems. NEC 2011 includes a requirement for new rooftop arrays to include UL1699B listed arc fault current interrupters (AFCI). NEC 2014 expands this requirement to include ground-mounted arrays as well.

Existing commercialized techniques that rely on pattern recognition in the time domain, or frequency domain analysis using a Fourier Transform do not work well because the signal to noise ratio is low, and the arc signal is not periodic. Instead, wavelet transform provides a time-frequency approach to analyzing target signals with multiple resolutions.

In this work, a technique for arc fault detection photovoltaic systems by using discrete wavelet transform (DWT) for feature extraction and support vector machines for decision making is proposed.

The frequency characteristics of electric arcs in the PV systems are first studied. The fundamental feasibility of applying wavelet theory to detect arc fault and arc flash in solar PV power systems is then examined both in simulation using synthetic waveforms

generated in MATLAB / Simulink and experimentally using arc waveforms measured from actual dc PV systems with/without operating inverters.

In the later chapter, a supervised learning method for arcing/non-arcing event classification using support vector machines (SVMs) is introduced. SVMs are believed to be one of the best "off-the-shelf" supervised learning algorithms. The main concept behind SVM is to create a hyperplane with a maximum margin between the two adjacent classes which helps bound the generalization error of the classification model. Different combinations of mother wavelets, decomposition levels, and kernel functions are examined in this work. Some of the strategies have shown very promising results.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Dr. Robert S. Balog, for his academic guidance and funding support throughout the course of this research. His invaluable wisdom and patience inspired the completion of this work.

I would like to thank my committee members, Dr. Prasad Enjeti, Dr. Tie Liu, and Dr. Sivakumar Rathinam, for their advice throughout my entire graduate study at Texas A&M University.

Thanks also go to my fellow colleagues working at the Renewable Energy & Advanced Power Electronics Research Laboratory and the Power Quality Laboratory: Mehran Mirjafari, Souhib Harb, Mohammad Shadmand, Haiyu Zhang, Somasundaram Essakiappan, Harish Sarma, Poornima Mazumdar, Dibyendu Rana, Pawan Garg, Bo Tian, Hezi Zhu, Stephen McConnell, Xiao Li, Morcos Metry, Mostafa Mosa, Cooper Barry, and Phillip Hatton for their friendship and so many precious memories.

I am also very grateful to Ran Huang, Yingxue Shine Yang, and Jing Sun. Thank you for being my closest friends and your impact on every aspect of my life.

Finally, the completion of this work would not have been possible without the encouragement and love from my mother, Lirong Wang.

CONTRIBUTORS AND FUNDING SOURCES

This work was supported by a dissertation committee consisting of Dr. Rober S. Balog, Dr. Prasad Enjeti and Dr. Tie Liu of the Department of Electrical and Computer Engineering and Dr. Sivakumar Rathinam of the Department of Mechanical Engineering. The data analyzed for Chapter IV was provided by the Sandia National Laboratories.

All other work conducted for the dissertation was completed by the student independently.

This research was supported in part by the National Science Foundation under Grant No. ECCS-1238412. It was also partially funded by the Department of Energy's Office of Energy Efficiency and Renewable Energy through a collaboration with Sandia National Laboratories. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

The statements made herein are solely the responsibility of the authors.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION*

1.1 Arc Faults in DC PV Systems

There are three major catastrophic types of failures in photovoltaic (PV) arrays: ground faults, line-to-line faults, and arc faults [7]. In PV systems, a large number of electrical connectors and long wire runs are expected. Combined with the high DC voltage, deterioration of the wire insulation due to aging or other circumstances such as rodent bites and abrasion due to chaffing with trees, building walls, or conduit during

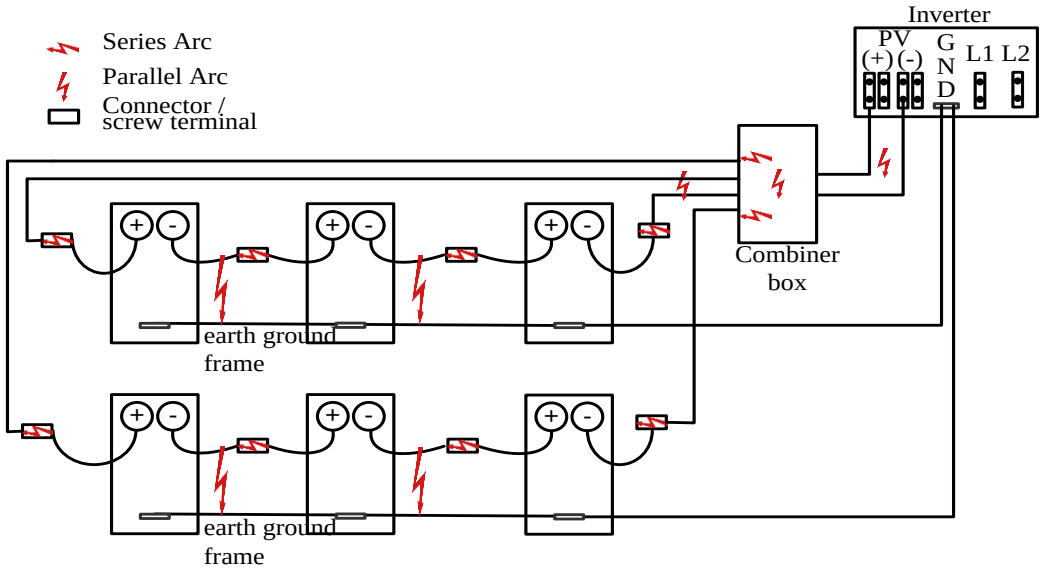


Figure 1: Example of locations where arcing may occur in a PV array.

* Reprinted with permission from "Arc Fault and Flash Detection in DC Photovoltaic Arrays Using Wavelets," by Z. Wang and R.S. Balog, 2013, IEEE 39th Photovoltaic Specialists Conference (PVSC), pp. 1619-1624, © 2013 IEEE. Reprinted with permission from "Arc Fault Signal Detection – Fourier Transformation vs. Wavelet Decomposition Techniques Using Synthesized Data," by Z. Wang and R.S. Balog, 2014, IEEE 40th Photovoltaic Specialists Conference (PVSC), pp. 3239-3244, © 2014 IEEE.

installation can cause electric arcs to occur. These DC arcs may result in shock hazards, fires, and system failures or faults in the PV systems [8].

Electrical arcs in PV systems can arise from series or parallel faults, as illustrated in Figure 1. Series arc faults often occur due to loose electrical connections while parallel faults can be caused by abrasion of conductors from thermal cycling or vibration, puncture of the insulation by rodents, or other failures within the PV system [9, 10].

Figure 2 illustrates possible sources of arc faults due to the way a utility-scale DC PV power system is installed. Figure 3 [3] provides an example of aging cable connectors which are very likely to cause loose connections in an operating PV system. As long as this arc fault problem exists, PV systems face significant concerns about liability which threaten the extensive use. Thus, arc fault detection is extremely important for reliable and safe system operation [11, 12] and is a prerequisite for widespread adoption of DC PV systems [13-16].

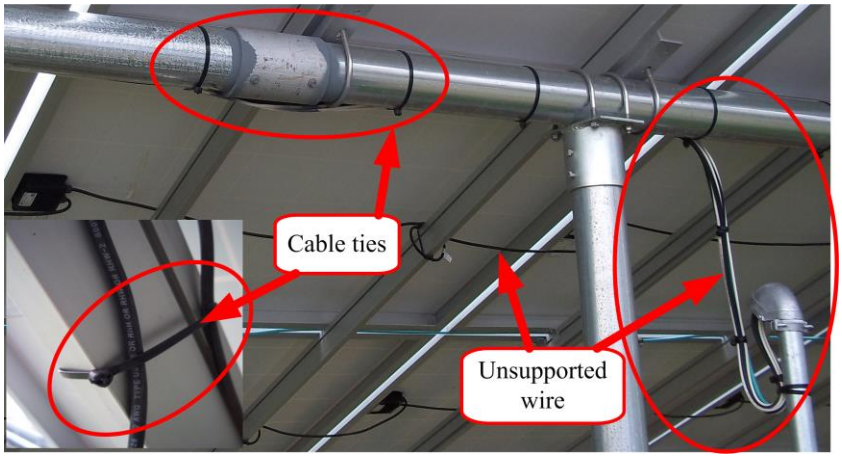


Figure 2: Example of DC wiring in a ground-mounted PV array (photo credit: Robert S. Balog).

Figure 3: An example of aging cable connectors which are very likely to cause loose connection [3].

Even more important is to detect arc flash, the pre-fault (before a sustained arc forms) events of sparking and dielectric breakdown. Arc flash may only last for a short duration (less than a second) but serves as an early indicator of incipient arc faults. Detecting arc flash is a difficult problem because unlike a bolted "hard short" fault in which high current flows through a metal-to-metal connection [17]. Arc flash involves short-term current flowing through the ionized air or along an ion path and may not draw sufficiently high RMS current, or have a high enough I^2t energy to trip a thermal circuit breaker. This is particularly true in finite-energy systems, such as many of the dc microgrids and systems energized by renewable energy sources. In these cases, an arc, like the one shown in Figure 4, can be sustained for hours or even days because the

Figure 4: Arcing persists in the DC wiring even after a fire consumes a portion of the combiner box [4].

overcurrent protection devices never activate [4]. Thus the fire and safety hazards are left undetected and unmitigated.

Unlike an AC system in which power electronics are typically only found at the point-of-load, a dc system requires the use of dc/dc converters throughout the distribution systems [18] which adds distributed capacitance throughout the system providing numerous coupling pathways for high-frequency signals. High-frequency noise from the dc/dc converter switching and other electromagnetic interference could obfuscate the arc signature, allowing an arc to establish and be sustained undetected [19].

In order to meet the goal of the SunShot Initiative launched by the U.S. Department of Energy in 2011, arc fault protection must be provided within the context of $0.40/watt balance of system and $0.10/watt power electronics costs. A highly integrated arc fault detector and circuit interrupter is the best solution to address the safety needs of the industry installations and retrofit applications [20].

1.2 Outlook for Photovoltaic Industry

The growth of photovoltaics (PV) has been increasing exponentially for more than two decades around the world, as illustrated in Figure 5 [2]. During this period of time, PV has evolved from a small market of applications by early adopters towards a mainstream electricity source. Economic incentives, such as feed-in tariffs, were implemented by a number of governments when PV systems were first recognized as a promising renewable energy source. Consequently, the cost of PV installation declined significantly due to the improvements in the technology and economies of the scale [2, 21, 22].

PV deployment has grown rapidly in the United States over the past several years. As shown in Figure 6 [5], the compound annual growth rates (CAGR) for the U.S. residential commercial and utility-scale PV sectors from 2010-2015 were 46%, 43%, and 101%, respectively. Wide adoption of solar energy has been acknowledged to strengthen

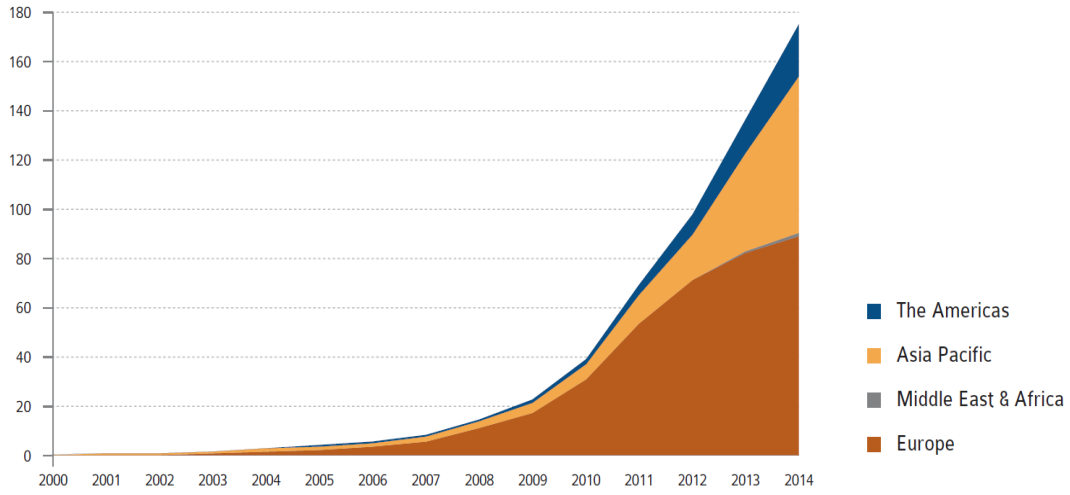


Figure 5: Cumulative regional PV installation [2].

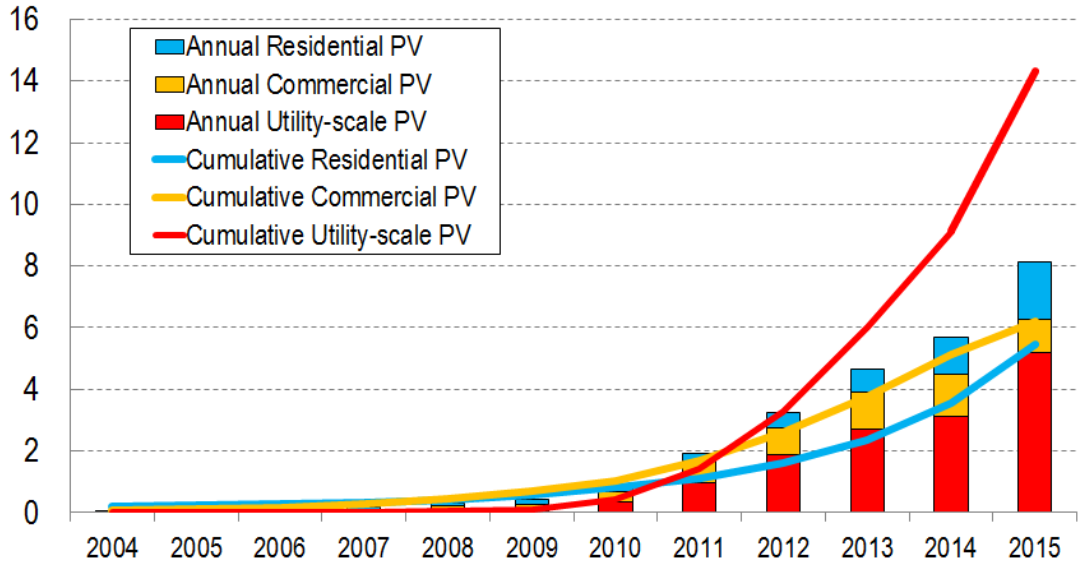


Figure 6: U.S. PV market growth, 2004-2015, in gigawatts of direct current (DC) capacity [5].

U.S. economic competitiveness in the renewable energy race, help cut carbon pollution to combat climate change, and secure America's energy future. The United States has immense potential to power the country with solar energy: photovoltaic (PV) panels on just 0.6% of the nation's total land area could supply enough electricity to power the entire United States [23]. PV can also be installed on rooftops without actually using any land. Further, as a domestic energy source, solar supports broader national priorities, including national security, economic growth, and job creation [23-25].

The SunShot Initiative was launched in 2011 by the U.S. Department of Energy launched with the goal of making solar electricity cost-competitive with conventionally generated electricity by 2020 [25, 26]. At the time, this meant reducing photovoltaic (PV) and concentrating solar power (CSP) prices by approximately 75% - relative to 2010 prices – across the residential, commercial, and utility-scale sectors. The SunShot Vision

Study published by the Department of Energy's Solar Energy Technologies Office (SETO) in 2012 projected that achieving the SunShot price-reduction targets could result in solar meeting roughly 14% of U.S. electricity demand by 2030 (11% PV, 3% CSP) and 27% (19% PV, 8% CSP) by 2050 – while reducing fossil fuel use, cutting emissions of greenhouse gases and other pollutants, creating solar-related jobs, and lowering consumer electricity cost [25, 26].

The SunShot initiative aims to reduce the total installed cost of solar energy systems to $0.06 per kilowatt-hour (kWh) by 2020 [25, 26]. Over the past five years, cumulative U.S. solar deployment has increased more than tenfold. The system price has dropped by 54% for utility scale, 63% for commercial, and 55% for residential systems (shown in Figure 7) and the levelized cost of energy (LCOE) has dropped by as much as 65% (shown in Figure 8) [6]. The recent report series "On the Path to SunShot" [6, 27-33] shows that today, at the halfway mark of the SunShot Initiative's 2020 target date, SunShot is about 70% of its way toward achieving the program's goal. Since SunShot's launch in 2011, the average price per kWh of a utility-scale photovoltaic (PV) project has dropped from about $0.21 to $0.11 [6].

Although the remarkable reduction in PV system prices and LCOE have been archived since the launch of the SunShot Initiative, significant additional reductions are needed to realize the 2020 SunShot Initiative. Among the various research to reach the SunShot goal, intelligent power electronics devices that maximize the power output from the PV arrays while ensuring overall system safety, reliability, and controllability are indispensable. A smart inverter functionality of predicting faults and improve system

reliability could effectively extend the lifetime of the PV installations, and more importantly, guarantees human safety around the systems.

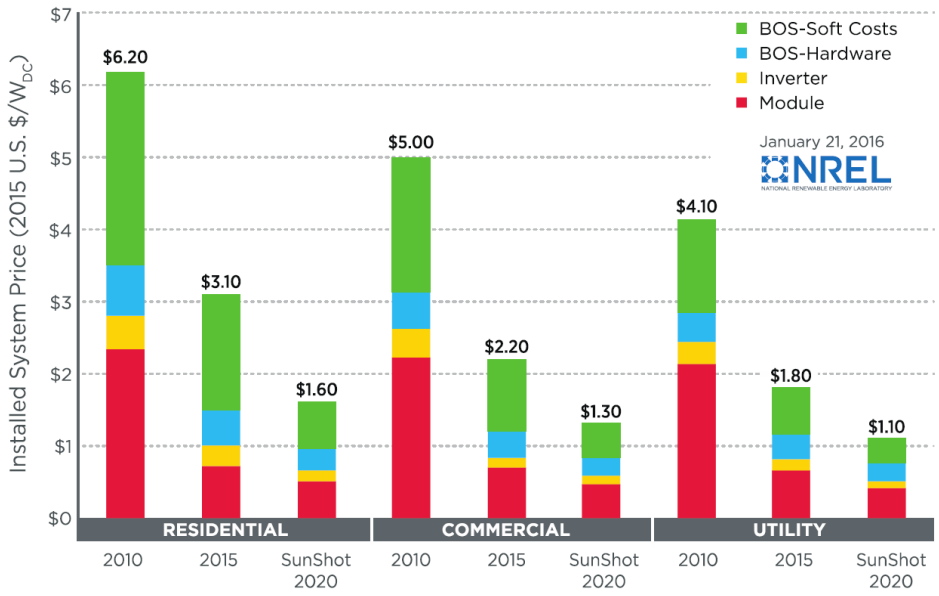


Figure 7: Historical, current, and SunShot 2020 target system prices for the utility, commercial and residential sectors (weighted national average for fixed-tilt systems) [6].

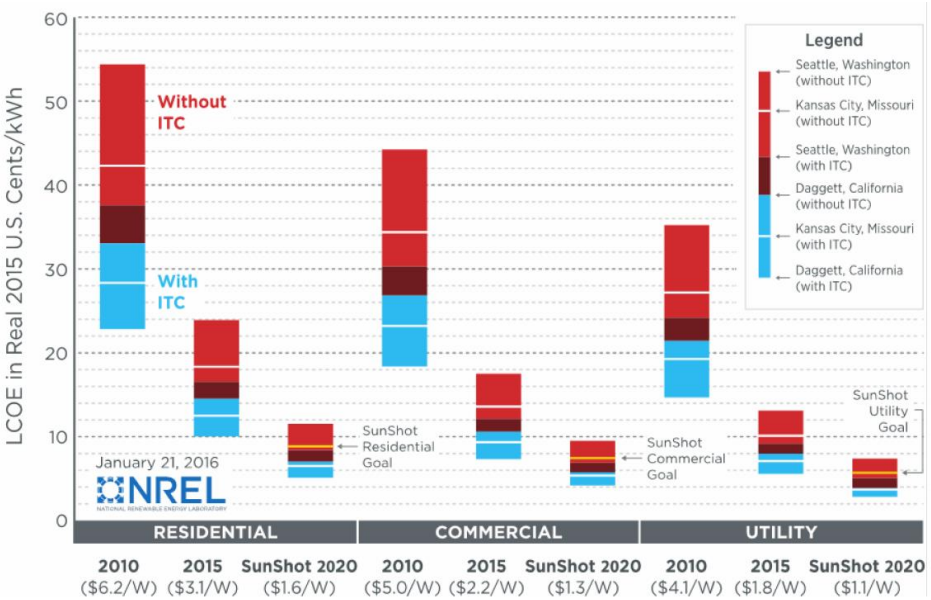


Figure 8: Historical, current, and SunShot LCOE calculations [6].

1.3 Arc Faults in AC Systems

AC arc faults have been well studied. The detection of AC arc faults has been well developed [12, 34-38] with commercial products designed and UL listed [39] for safety [40-42]. Comparatively, a much smaller body of work pertaining to arcs in dc electrical systems and commercialization of sensing and protecting devices has only recently begun [43]. A significant complication to their detection is that arcs in dc systems are not periodic, and thus may not have easily recognizable amplitude or frequency signatures for pattern recognition-based detection techniques. Spectral analysis using Fourier techniques to decompose the frequencies of a sustained arc or bolted fault requires a linear system and a stationary signal, and therefore Fourier techniques are not capable of reliably detecting arc flash.

In AC systems, signal decomposition using wavelet transform and wavelet packet have been proposed and worked well to detect the impulse-like effect of the discontinuous arc due to periodic extinguishing and re-ignition associated with the main frequency zero-crossing [11, 12, 34-36, 44]. Waveforms shown in Figure 9 are wavelet decomposition result of an AC arcing event at a residential house in College Station, Texas. The arcing was occurring within a 15A standard-duty single-pole light switch in a bathroom location. It is worth noting that the UL listed arc-fault current interrupter installed upstream in the circuit breaker box, in accordance with NEC and local code requirements, failed to detect the arc and de-energize the branch circuit. The arcing had been occurring sporadically for many weeks before this data being captured. The arc created enough energy to make the switch warm to the touch and created audible “cracking” noise.

While the exact algorithm in the circuit breaker is unknown, referring to Figure 9, the left is a healthy signal on the bottom, and above it shows a 3-level wavelet transform result of this non-arcing signal. On the right is the 3-level wavelet transform on the arcing signal with the same wavelet and the original time domain waveform of the arcing event.

It is worth noting that the comparison of the wavelet analysis results of these real-world signals matches what was demonstrated in the aforementioned literature: under AC conditions, the electric arc phenomenon can be described as “spark gap”. The spark gap will not conduct until the applied voltage across the gap reaches the breakdown point. The spark gap stops conducting once the voltage decreases to the breakdown point. This “re-ignition and extinction” process occurs every time the line voltage waveform hits the zero crossing. Since wavelet transform is particularly effective at discriminating discontinuous transients in the signal, impulse-like signals show up in the wavelet transform whenever the arc re-ignites or extinguishes. The pattern of this arcing event is periodic with two times the frequency of the supplied voltage. Thus, it is fairly easy to be detected in practice.

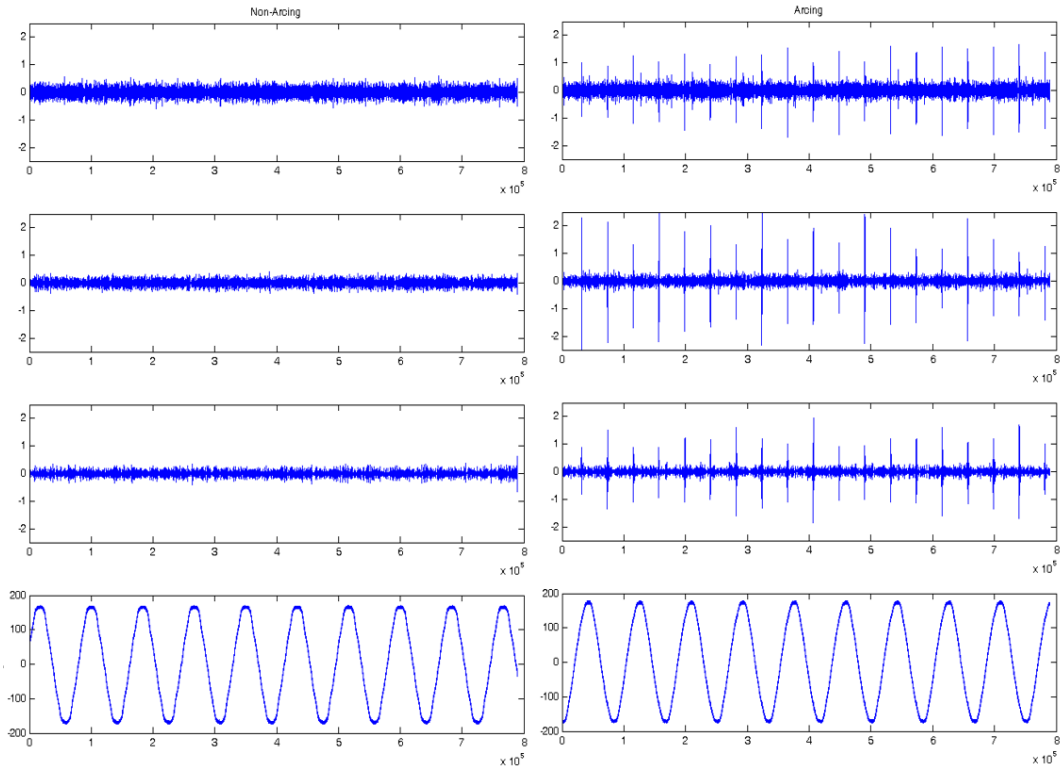


Figure 9: Electric arc in AC systems.
On the left is the 3-level wavelet decomposition and the original signal of a healthy 60Hz voltage signal; on the right is the 3-level wavelet decomposition and the original signal of the same system with arc happening.

1.4 Motivation of This Work

While arc faults are rare in photovoltaic installations, more than a dozen documented arc-faults have led to fires and resulted in significant damage to the PV system and surrounding structures [45].

Arc faults can occur in PV systems for a variety of reasons. In the exposed wires, there can be chaffing or abrasion to external conditions. For wires in a conduit, thermal expansion may be sufficient to cause abrasion, particularly if wire insulation was damaged during installation. Also, the modular nature of the string design means that there are a large number of connectors in the wiring between modules. Each module connector is a potential point of failure. In addition, modern grid-interactive PV system designs employ voltages as high as 600V [46]. High-voltage DC arcs are difficult to extinguish while the system is energized. Moreover, a traditional PV module has no means to disconnect the power source from the DC conductors. Even when the system is shut down, the conductors remain energized by the solar cells.

The PV electrical fire incidents caused by arc faults are the result of high-temperature plasmas produced as current passes across separated and damaged conductors. In response, the 2011 National Electrical Code Section 690.11 – requiring listed PV arc-fault circuit interrupters on PV installations – was created to reduce the likelihood of an electrical fire [46]. For PV inverter, smart combiner box, and original equipment manufacturer (OEM) products to become listed, the device must undergo a

sequence of tests defined in UL 1699B to verify its safety, ability to detect arc faults, and ensure a basic level of unwanted tripping [43].

From Sandia's unwanted tripping survey with UL 1699B-listed products [47], it is evident that there are limitations in algorithms of most of the listed products on the market. The motivation of this work is to study the electrical characteristics of the arc, extract salient features, and develop a more robust solution to avoid no-detection and false detection in real-world applications.

1.5 Dissertation Outline

This dissertation is organized as follows:

Chapter II: Existing arc fault detection approach in PV

- An existing arc fault detection approach which is widely adopted on many AFCIs on the market is studied.
- The intrinsic limitations of the tradition approach are explained.

Chapter III: Arc signal analysis using wavelet transform

- The fundamental theory of discrete wavelet is introduced.
- The reason why wavelet transform might be a better fit for this application than the traditional Fourier-based approaches is analyzed.
- Hardware implementation strategy of the wavelet transform is illustrated. Its feasibility for sample-by-sample real-time analysis is explained.
- A mathematic model of the arc is simulated in MATLAB / Simulink.
- Real-world waveforms measured from photovoltaic modules and an experimental arc generator are tested.

Chapter IV: Arc fault detection using wavelet transform and support vector machines

- A few basic concepts of machine learning are introduced.
- Feature extraction using wavelet transform and Parseval's theorem is proposed.
- Theory and derivation of support vector machines are briefly described.

- A couple of feature rescaling strategies are evaluated using cross-validation.
- An exhaustive search for the optimal feature extraction and feature selection strategies is performed.

Chapter V: Contributions and future work

- The contributions of the work are summarized.
- Future research interests are presented.

2. EXISTING ARC FAULT DETECTION APPROACH IN PV*

There are currently commercial products available and even required in some applications for AC arc detection in residential AC systems. Known as arc fault circuit interrupters (AFCIs), these products are required to detect both series and parallel arc faults [39]. AFCIs typically use current sensors and analog filters to acquire a filtered analog current signal in a specific frequency band where the arc fault signal is assumed to be most detectable. The filtered time-domain current signal is then processed, usually by proprietary detection algorithms and carefully tuned threshold setting in a digital signal processor (DSP) or microprocessor [40, 42]. Some research, however, has shown that neither branch/feeder AFCI nor combination AFCI would accurately detect all series arc faults [48]. This could be in part due to how the threshold of the detection algorithm was tuned and the assumptions made in the filter as to the frequencies in which the arc signature signal appears.

To give an example, a commercially available solution is designed to detect arc fault in a dc PV system using fast Fourier transform (FFT) as the detection method. The process, shown in Figure 10, uses a wide bandwidth coupled inductor circuit. An isolation transformer is used to isolate the high DC voltage and current from the arc monitoring circuit. The application in a PV string array is shown in Figure 11. The detection method

* Reprinted with permission from "Arc Fault and Flash Signal Analysis in DC Distribution Systems Using Wavelet Transformation," by Z. Wang and R.S. Balog, 2015, IEEE Transactions on Smart Grid, vol. 6, pp. 1955-1963. © 2015 IEEE

assumes that the arc signature lies predominantly in the frequency band between 40 kHz and 100 kHz and uses a pre-filter to condition the analog signal [49]. Nevertheless, other non-arcing related signals, such as switching harmonics from inverters and DC/DC optimizers may also generate signals in this frequency band which can lead to false detection or non-detection by masking the arc signature. Obviously, non-detection is detrimental since the hazard is undetected. False detection is also undesirable because the response may unnecessarily shut down the system, causing loss of revenue or even the potential for grid instability when the PV generation trips offline unexpectedly and needlessly.

Although the conventional Fourier transform is deeply researched and widely used, the fact that it works best for periodic signals is a significant limitation. The nature of arc faults in power systems is not periodic [11]. Further, only frequency information is given by traditional Fourier transform approaches; not enough time-domain information is provided to find out exactly when the event occurs. Such temporal localization could help correlate the electrical arc characteristics with other accompanying events such as lighting or fast transients that couple from other devices in the system.

The short-time Fourier transform (STFT) is a Fourier-related transform used to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time. This transform still has a fundamental drawback in that the length of the window used in the STFT is the same for all frequencies which leads to a fixed resolution. The window length selection then becomes a tradeoff between frequency resolution and time resolution. A large number of samples is required to obtain high

frequency resolution, which in turn causes low time resolution. A shorter window provides better time resolution but inferior frequency resolution [50, 51].

It is also worth pointing out that in order to minimize the spectral leakage, window size usually has to be chosen carefully to meet the coherent sampling requirement. However, the arc fault signature is distributed in a wide frequency band [49, 52]. It is impossible to choose a perfect window to accurately extract all the relevant information using Fourier transform based methods.

In conclusion, discrete STFT might be suitable for time-frequency domain analysis of harmonic related disturbances, but it is not ideal for capturing abrupt disturbances or short transient signals.

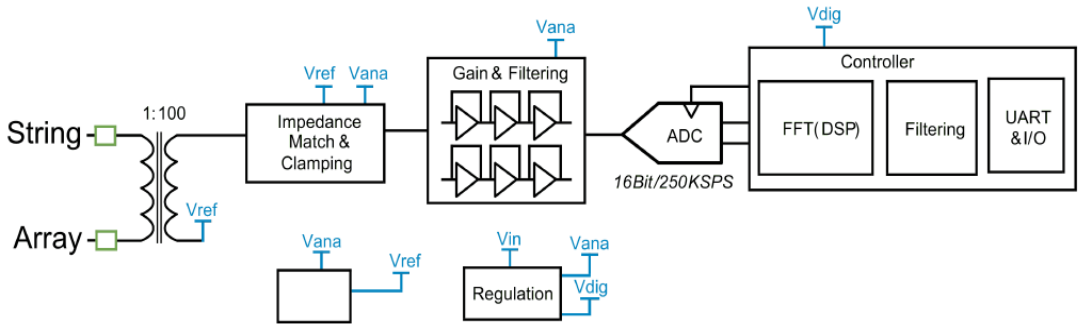


Figure 10: System diagram of a commercially available solution.

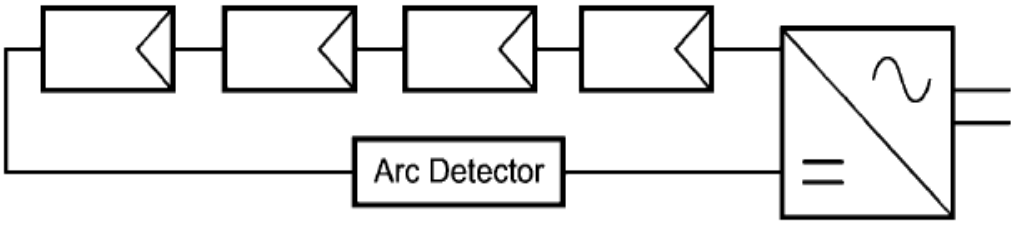


Figure 11: System application of SM73201 to detect series arc faults by sensing current.

3. ARC SIGNAL ANALYSIS USING WAVELET TRANSFORM*

3.1 Discrete Wavelet Transform

The wavelet transform is a mathematical tool for signal analysis first introduced in 1909 and further developed gradually since the 1970s. Wavelet theory establishes that a general transient signal can be constructed by the superposition of a set of special pieces of signals that occur with different time scales and at distinct times. A few typical continuous wavelet functions are given in Figure 12 [53]. For any signals that can be used as wavelets, they must satisfy the admissibility criteria:

- The signal must have a zero mean;
- The signal must be oscillatory;
- The signal must decay to zero quickly.

The wavelet transform (WT) is a linear transformation like the Fourier transform. Unlike FFT, it allows time localization of different frequency components of a given

* Reprinted with permission from "Arc Fault and Flash Detection in DC Photovoltaic Arrays Using Wavelets," by Z. Wang and R.S. Balog, 2013, IEEE 39th Photovoltaic Specialists Conference (PVSC), pp. 1619-1624, © 2013 IEEE. Reprinted with permission from "Arc Fault Signal Detection – Fourier Transformation vs. Wavelet Decomposition Techniques Using Synthesized Data," by Z. Wang and R.S. Balog, 2014, IEEE 40th Photovoltaic Specialists Conference (PVSC), pp. 3239-3244, © 2014 IEEE. Reprinted with permission from "Arc Fault and Flash Signal Analysis in DC Distribution Systems Using Wavelet Transformation," by Z. Wang and R.S. Balog, 2015, IEEE Transactions on Smart Grid, vol. 6, pp. 1955-1963, © 2015 IEEE. Reprinted with permission from "High Fidelity Replay Arc Fault Detection Testbed," by H. Zhu, Z. Wang, S. McConnell, P.C. Hatton, R.S. Balog and J. Johnson, 2016, IEEE 43rd Photovoltaic Specialists Conference (PVSC), pp. 1767-1772, © 2016 IEEE. Reprinted with permission from "Arc Generator for Photovoltaic Arc Fault Detector Testing," by P.C. Hatton, M. Bathaniah, Z. Wang, and R.S. Balog, 2016, IEEE 43rd Photovoltaic Specialists Conference (PVSC), pp. 1702-1707, © 2016 IEEE. Reprinted with permission from "Real Time Arc Fault Detection in PV Systems Using Wavelet Decomposition," by H. Zhu, Z. Wang, R.S. Balog, 2016, IEEE 43rd Photovoltaic Specialists Conference (PVSC), pp. 1767-1772, © 2016 IEEE.

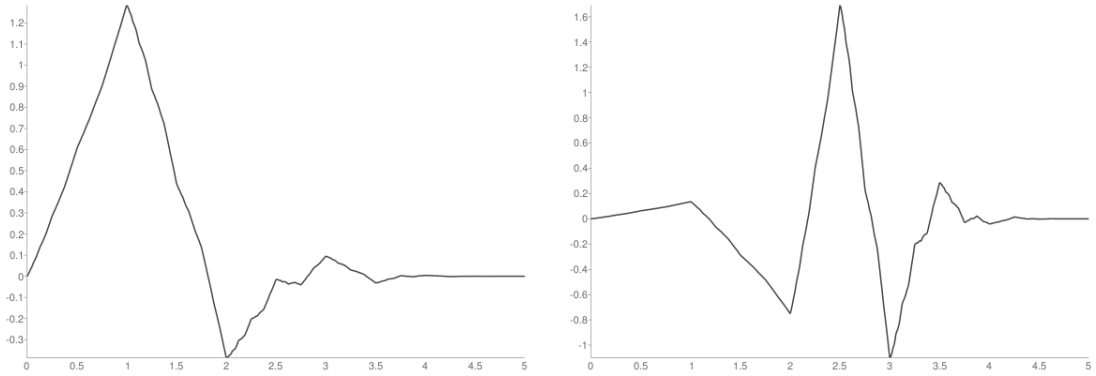
signal [54]. The wavelet transform cuts up data or functions or operators into different frequency components, and then studies each component with a resolution matched to its scale.

The wavelet analysis procedure is based on a wavelet prototype function called a "mother wavelet" which provides a localized signal processing method to decompose the differential signal into a series of wavelet components, each of which is a time-domain signal that covers a specific frequency band [55, 56]. Wavelets are particularly effective in approximating functions with discontinuities or sharp changes like power system fault signals [57]. With proper choice of the mother wavelet, the wavelet transformation is a good tool for signal analysis and fault feature extraction.

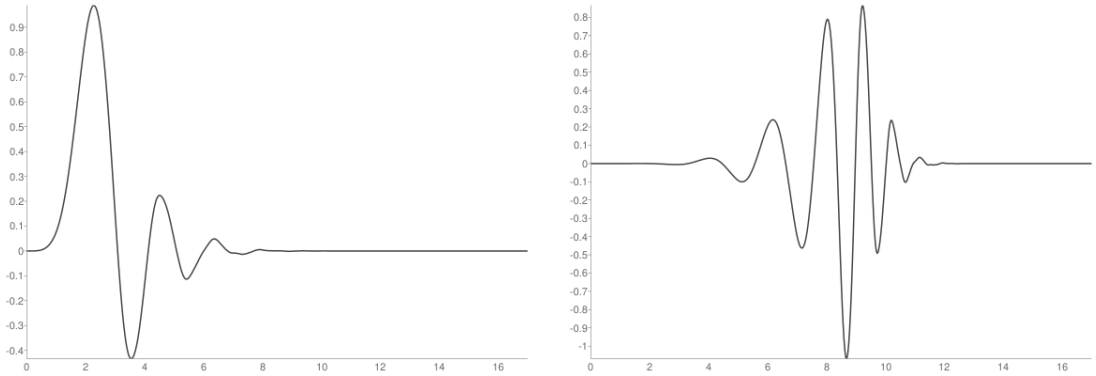
The wavelet transform is a powerful tool for statistical analysis in signal processing which have been adopted in a broad range of applications, for example:

- Data and image compression
- Partial differential equation solving
- Pattern extraction and recognition
- Texture analysis
- Noise/trend reduction
- Signal de-noising

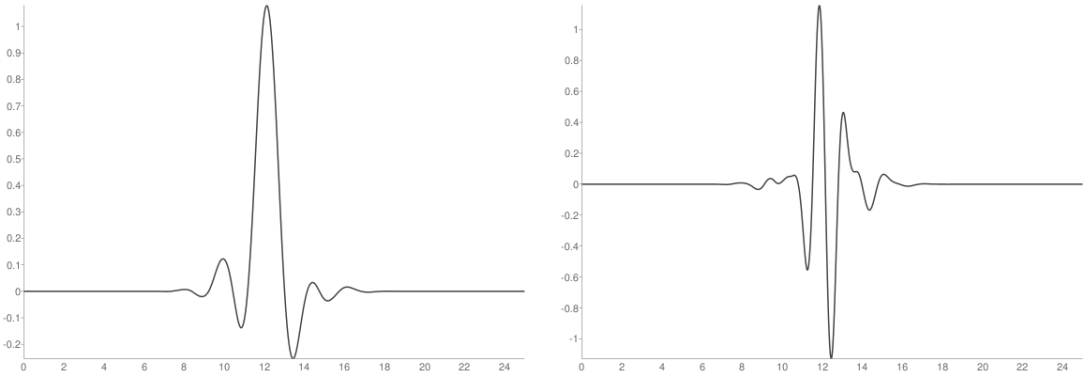
Due to the wide variety of signals and problems encountered in power engineering, there are various suitable applications of the wavelet transform, such as fault detection, load forecasting, and power system measurement. In addition, information about power disturbance signals is often a combination of features that are well localized



(a) Scaling function and wavelet function of Daubechies 3 (db3)



(b) Scaling function and wavelet function of Daubechies 9 (db9)



(c) Scaling function and wavelet function of Symlets 13 (sym13)

Figure 12: Scaling and wavelet functions of db3, db9, and sym13.

temporally or spatially such as power system transients. This requires the use of versatile analysis methods to handle signals regarding their time-frequency localization, which is an excellent area to apply the special property of wavelets [58].

The discrete wavelet transform is any wavelet transform for which the wavelets are discretely sampled. The discrete wavelet transform (DWT) is defined as

$$C(j,k) = \sum_{n \in Z} s(n)\varphi_{j,k}(n) \qquad j \in Z, k \in Z$$

$$D(j,k) = \sum_{n \in Z} s(n)\psi_{j,k}(n)$$

$$j \in Z, k \in Z$$

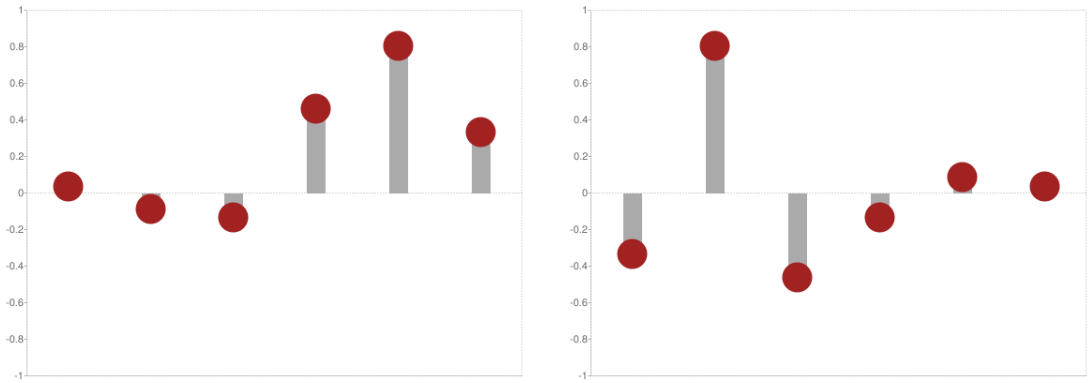
where $C(j,k)$ and $D(j,k)$ are the corresponding approximation and detail coefficients, n is the sample number, $s(n)$ is the signal to be analyzed, $\varphi(n)$ is the discrete scaling function (also called the father wavelet), and $\psi(n)$ is the mother wavelet. For dyadic-orthonormal wavelet transform, a series of rescaled and shifted functions can be derived by:

$$\varphi_{j,k}(n) = 2^{-j/2}\varphi(2^{-j}n - k)$$

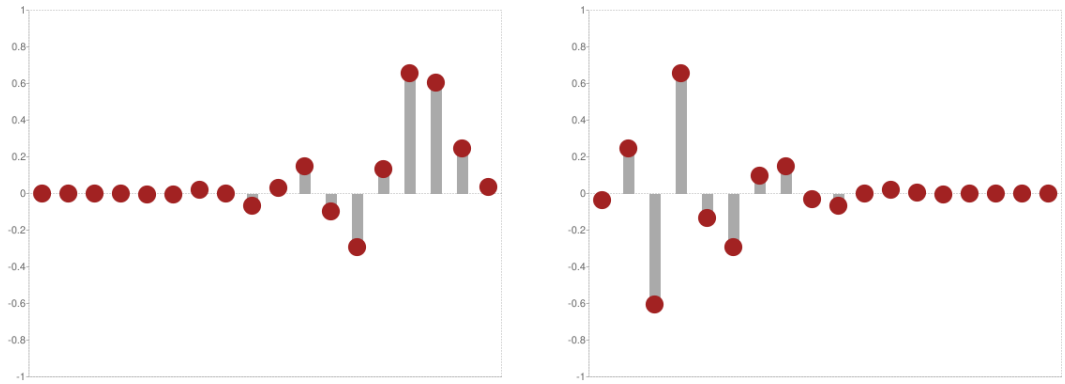
$$\psi_{j,k}(n) = 2^{-j/2}\psi(2^{-j}n - k)$$

where the set of $\psi_{j,k}(n)$ s are called daughter wavelets

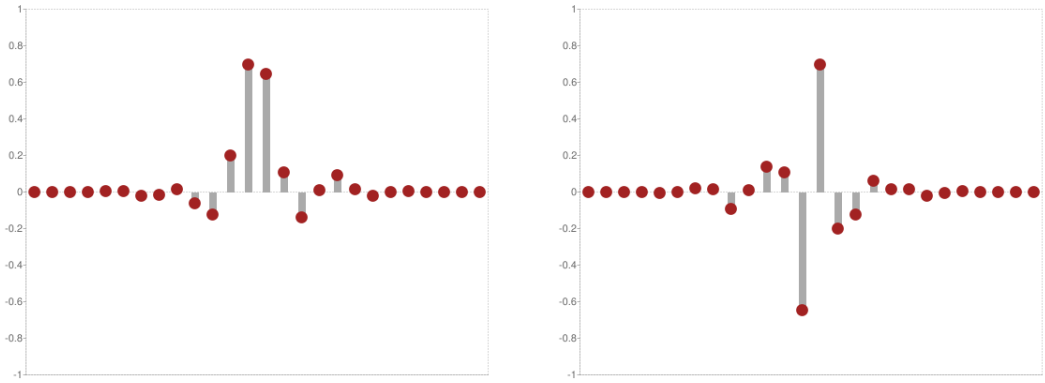
With this initial setting, the DWT can be easily implemented by the multi-resolution analysis (MRA). Impulse responses of the low-pass and the high-pass filters corresponding to the continuous wavelets in Figure 12 are listed in Figure 13. As shown in Figure 14, at each level j, approximation signal A_j (represented by linear combinations



(a) Impulse response of the low-pass and the high-pass filter of Daubechies 3 (db3)



(b) Impulse response of the low-pass and the high-pass filter of Daubechies 9 (db9)



(c) Impulse response of the low-pass and the high-pass filter of Symlets 13 (sym13)

Figure 13: Impulse response of low-pass and high-pass filters used db3, db9, and sym13.

of father wavelets at the j^{th} level) and detail signal D_j (represented by linear combinations of mother wavelets at the j^{th} level) can be created：

$$s = A_J + \sum_{j \le J} D_j$$

The function above signifies that s is the sum of its J-th level approximation A_J improved by the fine details D_js[59].

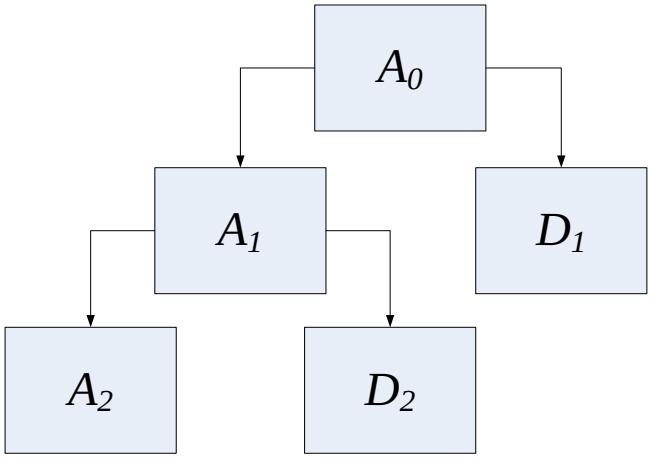


Figure 14: Dyadic wavelet decomposition tree.

3.2 Filter Banks and Selection of Mother Wavelet

Multi-resolution signal analysis using DWT can be implemented by filter banks, where a wavelet and a scaling function are associated with a high-pass and a low-pass filter respectively. As shown in Figure 15, on each level of decomposition, the input signal is split into a lower frequency component and a higher frequency component. With dyadic wavelet filters (wavelet transform), only the low-frequency part is further decomposed. In comparison, binary-tree wavelet filters (wavelet packets), which split both low- and high-frequency components on each level, lead to decomposed signals with an equal bandwidth [50]. In this thesis, only dyadic wavelet filter implementation is discussed.

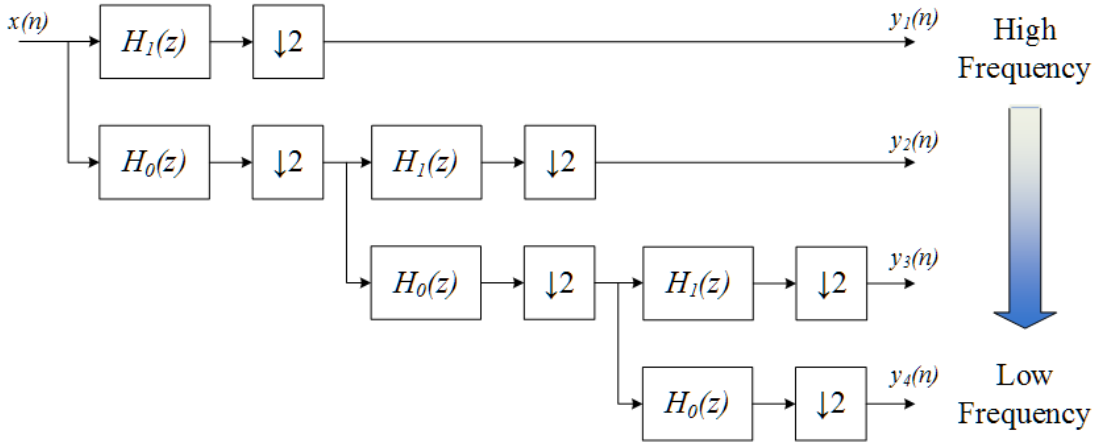


Figure 15: Dyadic filter bank.

The criteria for selecting the mother wavelet adopted in this dissertation are summarized in [60, 61]:

1) The wavelet function should have a sufficient number of vanishing moments to represent the salient features of the disturbances.

2) The wavelet should provide sharp cutoff frequencies to reduce the amount of leakage energy into the adjacent resolution levels.

3) The wavelet basis should be orthonormal.

4) For applications where the information lasts for a very short instant, wavelets with fewer numbers of coefficients are better choices; on the other hand, for signal signature spread over a longer period, wavelets with larger numbers of coefficients tend to show smoother results.

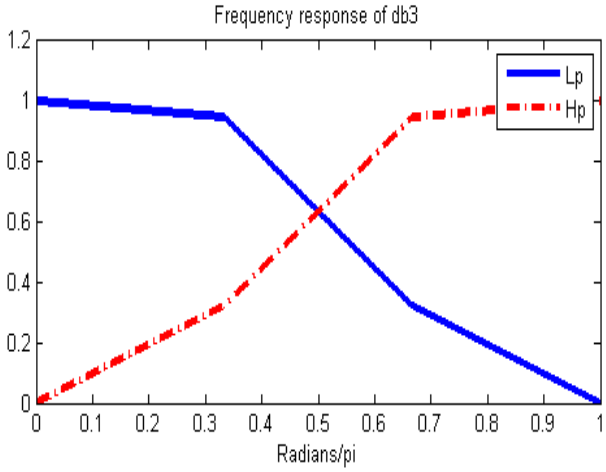
There are several well-known families of orthogonal wavelets named after their inventors. An incomplete list includes Harr, Meyer family, Daubechies family, Coiflets family, and Symmlets family [62]. Daubechies wavelets are chosen in this paper due to their outstanding performance in detecting waveform discontinuities [60, 63].

The frequency response of filter banks of Daubechies 3 (db3), Daubechies 9 (db9), and Daubechies 19 (db19) are shown in Figure 16. It can be seen that the frequency response of db9 filters has a significantly sharper cutoff frequency in comparison with that of db3 filters. But db19 does not provide an equally significant improvement over db9. Considering the extra computational load brought on by wavelets with more coefficients, db9 is a good compromise.

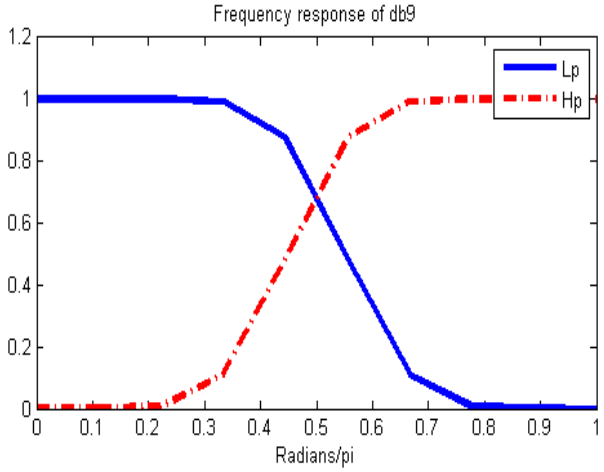
For a dyadic-orthonormal wavelet transform, the first level detail signal has a frequency range of $f_s/4$-$f_s/2$, where f_s is the sampling frequency of the time domain disturbance signal. The second-, third-, fourth-, fifth-, and higher-level signals have frequency ranges of $f_s/4$-$f_s/8$, $f_s/8$-$f_s/16$, $f_s/16$-$f_s/32$, $f_s/32$-$f_s/64$, respectively.

By way of example, consider a DC system in which there is switching noise introduced by the power electronic converters in the system. The original time domain waveform and the spectrum of this signal are plotted in Figure 17(a). Wavelet analysis using db3, db9, and db19 is performed on this signal. The 4^{th} level detail component was designed to span the frequency band 31.3 kHz – 62.5 kHz. The FFT in Figure 17(a) shows that only one harmonic resides in this frequency band and has a magnitude of only about 0.002 Amps. As shown in Figure 17(b), it obvious by inspection that db9 and db19 provide better approximations of the designated frequency band of the original signal than db3, but db19 does not provide a substantial improvement in comparison with db9.

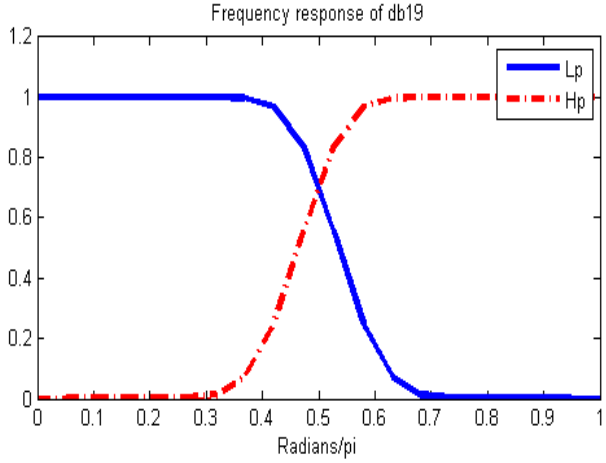
Since the goal of wavelet analysis is to separate the arc fault signal from electronic converter noise (which resides in specific frequency bands) and other electrical disturbances (which usually vary slowly), a narrower transition frequency region leads to less information leaking into other decomposition levels and a more accurate signal approximation. While the db9 and the db19 filter banks are better choices than the db3 filter bank, from a hardware implementation standpoint, db9 filters require less mathematical operations than db19 wavelet. Thus we can tradeoff the accuracy of the wavelet decomposition with processing overhead of the real-time wavelet filter banks implemented in a microcontroller (MCU) or digital signal processor (DSP).



(a) db3

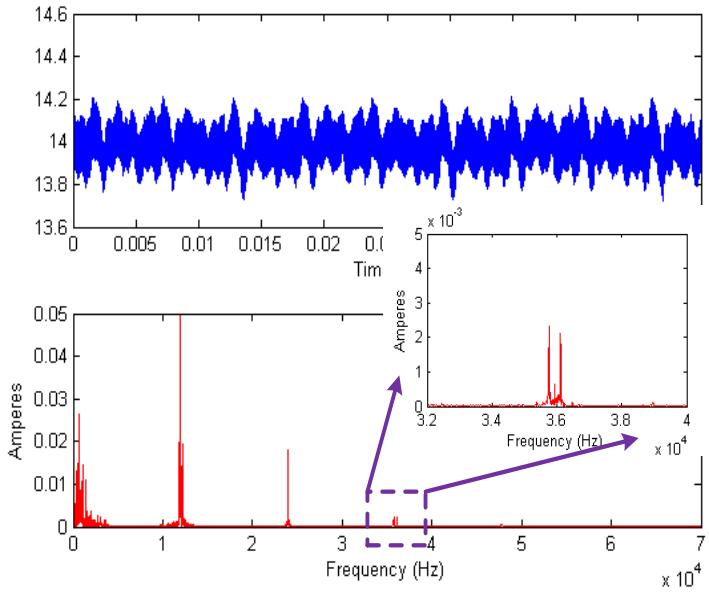


(b) db9

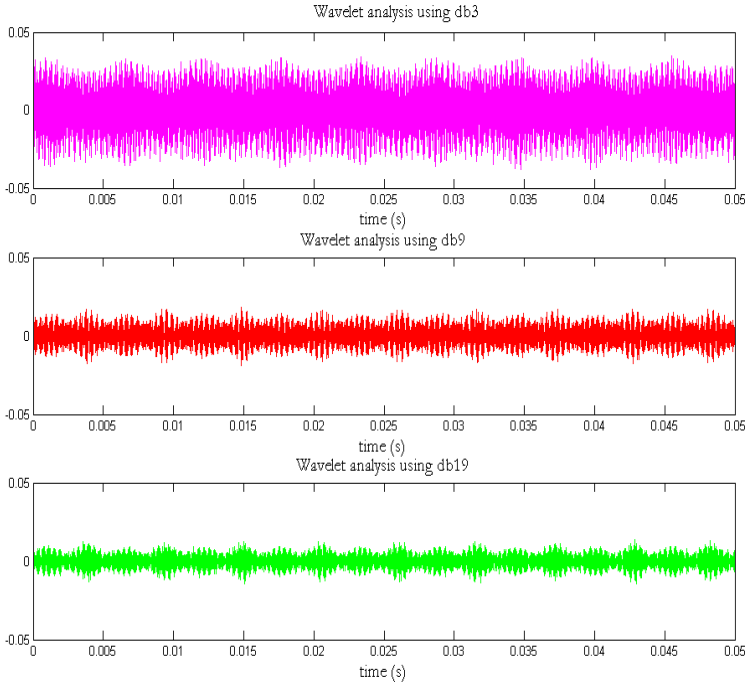


(c) db19

Figure 16: Frequency response of filter banks using db3, db9, and db19.



(a)



(b)

Figure 17: (a) DC system with switching noise: Time trace of the signal (blue); FFT spectrums of the signal (red).
(b)Frequency response of filter banks using db3, db9, and db19.

3.3 Hardware Implementation of DWT

The hardware implementation of a 3-level DWT filter bank is illustrated in Figure 18. In order to implement the process with optimal efficiency, the cascaded filtering process can be reassigned into different sampling periods instead of completing the entire decomposition in a single sampling period.

As shown in the figure above, the digital signal processing device only needs to perform one or zero filtering process in one sampling period with the optimized implementation strategy. Given the length of the filter M (number of filter coefficients) and the total number of decompositions D, the average number of complex multiplications in one sampling period is $(1-2^{-D})\times 2M$, which asymptotically approaches $2L$. The memory space needed for the process is $D\times M$ samples.

In comparison, the total number of complex multiplications needed for radix-2 FFT over a sequence of N complex-valued numbers is $(N/2)\log_2 N$, and the memory space needed would be $2N$. Considering the number of samples needed in STFT to achieve a fine frequency resolution, N is a large number comparing with the filter length M, which makes the FFT not ideal for sample-by-sample real-time implementations.

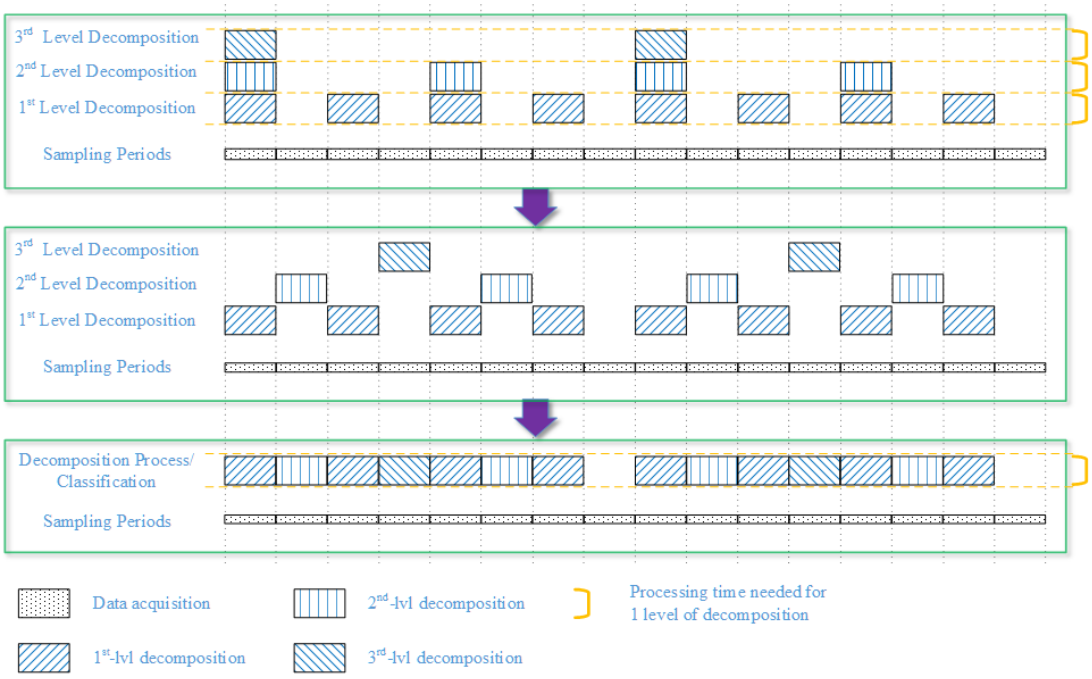


Figure 18: Hardware Implementation of a 3-level DWT.

3.4 Arc Signal Analysis Using DWT - Simulation

To simulate the arcing condition, black box modeling is commonly used to describe the arc interaction with the electrical network. The black box models use voltage and current traces from a circuit breaker test, together with a select differential equation, to produce a mathematical model for the desired arc under study. Most published work using black box models is based on the well-known Cassie and Mayr models. The mathematical model is tuned to a set of measured data by means of a proper selection of arc parameters including the time constant and the cooling power, which is normally taken as a function of arc current and voltage [64].

An example circuit with series arcing is created in Simulink, as shown in Figure 19, which is based on the Cassie arc model blockset developed in [19]. The Cassie arc model is written as a differential equation [65]:

$$\frac{1}{g}\frac{dg}{dt} = \frac{d\ln g}{dt} = \frac{1}{\tau}\left(\frac{u^2}{U_C^2} - 1\right)$$

g the conductance of the arc

u the voltage across the arc

i the current through the arc

U_c the constant arc voltage

τ the arc time constant

The system bus voltage source is comprised of a dominant 1,000V DC component with small-amplitude AC components at 120Hz and 2,000Hz which represent single-

Table I: Parameters used with the Cassie arc model

τ	1.2e-6 s	U_C	100 V
g(0)	1e4 S	Separation time	0.5 s

phase double-frequency power ripple and power electronic switching noise respectively. The sustained series arc starts at 0.5s.

The FFT results are shown in Figure 20. The DC voltage, without arc fault, is shown on the left side of Figure 20 while the voltage with the arc is shown on the right-hand side of the figure. There is no easily distinguishable change in the FFT spectrum from before and after the onset of the arc.

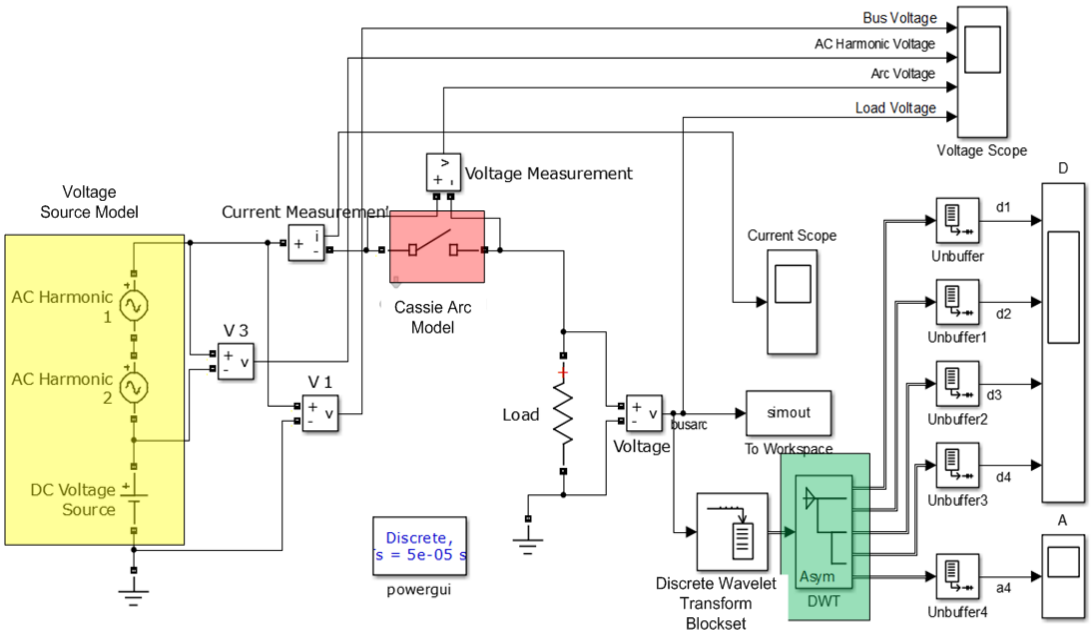


Figure 19: Simulink model.
DC array system with 120Hz double-frequency power line ripple (AC voltage 2), 2 kHz switching ripple (AC voltage 1), and series arcing (Cassie arc model details listed in Table I). The arc model initially behaves as an ideal conductance with the value 1e4 Siemens until the arc "switches on" and then is governed by (4). This simulates the separation of the electrodes that initiates the arc.

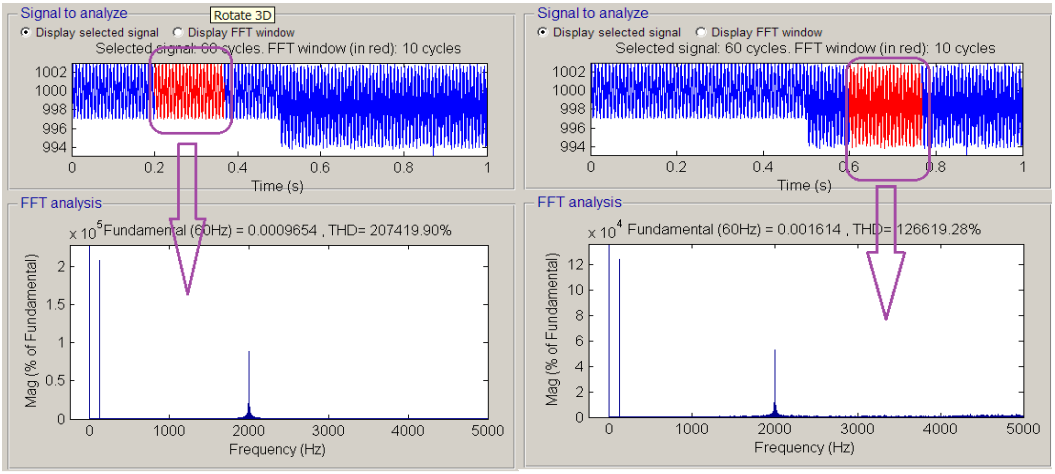


Figure 20: FFT analysis of the synthetic DC voltage.
(a) Voltage FFT signal before onset of arcing (analysis window starts at 0.2s); (b) Voltage FFT after onset of arcing

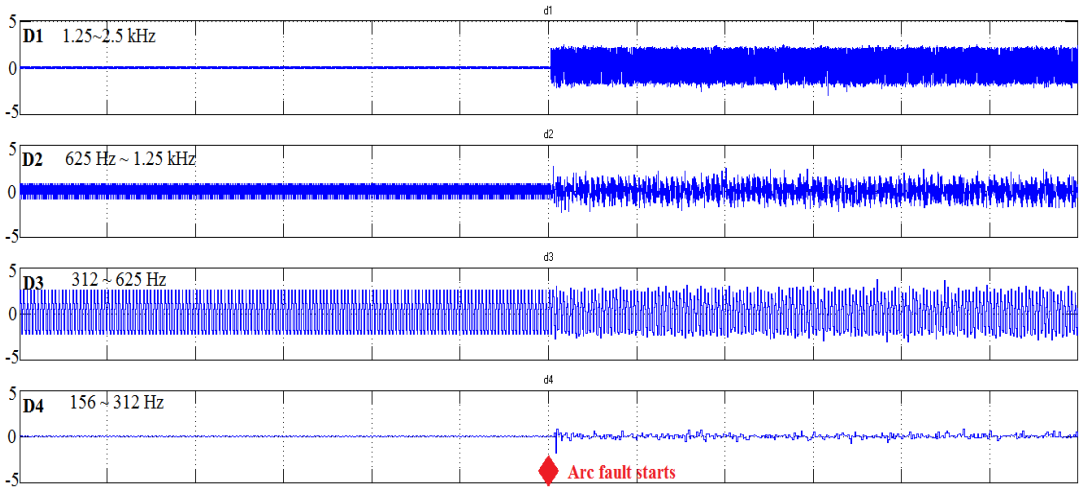


Figure 21: Wavelet decomposition result of synthetic arcing waveform.
(At 0.5s, the switch across the arc generator opens and current begins to flow in the arc. Load voltage signal is processed by wavelet transform by Daubechies 9 as above.)

By contrast, the result of the wavelet transform (Daubechies 9) shown in Figure 21 clearly demonstrates an obvious difference in the wavelet transform from before and after the onset of the arc. This wavelet transform provides an easily observable signal. The chaotic characteristic of arc should be easily detected by DSP/microprocessor programs.

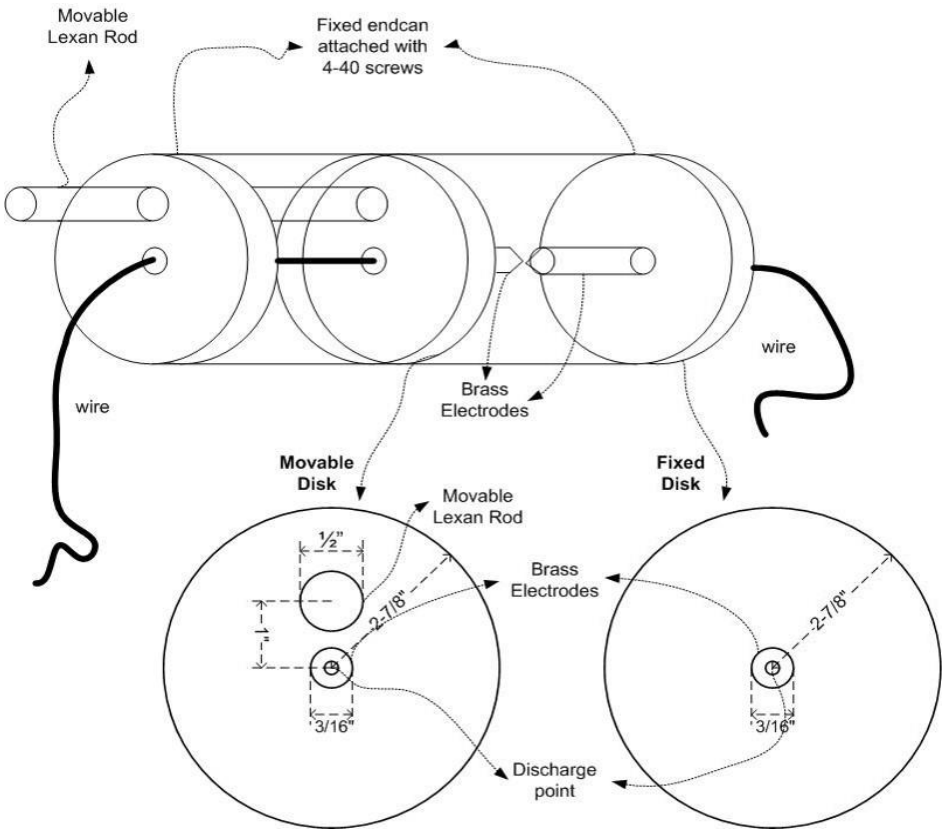
3.5 Arc Signal Analysis Using DWT – Real-World Signal

3.5.1 Experiment Setup and Test Procedure

The experiment performed demonstrates a series arc sustained for about half a second. The arc fault generator (shown in Figure 22) is comprised of two brass electrodes (a stationary electrode and a movable electrode) contained in a polycarbonate tube for safety. The series arcing condition is created initially placing the electrodes in contact and then gradually separating the two electrodes until a desired air gap, in this case 2mm, is achieved. This is the so-called "pull-apart method" for generating a series arc. The system is powered by two PV panels connected in series (details are in Table II). A resistor serves as the load in this experiment. The brass electrodes are polished before every experiment trial to remove the oxidized and melted brass from the previous experiment. Figure 23(a) shows the electrical circuit design; Figure 23(b) shows the experiment station; Figure 23(c) arcing event.

3.5.2 Load Voltage Signal Processing

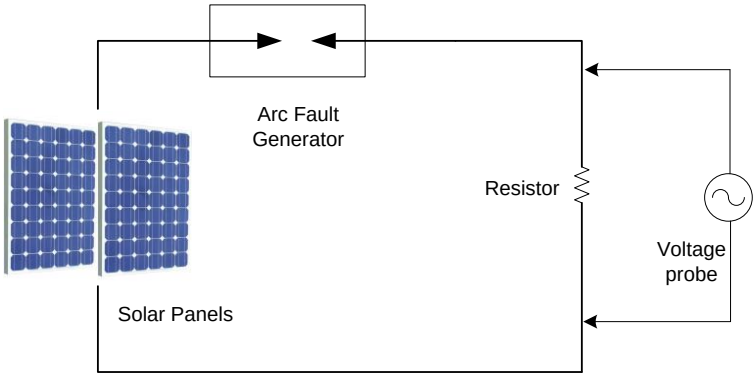
According to some previous research study and arc fault signal analysis on DC systems, the bandwidth for the FFT analysis is adjusted as 1.5 kHz – 45 kHz [52]. Since the sampling frequency of the acquired data is 1 MHz, the frequency band of the 4^{th} – 6^{th} level of wavelet decomposition was selected to be approximately 7.8 kHz – 62.5 kHz. Waveforms shown in Figure 24 are the sensed load voltage signal and the processed FFT



(a) Design of the arc generator

(b) The arc generator

Figure 22: The arc generator made for the arc signal acquisition experiment.



(a) Circuit diagram of the experiment

(b) Experiment Setup

(c) Arc created in the experiment

Figure 23: Experiment setup for arc signal acquisition.

Table II: Experimental conditions

Load type	Fixed resistance: 10 Ohms
DC source type	2 PV panels connected in series
Panel type	1STH-245-WH
Radiance	815 W/m^2
Panel temperature	39℃
Ambient temperature	26℃
Oscilloscope	Tektronix MSO 4034
Sampling rate	1 MHz/S
Bandwidth	350MHz
Memory size	10M points

results for both arcing and non-arcing portions of the waveform as computed using MATLAB. Figure 25 shows the 4^{th} – 6^{th} level of wavelet decomposition using Daubechies 9 mother wavelet.

3.5.3 Result Discussion

From the result in the previous section, it can be seen that it is possible to detect arc faults using FFT, but it is not as significant as using wavelet transform, especially when it comes to the problem for a threshold setting for arc fault determination. Setting the detection thresholds involves consideration of the signal-to-noise ratio, which may change from application to application. However, when using wavelet transformation, the arc signature is significantly distinguished from the non-arcing signal and is easy to be detected when the detection method is embedded in a microcontroller for real-time arc fault detection. Further, since the wavelet transform preserves the time-domain localization information, the precise time of the arc is available for cross-correlation with

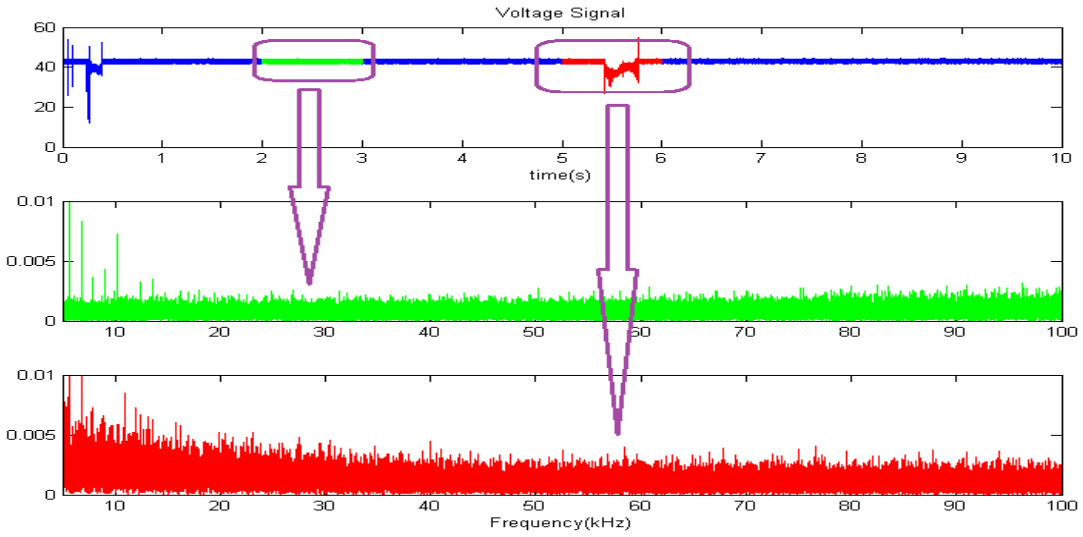


Figure 24: The load voltage and FFT results for non-arcing and arcing part of the signal.

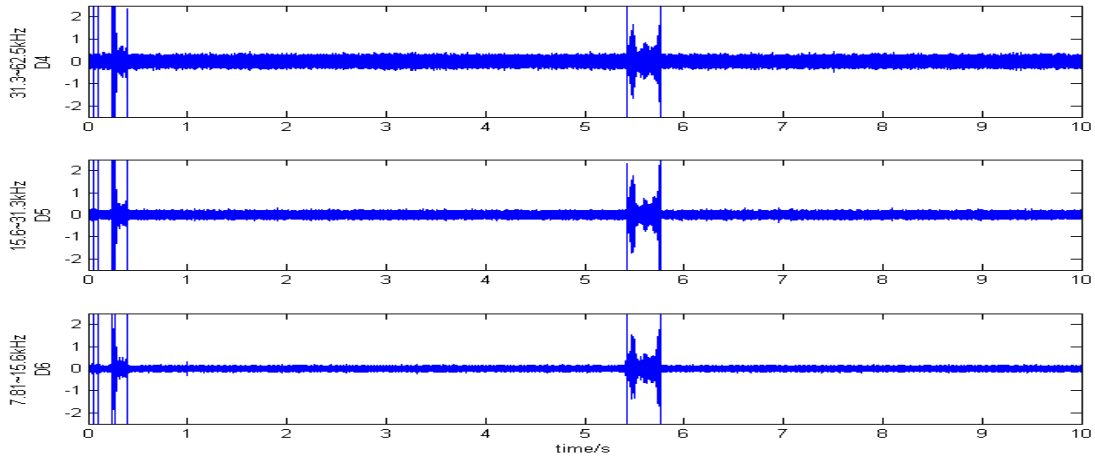


Figure 25: Wavelet transform result.
Top: 31.3 ~ 62.5 kHz; Middle: 15.6 ~ 31.3 kHz; Bottom: 7.81 ~ 15.6 kHz

other system events to improve the confidence of arc fault detection rather than some other benign electrical events.

3.5.4 Signal Analysis Using Composite Signal

To better approximate real-world input for arc fault detectors in PV systems, the arcing information needs to be combined with the real-world background noise. Rather than combining the inverter noise data and synthetic arc signal using simple amplitude superposition, we can combine the signals at specific relative power magnitudes. Knowing the relative strength of one signal versus the other allows one to explore the limits and therefore the range of detectability. We define a metric called the arc-signal-to-noise ratio (ASNR) which determines the proportion of power from each source in the composite signal.

$$ASNR = \frac{P_{arc}}{P_{noise}}$$

Adjusting this user-specified parameter in the synthesizing process enables the synthesis of a family of test signals for validation, sensitivity, and efficacy studies of the detection algorithm based on real-world signals and scenarios. The composite signal is synthesized by using the function:

$$Arc_{composite} = Inv + Arc \times \sqrt{\frac{ASNR_{desired}}{ASNR_{original}}}$$

The performance of the FFT arc detection method compared to the wavelet decomposition method are further studied using the synthetic waveforms.

In this work, a test signal is synthesized using time-domain inverter noise signal data measured from a PV array and time-domain arc signals obtained from an arc generator. Once these synthetic signals are created with specified ASNR levels, discrete

Fourier transform and discrete wavelet transform are comparatively studied. The influence of sampling frequency on the two analysis approaches is examined. The wavelet transform analysis with distinct types of wavelets are also evaluated and compared. Some analysis results are shown below.

3.5.4.A. Composite Signal with High-Rate Data (F_s=1MHz)

A composite signal with a duration of one second is synthesized by combining inverter noise and arc fault signals at a sampling rate of 1MHz to achieve an ASNR of 0.1. FFT analysis, shown in Figure 26, is first performed on the entire one-second sample (second from top), the non-arcing portion (third from top) and then the arcing portion (bottom) of the waveform. The strong presence of the inverter switching frequency and harmonics appears to overshadow the arc noise, making detection difficult.

By contrast, the 7^{th} decomposed signal (covers the frequency band of 3.9 kHz – 7.8 kHz) from the wavelet transform is selected. Different decomposition results using db3, db9, and db19 are shown in Figure 27. The temporal waveforms for the selected frequency band clearly indicate the causality and timing synchronization of the initiation and extinction of the arc.

3.5.4.B. Composite Signal with Downsampled Data (F_s=100 kHz)

The composite signal from part *A* is downsampled by a factor of 10 to produce a composite signal with a sampling rate of 100 kHz and a total of 100k sample points for the one-second signal. The FFT and wavelet analysis results are shown in Figure 28 and

Figure 29 respectively. The decomposed signals cover the band of 3.125 kHz – 6.25 kHz are selected.

3.5.4.C. Result Comparison

From the FFT analysis results shown in Figure 26 and Figure 28, it is difficult to find any significant detectable arc fault features by comparing the FFT result of the non-arcing part and the arcing part of the signal, especially when the sampling rate is decreased (Figure 28). Slight differences do exist between the two spectral analysis graphs, but the fault detection threshold can be very difficult to select, particularly if a detection technique using limit-lines is used. Detection threshold setting involves consideration of the signal-to-noise ratio, which may change from application-to-application. Selecting a threshold without delicate calculation and thorough understanding of the system behavior would lead to not triggering or false triggering of the protection mechanism.

However, from the wavelet analysis plots, not only arc features can be easily distinguished from the non-arcing signal, but the exact moments when the arc fault ignites and extinguishes can also be observed. This facilitates selection of a detection threshold for an embedded microcontroller for real-time arc fault detection. It also enables characterization of the arc event.

By comparing the analysis result using three different Daubechies wavelets, we can conclude that decomposition results using db9 and db19 are significantly better than using db3. But db19 doesn't provide much improvement to the result of db9. This is consistent with our frequency response analysis of the filter banks. By taking the DSP

computational load into consideration, db9 is a good compromise between calculation speed and decomposition quality.

As shown in part A and part B, the sampling frequency has a significant impact on both Fourier and wavelet detection approaches. With the signal sampled at 100 kHz, it is almost impossible for the Fourier transform to capture any arc fault features. While the sustained presence of the arc is not as obvious as when the sampling frequency is 1MHz for wavelet decomposition, we should still be able to draw enough information to detect the arc fault. Thus, detection approaches based on wavelet can use a lower sampling rate than Fourier transform to accomplish accurate arc fault detection if indeed Fourier methods can accomplish it at all.

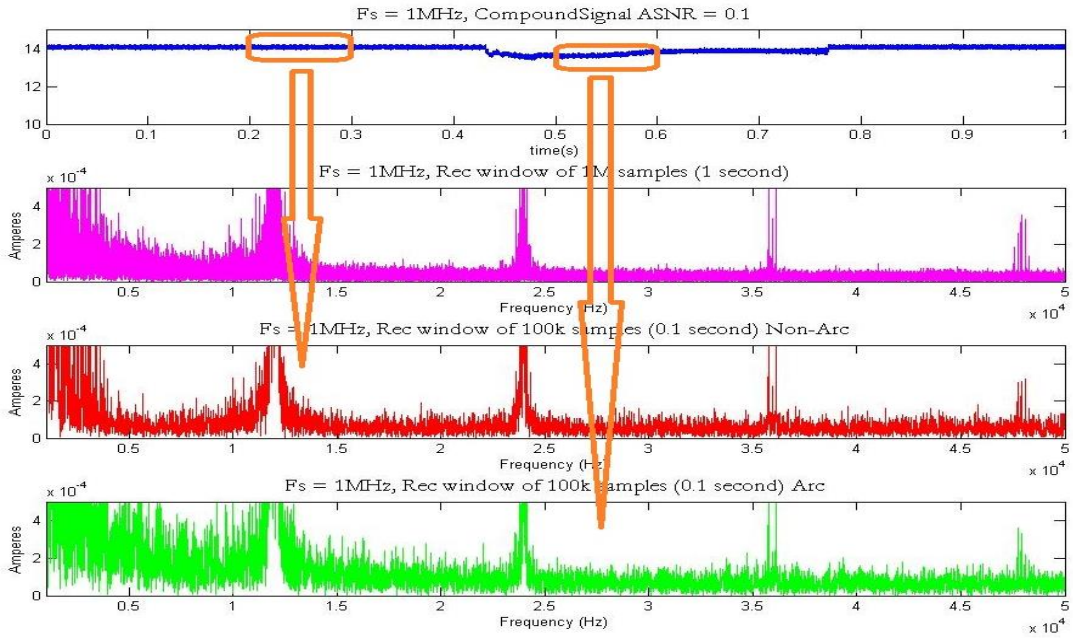


Figure 26: Composite signal (Fs = 1MHz); FFT analysis of the entire composite signal (red); FFT of the non-arcing part of the signal (red); FFT of the arcing part of the signal (green).

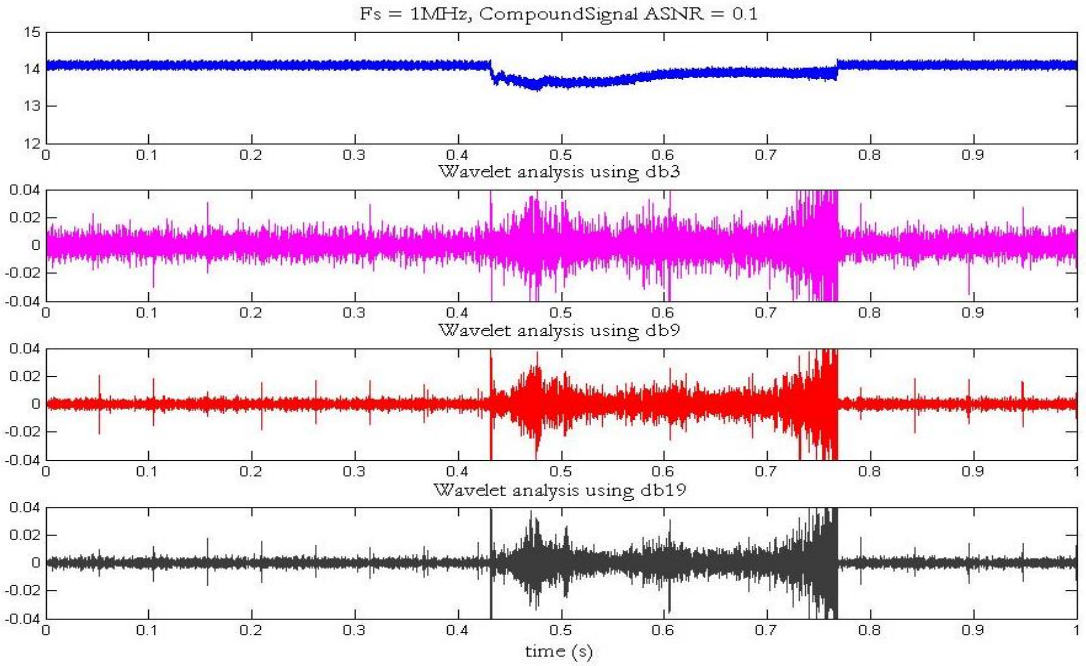


Figure 27: Wavelet analysis (db3 – magenta, db9 – red, db19 - grey) of the composite signal (Fs = 1MHz).

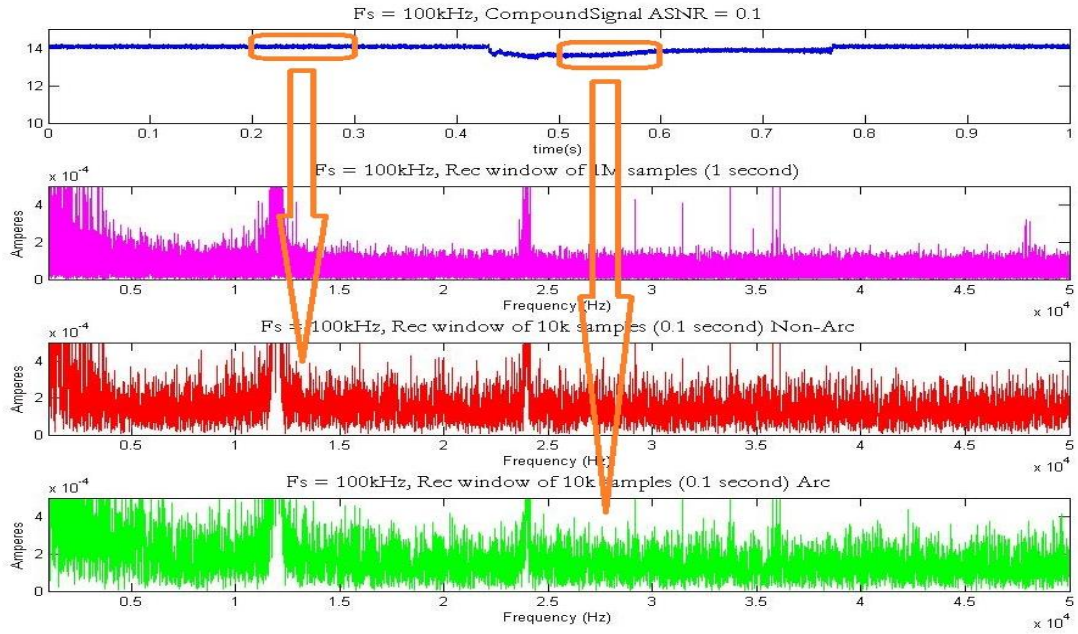


Figure 28: Composite signal (Fs = 100kHz); FFT analysis of the entire composite signal (magenta); FFT of the non-arcing part of the signal (red); FFT of the arcing part of the signal (green).

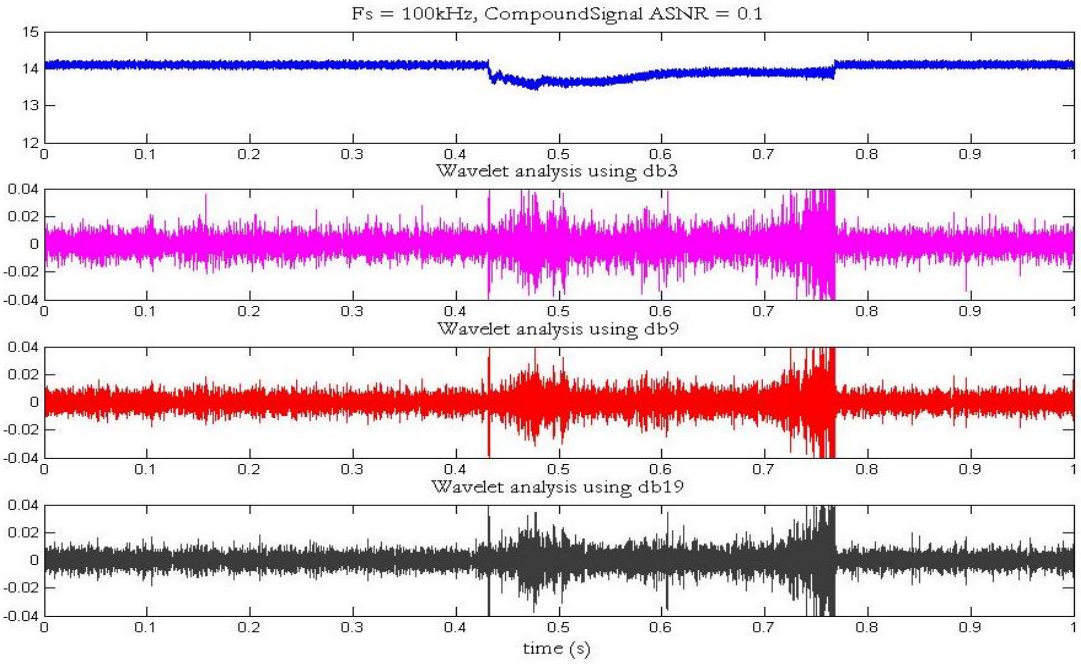


Figure 29: Wavelet analysis (db3 – magenta, db9 – red, db19 - grey) of the composite signal (Fs = 100kHz).

3.6 Revised Arc Generator

The arc generator (AG) is a platform for producing arc faults for research, testing, and standardization for the PV industry. Previous drafts of the standard by the UL 1699b committee included a “steel wool” method to initiate the arc, which works for producing both series and parallel arc faults [43]. One problem with this approach is that the initial arc current must flow around the AG through a different path and then switch to the AG which could alter the system operation (such as maximum power point tracking) or introduce noise into the system that is not part of the arc itself. This can create a false detection or mask detection of the true arc. Another problem is that the steel wool was found to dominate the initial arc characteristics [43]. As such, the device under test (DUT) affectionately became known as a “steel wool detector”.

Another approach to creating an arc is by using a pull-apart method in which two electrodes are initially in contact, establishing the current path, and then separated thereby creating an arc [43]. Although this method only works for series arc faults, it does not depend on an ignition medium. The test bed presented earlier in section 2 of this dissertation was the first attempt. It consisted of brass electrodes and a polycarbonate tube with one fixed and one moveable plate. While it produced good arc results, there were several design limitations including limited thermal dissipation and manual electrode separation that limited it from becoming a robust testbed. Repeated arcs gradually ablated the profile of the electrodes, which alters the electric field, and hence could change the signature of the arc [66]. Further, sustained arcs, or high repetition rate of arcs, generate high temperature which melted the AG as shown in the figure below. Thus a new AG was

needed with improved electrodes, capability to dissipate heat generated from the arc and to provide precise and automatic control of the electrode separation.

In order to overcome the limitations encountered with the previous design and to add new features such as the ability to create a horizontal or vertical arc, a new AG was developed from the ground-up. The new AG allows for a variety of electrodes with varying geometries to be used. The electrodes are mounted on a linear actuator to precisely control their separation distance, velocity, and acceleration; the linear actuator is driven by a stepper motor which is controlled via a microcontroller and computer interface; and the entire assembly is contained within an enclosure for safety and to eliminate arc fluctuations from air movement.

Figure 30: Melted mounting plate on the first generation AG.

3.6.1 Arc Generator Enclosure

As previously mentioned, the first generation AG was limited by its ability to dissipate heat. The enclosed volume of the first generation AG was too small to reject heat into the ambient space, and there was not a sufficient heat sink to conduct heat from the electrodes. Thus the small volume, coupled with the low melting point of polycarbonate, resulted in a melted mounting plate during prolonged arc generation.

The new AG was designed with a substantial increase in ambient space. The first generation AG had approximately 150 in3, and the new AG has approximately 3,500 in3. The figure below displays a picture of the new AG with the base and lid. The base is L-shaped in order to allow the AG to be positioned in such a way that the electrodes are separating either horizontally or vertically

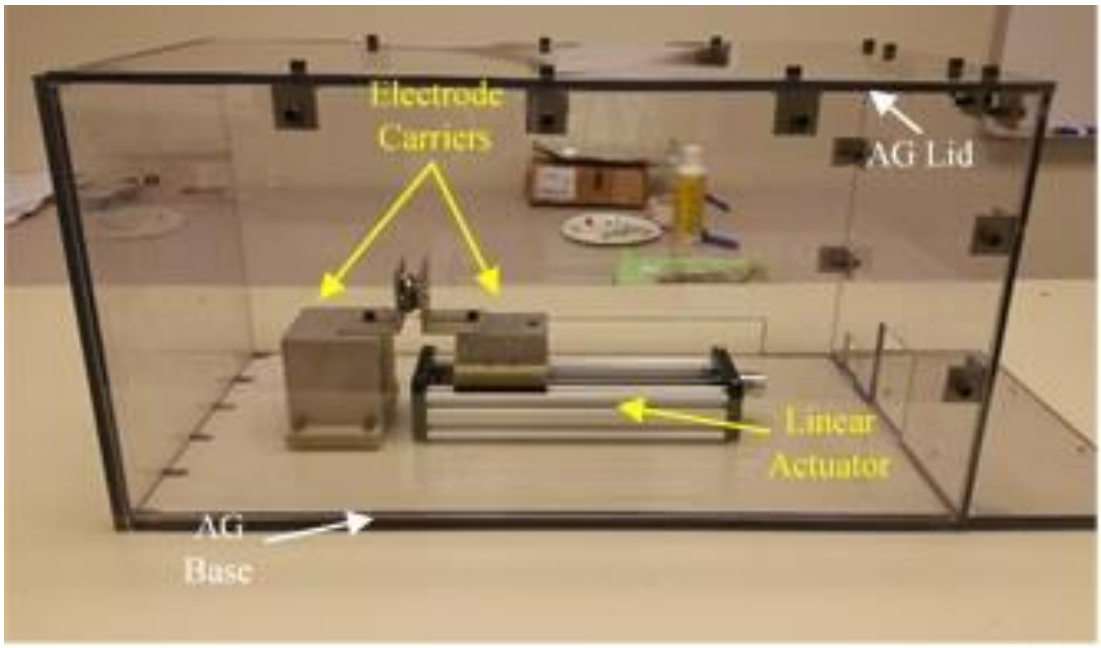


Figure 31: The enclosure of the new AG.

3.6.2 Electrode Carrier

The electrode carriers of the first generation AG were the major failures of the first generation AG. During arc generation testing where sustained arcs were repeated at a high rate or for prolonged periods, the electrodes increased to a temperature beyond the melting point and damaged the polycarbonate mounting plate. In order to prevent this type of damage occurring to the new AG, aluminum was chosen to mount the electrode carriers.

An important safety feature of the electrode carriers is electrical isolation. If the electrode carriers are electrically isolated from the rest of the AG, then the AG becomes safer and more reliable by reducing shock hazard to humans and chances of component damage. An isolation voltage of 900V was specified to provide a 50% safety margin because typical grid-tie systems are rated at a maximum of 600V DC [46]. According to McMaster-Carr, the dielectric strength of the electrical insulator chosen (Garolite) is 350

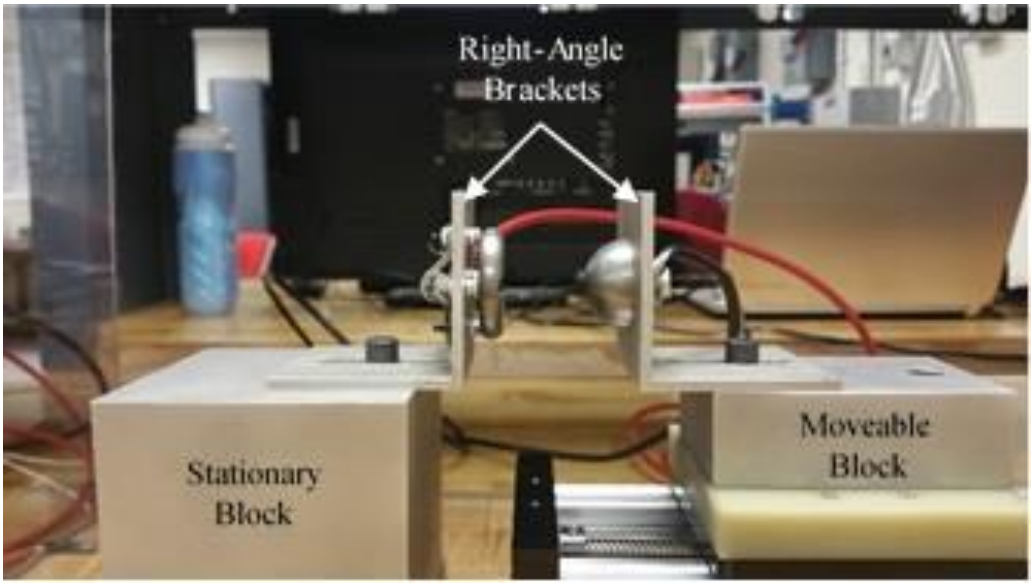


Figure 32: The electrode carriers.

V/.001". The Garolite used in the new AG is 0.5 inches thick, thus resulting in insulation capabilities of up to 175 kV (see Figure 33). This is well above the specification however, a thicker slab of Garolite was needed to prevent damage to the slabs during fabrication

A requirement for the new AG was electrode interchangeability and rotational adjustment. Since the new AG will (ideally) be used for UL standards development, it would provide robustness if the electrodes could be easily interchanged. The following figure displays an image of the ball-and-ring electrodes.

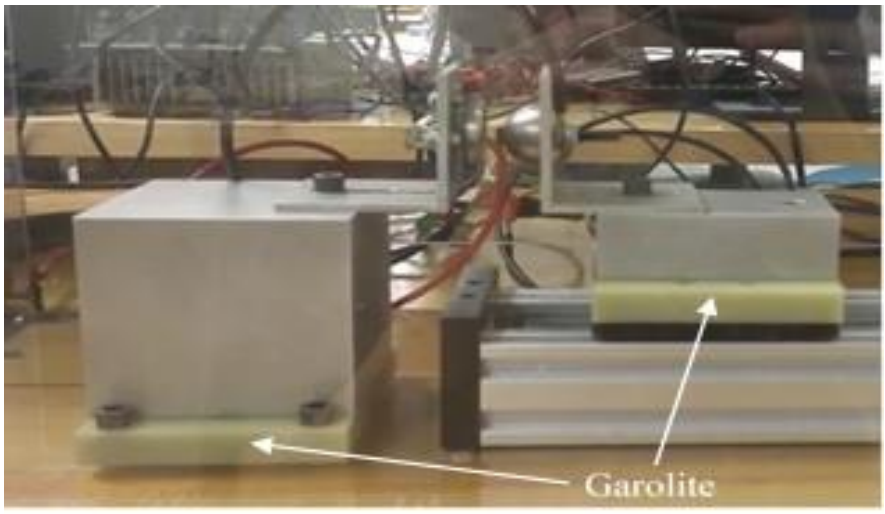


Figure 33: Garolite voltage insulators.

3.6.3 Motor Control

One of the issues encountered with the first generation AG was manual electrode separation. Although this method was easy to perform, it is not precise nor was it automated. This created inconsistencies in the resulting arc signatures because the separation speeds for each test was impossible to reproduce perfectly. To combat this

issue, a stepper motor, controlled by a microcontroller-computer interface, was used to automate the process of electrode carrier separation.

The Texas Instruments' DRV8412 evaluation kit was used for this application. Relatively low-cost, this evaluation kit provides the user with plug-and-play capabilities that does not require much in-depth knowledge of stepper motor programming.

The kit comes with a NEMA-23 stepper motor, two brushed DC motors (although included with the kit, they were not used in the presented AG setup), motor driver, C2000 Piccolo F28035 MCU controlCARD, and motor control graphical user interface (GUI). More information about the DRV8412 evaluation kit may be found at http://www.ti.com/tool/DRV8412-C2-KIT.

A picture of entire AG setup is shown in Figure 33.

The full bill of materials (BOM) for the AG presented in this project is provided in Appendix B.

3.6.4 Lessons Learned

A number of issues were encountered with the AG during initial experimentation. The most prominent issue was the rigidity of the electrode carriers. The right-angle brackets were not mounted firmly enough, thus causing the brackets to move upwards relative to their axis of movement. This created inconsistencies in the number of motor steps required to generate an arc. To fix this problem, additional fastening points were added to the right-angle brackets to counteract the issue.

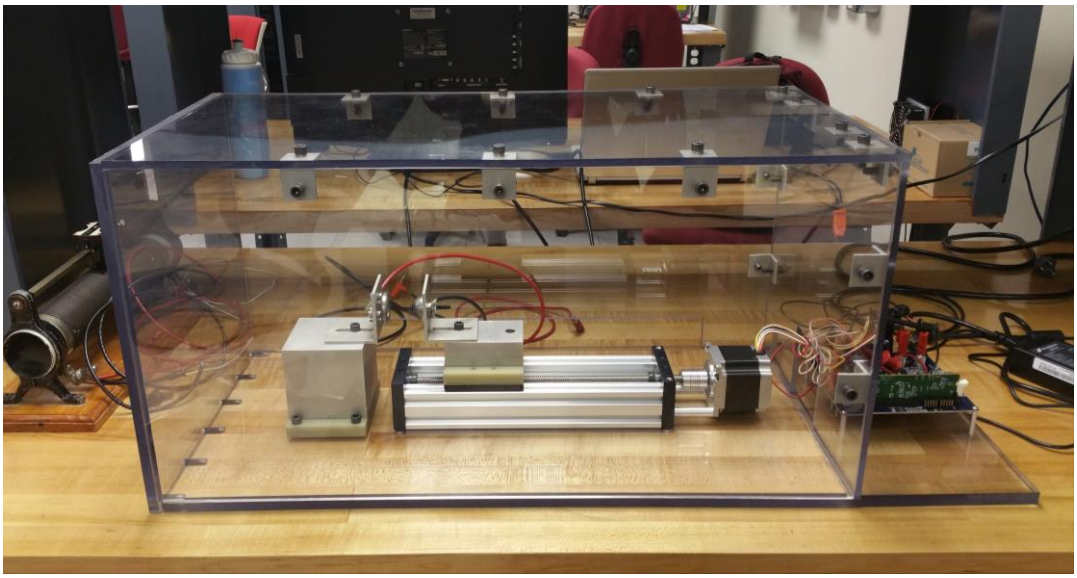


Figure 34: The revised arc generator.

Another issue involved the electrical connections of the electrodes. The ball-and-ring electrodes are electrically connected by wrapping wires around the screws used to mount the electrodes to the right-angle brackets. Although this method is capable of adequate current flow for low-power arc generation, it is rather crude and could be improved. The ring electrode severely limits the options for electrical connections because the mounting screws are so small. Different choice of electrodes could provide more flexibility for electrical connections.

3.7 "Replay" Arc Fault Detection Testbed

As an investigation and verification practice, arc generation using arc generator connected to a PV system is a common method used to study the characteristics of an electric arc [20, 45, 66-71]. However, the amount of effort required for the arc generation practice and the uncontrollability of arc power impede researchers to develop accurate detection algorithms. Controllable recreation which can genuinely represent the arc signal is a prerequisite to a successful detection: accurate evaluation of a valid arc and minimized the chance of nuisance tripping.

In this work, a testbed is developed for the recreation of prerecorded real arc signals. The input to the system is a prerecorded signal acquired at a sufficiently high sampling rate. The desired output should be an analog signal that carries the same voltage/current of the original arc signal, and therefore, can be viewed as a truthful representation. An ideal replay system could enable a highly automatable and reproducible system for validation and efficacy studies of the detection algorithms based upon database with arcing and non-arcing events from different realistic fault scenarios.

3.7.1 System Configuration

As shown in Figure 35, the testbed system created in this work is comprised of a computer, a digital-to-analog converter (DAC), a resistive load, and an oscilloscope. Considering the DAC device might not be able to provide sufficient output power to recreate the original signal, an extra power amplifier is added to the system.

There are several concerns need to be addressed when selecting equipment.

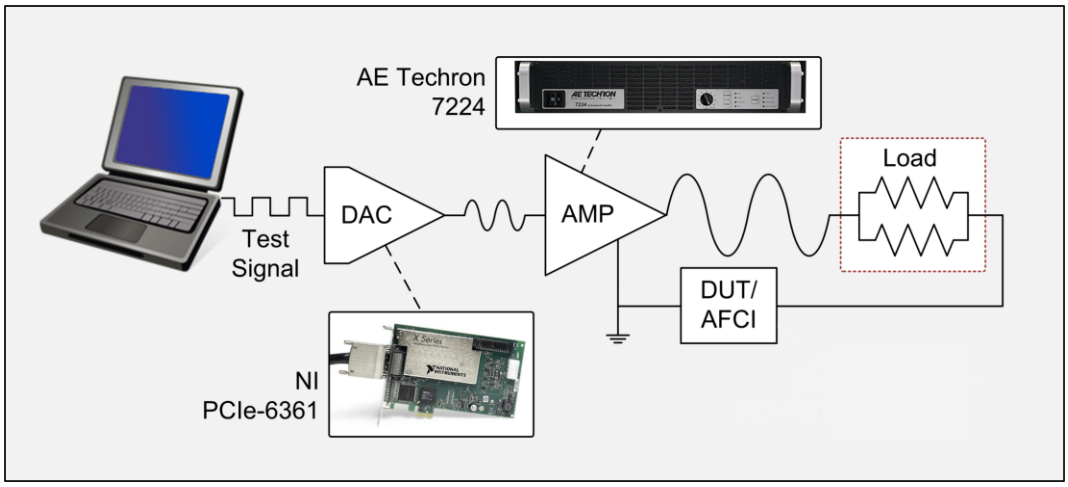


Figure 35: Structure of the testbed.

1) The output rate of the DAC should be high enough to maintain the true frequency characteristics of the original signal. If the sampling rate of the DAC device is not high enough, the prerecorded signal will have to be down-sampled. Therefore, not only is the information in the higher frequency bands lost in the process, but also the lower frequency components can be contaminated because of the aliasing effect.
2) The chosen DAC device would need to have a high enough bit resolution to truthfully (to some extent) represent the amplitude and the minor variations in the original signals.
3) The power amplifier would need to have dc output capability with output power rating higher than the prerecorded signals.

In this work, PCIe-6361 from National Instruments and AE 7224 from Techron are used as the DAC and the power amplifier respectively. The NI PCIe-6361 offers up to

2.86 MS/s output sampling rate for 16-bit data. The output voltage range is ±10 V. The AE 7224 is capable of amplifying the voltage up to 140 V with its maximum output power at 1.2 kW.

The input to the testbed is prerecorded current signal sequence acquired from non-arcing and arcing events. The digital sequences are converted into analog signals in the DAC and then amplified by the power amplifier. A 5Ω wire-wound resistor is used in the testbed as the resistive load. By tuning the rotary knob on the power amplifier, the output voltage of the power amplifier can be modified in order to achieve a truthful representation of the original signal. The commercially available arc fault detector (AFD) RD-195 from Texas Instruments is used as a DUT example. Waveforms at each stage of the replaying process can be observed and recorded on the oscilloscope. The actual lab setup of the testbed is shown in Figure 36.

3.7.2 System Validation

Arc signals have a wide frequency spectrum. Both low and high frequency characteristics can be used for evaluation and detection of arcing events. Thus, the frequency response of the testbed is expected to be constant within the band of interest, which is considered to be dc to 100 kHz in this work. Therefore, an accurate replay system is needed to ensure minimum distortion is introduced into the reproduction. A frequency sweep from 10 kHz to 160 kHz with 10 kHz interval on the input side is conducted for frequency response measurement. The input signal of the frequency sweep is generated in MATLAB, and the output of the DAC and the power amplifier are observed on the

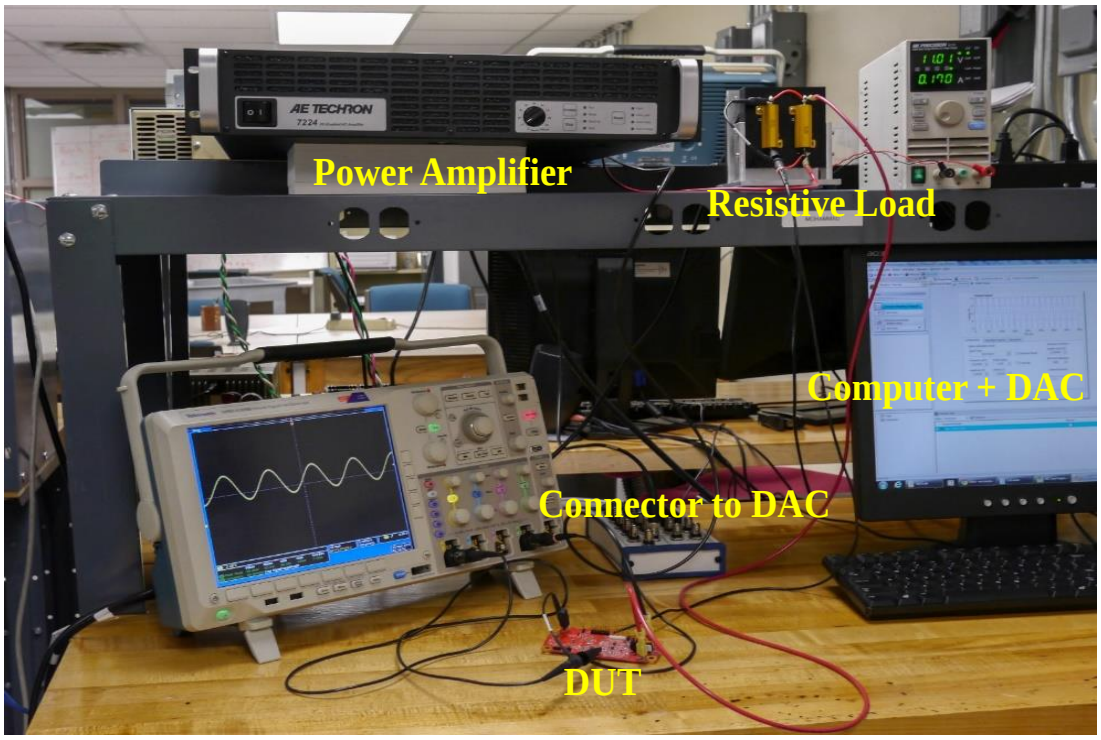


Figure 36: Picture of the testbed.

oscilloscope. To ensure the smoothness of the digitized sine waves, the sampling rate of the input signal is set at 2.5 MS/s, which is much higher than the upper bound of the frequency sweep. The sampling frequency of the oscilloscope is 1 MS/s.

Figure 37 shows the frequency response of replay system. Each test is performed three times and the result shown here is the average of the three trials. It can be seen from that the magnitude ranges from 0.978 to 1.005. Taking into consideration the possible measurement error, the magnitude response of the testbed is considered invariant within the band of interest. In the meantime, a linear phase where the phase response of the system is a linear function of frequency. The result is that all frequency components of the input signal are shifted in time (usually delayed) by the same constant amount. And

consequently, there is no phase distortion due to the time delay of frequencies relative to one another.

Besides the desirable frequency response for the steady-state signals, the accurate and fast response is also required since arc signals exhibit abrupt changes and chaotic behavior. The transient response properties include rise time, overshoot, settling time, delay time, peak time, and steady-state error.

Considering the possible measurement error from the oscilloscope, the testbed is believed to have an invariant magnitude gain and constant time delay within the frequency band of interests. Therefore, the testbed should have high fidelity reproduction capability-fast response with minimum magnitude and phase distortion is introduced for signal replay.

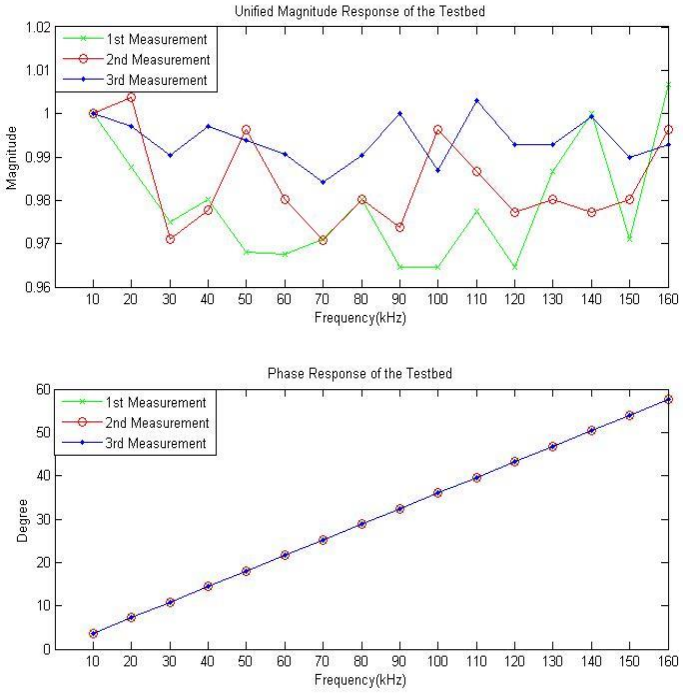


Figure 37: Frequency response test result for the replay system validation.

3.8 Real-Time Arc Fault Detection Using Wavelet Decomposition

Prior theoretical studies have implemented the DWT algorithm on prerecorded and synthesized arc signals. The results have demonstrated the potential of DWT for arc fault detection in PV systems. However, the reduction to practice into cost-effective real-time arc fault detection hardware, and the challenges involved in doing so, remains an open issue along with a direct comparison of hardware-based DWT to DFT. This section presents the results of developing a real-time arc fault detector (AFD) which uses DWT to analyze the measured PV current and a detection algorithm based on the power of the decomposed signal.

The proposed DWT method was tested using prerecorded signals acquired by Sandia National Laboratories from real direct-current photovoltaic (DCPV) systems with operating inverters. Subsequent testing compared the proposed DWT AFD to the TI RD-195 evaluation module, which is a commercially available AFD which adopts an FFT approach. All test signals here were prerecorded. And they are relayed using the replay testbed introduced in the previous section.

3.8.1 Hardware of the Arc Fault Detector

The RD-195 dc arc detection evaluation board from Texas Instruments is a commercially available AFD built on a C2000 DSP platform which adopts an FFT-based detection method. In order to implement the FFT, the PV current is sampled and block-processed on a frame-by-frame basis. The system is shown in Figure 38.

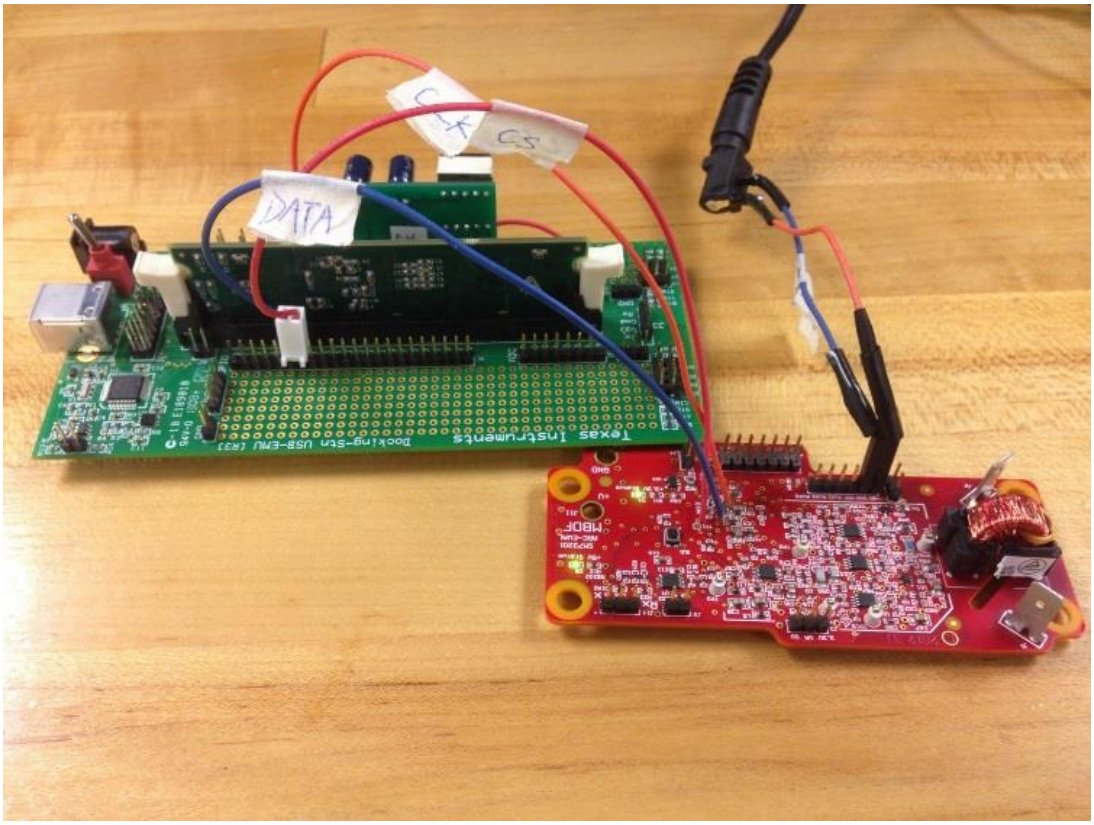


Figure 38: Picture of the testbed.

Sold as an evaluation kit, users are able to customize parameters such as sampling frequency, frame size, windowing technique, frequency band, and detection threshold value. Since this off-the-shelf hardware is specifically designed for acquisition of DC current and arc fault detection in dc systems, it is chosen in this project for rapid prototyping and verification of the proposed DWT-based arc fault analysis method.

A slight modification was made to upgrade the DSP from the original TMS320F29033 to TMS320F28335 while remaining in the C2000 product family. This upgrade enables a faster CPU clock rate which leads to a sampling rate up to 200 kHz while executing DWT filter bank in real-time on a sample-by-sample basis.

3.8.2 Threshold Setting Based on Power Ratio

From the wavelet decomposition results in the previous sections, it was found that the signal power from the 1st-level high-pass filter, which represents the frequency components roughly between 50 kHz and 100 kHz, demonstrates a noticeable increase when arcing occurs. In order to avoid false detection caused by accidental spikes in the signal, the detection decision is based on the average power of 128 consecutive samples in the first decomposed detail signal (*D1*). The signal power computation flow is illustrated in Figure 39. By calculating and comparing the average power of the processing frame and the average of power of a non-arcing frame, the threshold ratio between the two is chosen to be 1.4. That means:

$R = P_{processing}/P_{reference}$

If $R > 1.4$ ➔ the frame being processed is determined to contain a potential arcing event.

3.8.3 Implementation of DWT-Based Algorithm

On the RD-195 evaluation kit, the current signal is sampled using the external SM73201 ADC connected via SPI to the MCU. Figure 40 shows the overall signal flow of the wavelet detection process. The sampling frequency of the ADC is set to be 200 kHz by an interrupt routine triggered every 5 μs which initiates data conversation of the ADC. An SPI interrupt routine is triggered when the conversion is complete and is followed by DWT for the new sample. The MCU is programmed to calculate the DWT result of 128 consecutive samples per frame.

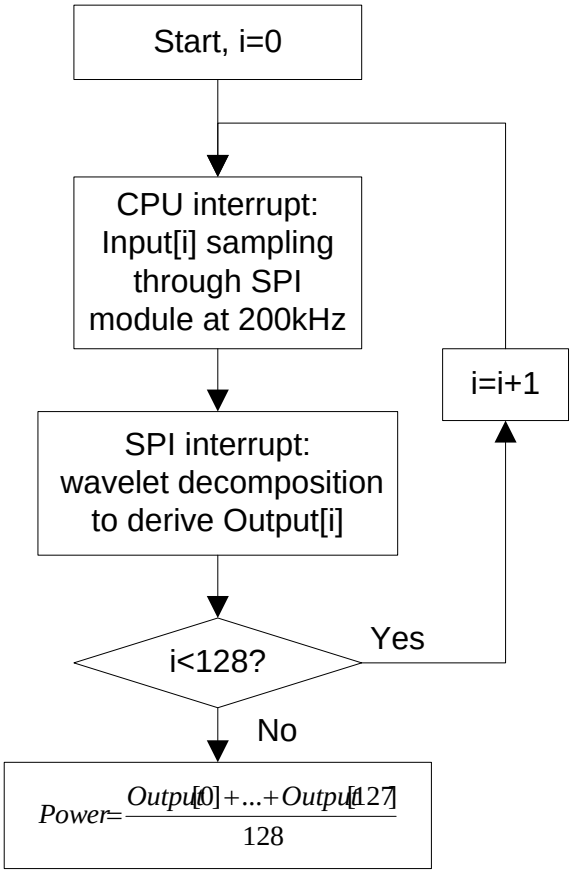


Figure 39: Signal power computation for each detection frame.

The time needed for the entire detection process has to be managed less than the sampling period (5 μs) in order maintain the sampling rate at 200 kHz. Data transmission through SPI takes 3 μs, which leaves us 2 μs to run the decomposition and detection algorithm. According to the implementation strategy in Section 3.3, one level of high-pass filtering and low-pass filtering should be accomplished in 2 μs. Processing time and corresponding bandwidth for a list of Daubechies (Daubechies 3 to Daubechies 9) wavelets are shown in Table III. It can be observed that decreasing the length of the filter

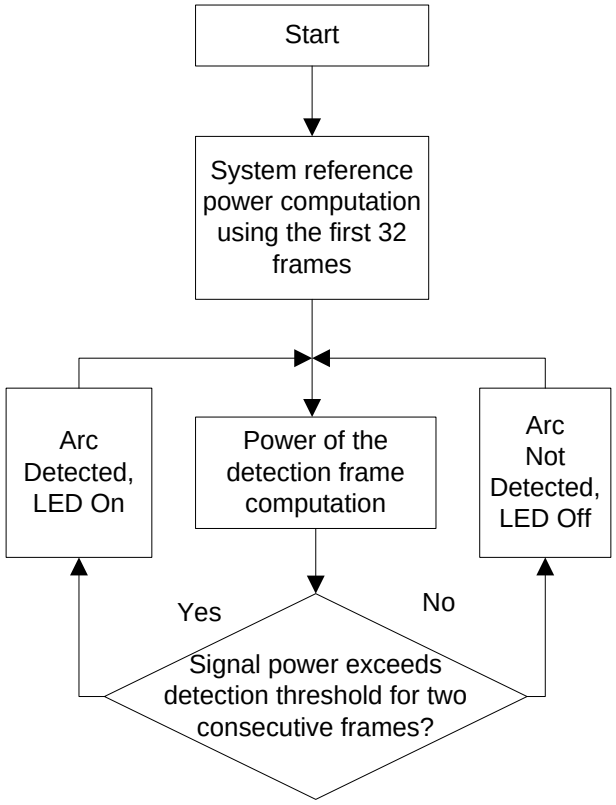


Figure 40: WT based arc detection program signal flow.

Table III: Processing duration and visible bandwidth for coefficients db3-db9

Selected Wavelet	db3	db4	db5	db6	db7	db8	db9
Time Needed for DWT (us)	0.6	0.72	0.84	1.02	1.16	1.32	1.46
Bandwidth (kHz)	138.9	127.6	130.2	124.4	120.2	116.7	112.1

does not boost the sampling frequency significantly. Thus, when it comes to the wavelet selection, the capability of extracting arc information would be our major consideration.

3.8.4 Implementation Result

As illustrated in Figure 41, the high frequency feature of the inverter noise looks similar to the arc signal. Therefore, to demonstrate that the information extracted from the wavelet transform are effective, both inverter noise and inverter + arcing events are tested here. Five sets of inverter noise and inverter + arcing signals are used as the test signals. Each set of inverter noise and inverter + arcing signal are acquired from the same system.

By using the inverter noise signal as a reference, oscilloscope screenshots in Figure 42 verified that with a little bit of threshold tuning, the information extracted from DWT could be effective even by using a simple detection method. Although no false tripping is observed, the algorithm might still be not robust enough since some parts of the arcing events are not detected.

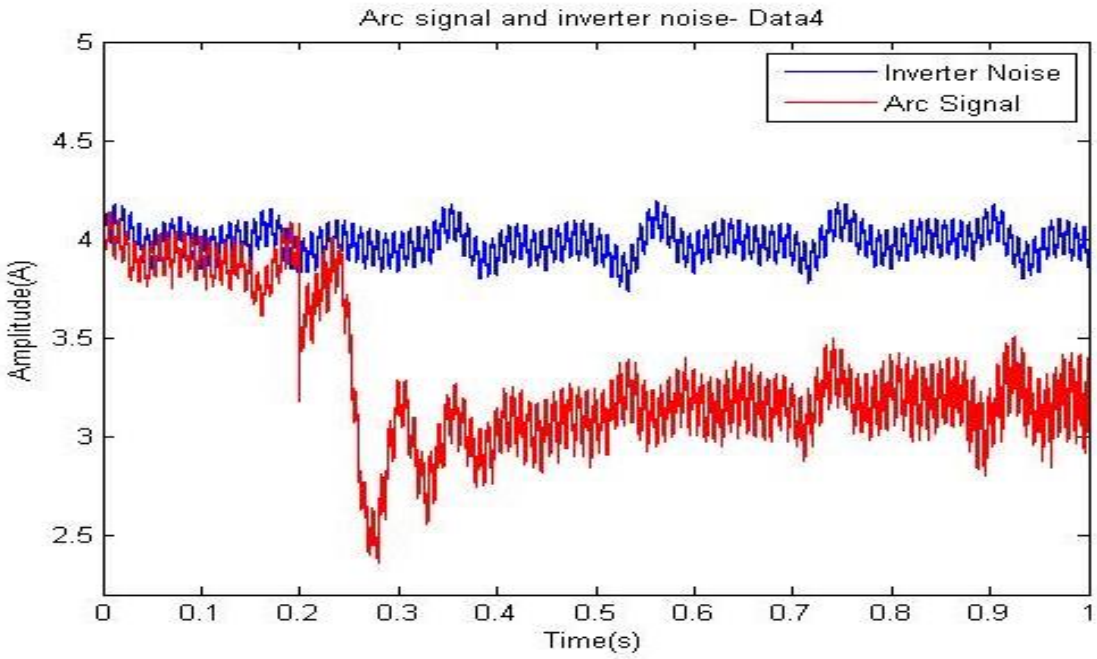
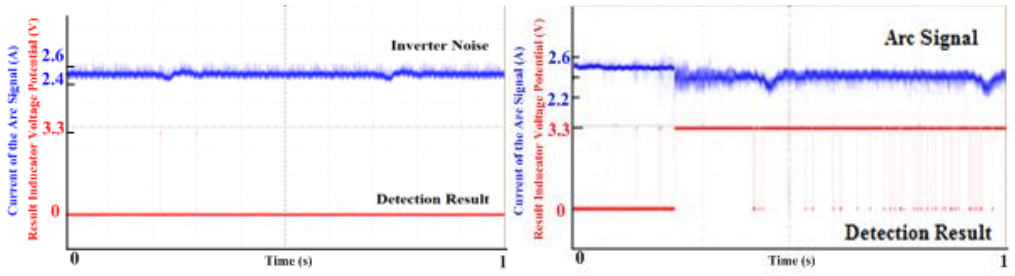
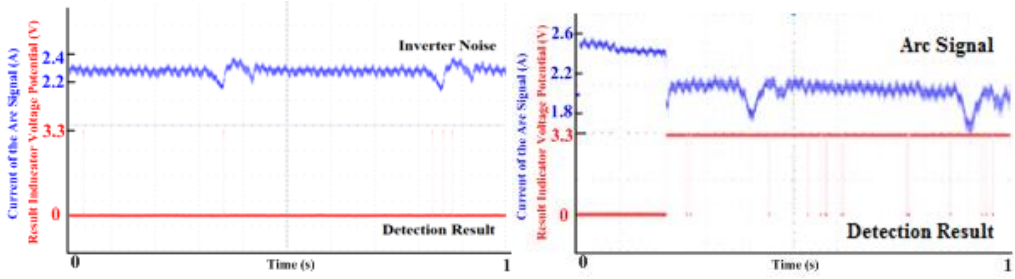


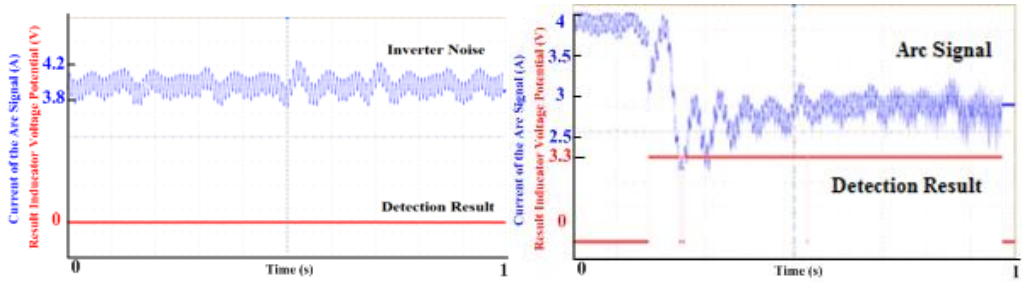
Figure 41: An example of real PV arc data (inverter noise included) and inverter noise.



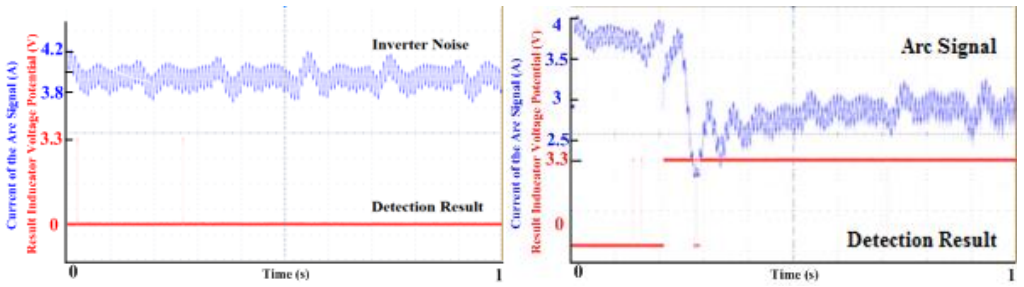
(a) Data1



(b) Data 2



(c) Data 3



(d) Data 4

Figure 42: DWT algorithm detection result for PV inverter noise.

3.9 Summary of the Chapter

From the analysis shown so far, several conclusions can be drawn:

- It is difficult to find any significantly detectable arc fault features by comparing the FFT results of the non-arcing part and the arcing part of the signals, especially when the sampling rate is decreased.
- However, from the wavelet analysis plots, arc features can be easily distinguished from the non-arcing signal and selection of detection criteria potentially suitable for an embedded microcontroller for real-time arc fault detection.
- Further, since the wavelet transform preserves the time-domain localization information, the precise time of the arc is available for cross-correlation with other events in the system and environment which may improve the confidence of arc fault detection rather than some other benign electrical events.
- By comparing the analysis result of 3 different Daubechies wavelets, we conclude that decomposition results the decomposition results from db9 and db19 are significantly better than those from db3, which is expected according to the frequency responses of the wavelet filters being used. But db19 doesn't provide much improvement over db9. This suggests that optimal selection is possible from the perspective of frequency response analysis of the filters. However, since

- The implementation process of DWT is ideal for real-time sample-by-sample implementation since the computation load is similar to a simple FIR filter. However, FFT is not suitable for sample-by-sample analysis in real-time.
- The sampling frequency has a significant impact on both Fourier and wavelet detection approaches.
    - With the signal sampled at 100 kHz, it is almost impossible for the Fourier transform to capture any arc fault features.
    - While the sustained presence of the arc is not as obvious as when the sampling frequency is 1 MHz for wavelet decomposition, we should still be able to draw enough information to detect the arc fault when the sampling rate is at 100 kHz.
    - Thus, detection approaches based on wavelet can use a lower sampling rate than Fourier transform to accomplish accurate arc fault detection if indeed Fourier methods can accomplish it at all.
- Real-time wavelet transform can be successfully implemented on a popular low-cost MCU on the market. This demonstrates that integrating wavelet-based signal analysis algorithm would not add much extra cost to the BOS hardware cost.

4. ARC FAULT DETECTION USING SUPPORT VECTOR MACHINES*

4.1 A Few Fundamental Concepts of Machine Learning

4.1.1 Supervised Learning and Unsupervised Learning

The widely quoted, formal, modern definition of machine learning is provided by Tom Michell as: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E" [72]. By way of example, in chess games, E stands for the experience of playing many games of chess; T represents the task of playing chess games; and the probability that the program will win the next game serves as P.

In general, any machine learning problem can be assigned to one of two broad classifications: 1) supervised learning, or 2) unsupervised learning [1, 73, 74].

In supervised learning, an input data set is given, and the correct output corresponding to the input is already known. This gives the idea that there is a relationship between the input and the output. Supervised learning problems are further categorized into "regression" and "classification" problems. In a regression problem, we try to predict results within a continuous output, meaning that we try to map the input variables to some continuous function. While in a classification problem, we try to predict results in a

* Reprinted with permission from "Arc Fault and Flash Detection in Photovoltaic Systems Using Wavelet Transform and Support Vector Machines," by Z. Wang and R.S. Balog, 2016, IEEE 43rd Photovoltaic Specialists Conference (PVSC), pp. 3275-3280, © 2016 IEEE.

discrete output instead. In other words, we try to map the input variables into discrete categories [1, 73].

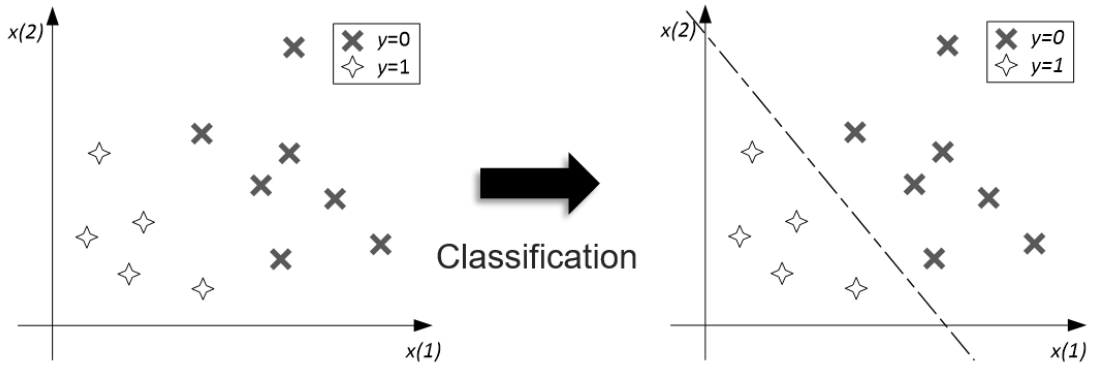


Figure 43: Typical supervised learning example.

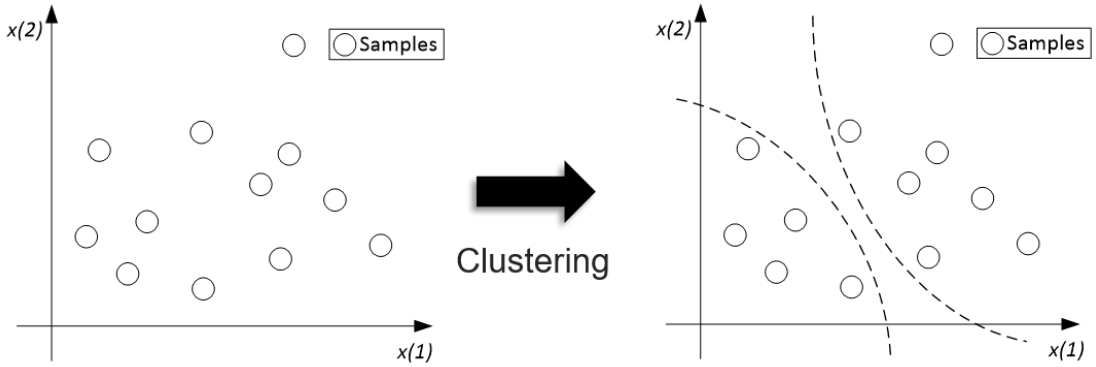


Figure 44: Typical unsupervised learning example.

Unsupervised learning, on the other hand, allows us to approach problems with little or no idea what the results should look like. We can derive structure from data where we do not necessarily know the effect of the variables. The structure can be derived by clustering the data based on the relationships among the variables in the data. With unsupervised learning, there is no feedback based on the prediction results [1, 73].

In this work, since all the raw data in the database, the correct label (category of non-arc event or arc event) of each data file is already provided, that makes this research problem a supervised learning problem.

A typical supervised learning procedure is shown in Figure 45. The training set is first fed into the learning algorithm to get an output "hypothesis". The new input data is then applied to the hypothesis to obtain an estimated output [73].

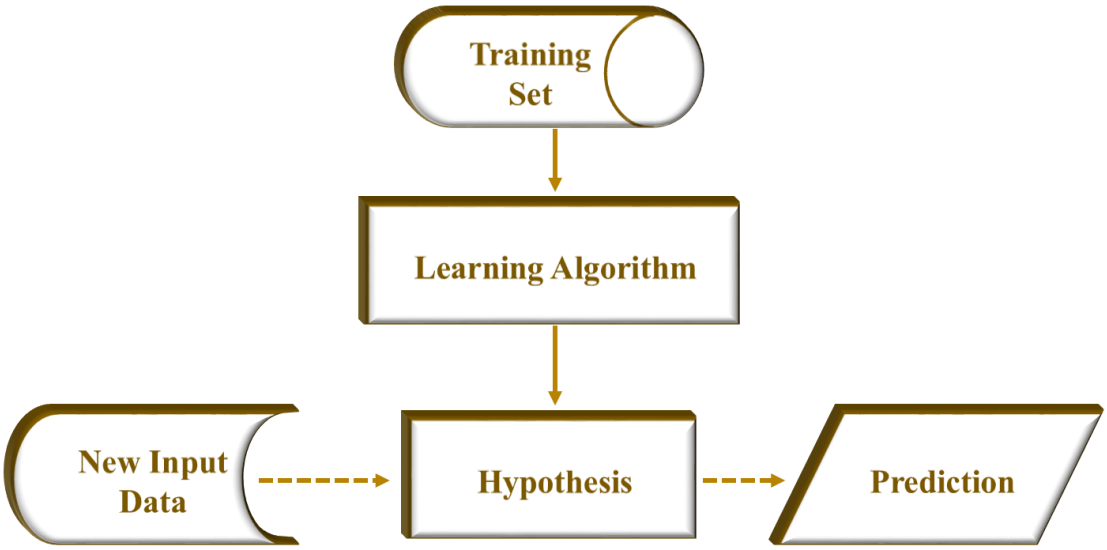


Figure 45: Illustration of supervised learning procedure.

4.1.2 Underfitting and Overfitting

The performance of a classifier depends on the interrelationship between sample sizes, number of features, and classifier complexity. A naive table-lookup technique (partitioning the feature space into cells and associating a class label with each cell) requires the number of training data points to be an exponential function of the feature

dimension. This phenomenon is called as “curse of dimensionality”, which leads to the “peaking phenomenon” in classifier design.

It is well-known that the probability of misclassification of a decision rule does not increase as the number of features increases, as long as the number of training samples is arbitrarily large and representative of the underlying class-conditional densities. However, it has often been observed in practice that the added features may degrade the performance of a classifier if the number of training samples that are used to design the classifier is small relative to the number of features. This paradoxical behavior is referred to as the peaking phenomenon. A simple explanation for this phenomenon is as follows: The most commonly used parametric classifiers estimate the unknown parameters and plug them in for the true parameters in the class-conditional densities. For a fixed sample size, as the number of features is increased (with a corresponding increase in the number of unknown parameters), the reliability of the parameter estimates decreases. Consequently, the performance of the resulting plug-in classifiers, for a fixed sample size, may degrade with an increase in the number of features (dimensionality and sample size consideration in pattern recognition practice).

Once the classification model is selected to fit some set of data (training set), the error of the parameters as measured on the data (the training error) is likely to be lower than the actual generalization error.

The generalization error can be decomposed as follows:

$$Err(\hat{f}) = Bais[\hat{f}]^2 + Var[\hat{f}] + \sigma^2 = Bias^2 + Variance + IrreducibleError$$

where $Bias[\hat{f}(x)] = E[\hat{f}(x)] - f(x)$, is the error from erroneous assumptions in the learning algorithm. High bias can cause an algorithm to miss the relevant relations between features and target outputs (underfitting); $Var[\hat{f}(x)] = E[\hat{f}(x)^2] - E[\hat{f}(x)]^2$

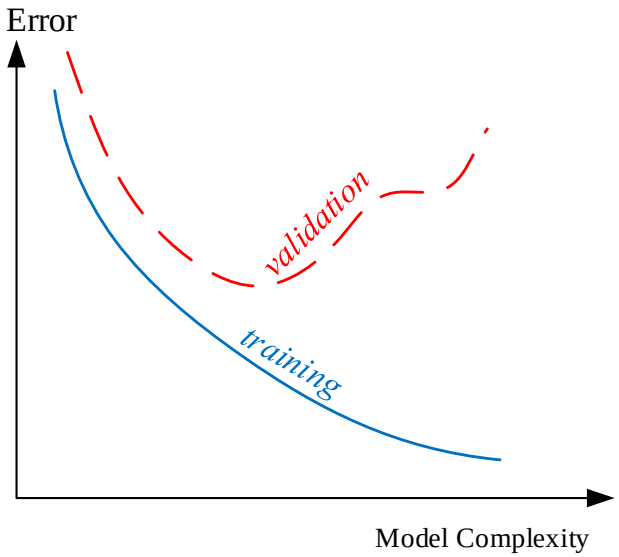


Figure 46: Model complexity vs. training error and validation error.

represents error from sensitivity to small fluctuations in the training set. High variance makes the classifier model the random noise in the training data, rather than the intended outputs (overfitting).

4.1.3 Error Metric and Analysis

It is important to have specific error metrics that evaluate the performance of the classifier. Conventionally, classification accuracy or traditional Receiver Operator Characteristic (ROC) curves are commonly used as the evaluation metric for binary classification problems. However, when dealing with situations where the number of

negative examples greatly exceeds the number of positive examples (the dataset is thus called highly skewed), the metrics pair Precision-Recall (PR) would give a more informative picture of the performance of an algorithm. A detailed comparison of the ROC and PR metrics is provided in [75].

The definition of precision and recall can be represented with the help of a confusion matrix as shown in the table below. *TP* stands for number of true positive samples, which are examples correctly labeled as positive; *FP* corresponds to false positive samples (negative examples incorrectly labeled as positive); *TN* (true negative) refers to the number negative samples which are correctly labeled as negative; and *FN* (false negative) equals to the number of actual positive samples incorrectly labeled as negative. Now we have:

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

Table IV: Confusion matrix

	Actual Positive	**Actual Negative**
Predicted Positive	*TP*	*FP*
Predicted Negative	*FN*	*TN*

In the application of arc fault detection, not only precision and recall suits the classification of the skewed dataset (meaning there are a lot more data samples labeled as

non-arcing events than the ones labeled as arcing signals), but the pair also represents the kind of accuracies we are most concerned about when it comes to fault identification.

UL1699B requires that the arc fault detectors should be designed to ensure basic arc-fault detection capabilities with resistance to unwanted tripping [43]. As shown in Figure 47, in this project, precision gives us the proportion of samples classified as arc faults are actual arcing events, while recall presents the proportion of actual arcing events

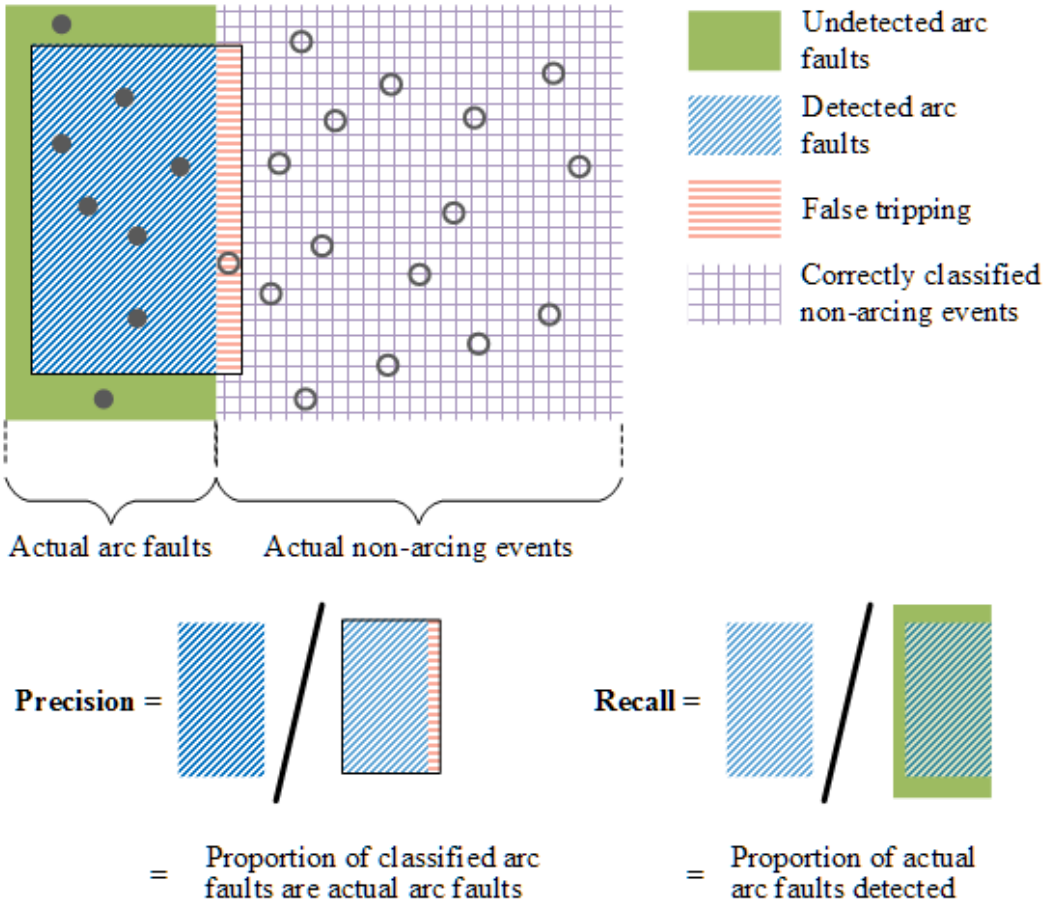


Figure 47: Illustration of precision and recall.

are correctly identified. In other words, higher precision value means lower unwanted tripping rate and higher recall value indicates lower no-detection rates.

4.1.4 Evaluating a Learning Algorithm

In practice, the error rate of a classification model must be evaluated from all the available samples which are partitioned into training and test sets. The classifier is first designed using training samples, and then it is evaluated based on its classification performance on the test samples.

Cross-validation is probably the most widely used method for estimating generalization error and evaluating a classification model.

Workflow of K-fold cross-validation is as follows (also illustrated in Figure 48):

1. Randomly split the entire training set into k disjoint subsets of m/k training examples each: $S_1, \ldots, S_k$.
2. For each model M_i, we evaluate it as follows:

   For $j = 1, \ldots, k_j$

   1) Train the model M_i on $S_1+\ldots+S_{j-1}+S_{j+1}+\ldots+S_k$ (i.e., train on all the data except S_j) to get some hypothesis h_{ij}
   2) Test the hypothesis h_{ij} on S_j, to get validation error ε_{ij}
   3) The estimated generalization error of model M_i is then calculated as the average of the ε_{ij} 's (averaged over j)

3. Pick the model M_i with the lowest estimated generalization error, and retrain that model on the entire training set S. The resulting hypothesis is then output as our final answer.

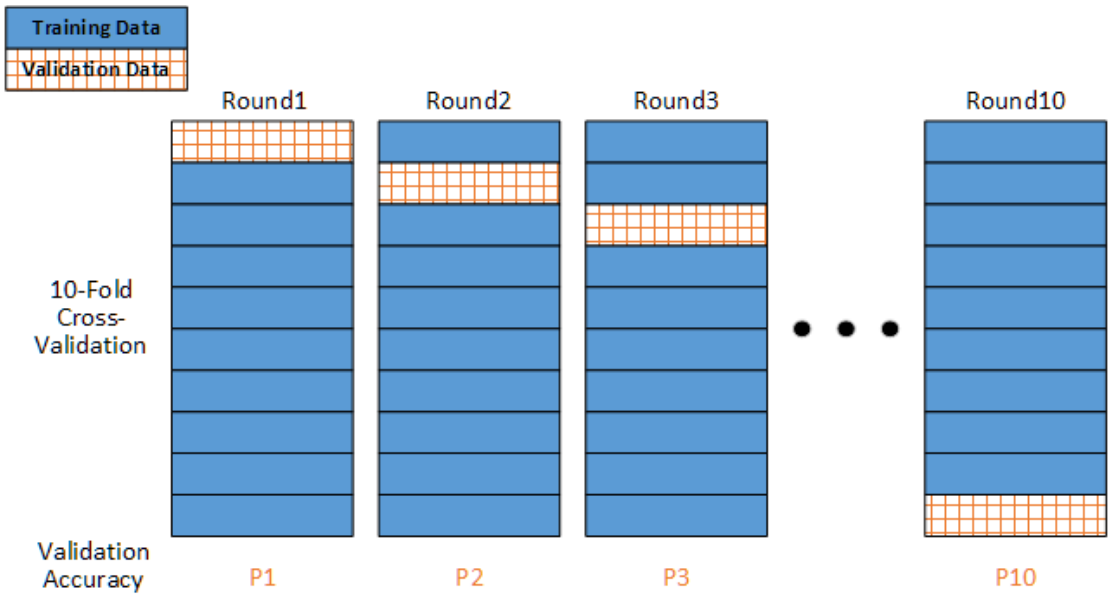


Figure 48: 10-fold cross-validation.

4.2 Feature Extraction Using DWT Results

Although wavelet analysis has been demonstrated to be effective with arc fault signal analysis, practical implementation needs a way to distinguish arc/non-arc events and provide annunciation that can be included in a DSP or microcontroller. A more robust way is to classify the feature vectors acquired by the wavelet transform.

The flowchart in Figure 49 [1] illustrates the workflow of the detection process which involves data acquisition, feature extraction (including data preprocessing and rescaling), classification model learning, and classification model evaluation. The output of this process will be the classification model which includes both the model structure and specified values of the coefficients.

The feature extraction process involves reducing the amount of resources required to describe a large set of data. Analysis of a large number of variables generally requires a large amount of memory and computation power [76], and it may also cause a classification algorithm to overfit the training samples and generalize poorly to new samples. An ideal feature extraction would make the job of the classifier trivial, and thus, makes the real-time classification much easier.

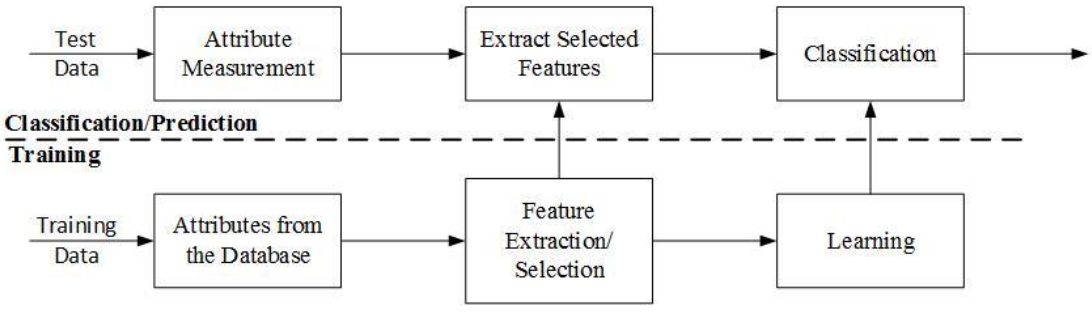


Figure 49: Training and test procedure for classification problems [1].

Typical features of arc fault signals include conductor temperature, voltage/current signals, dc voltage/current level, energy, frequency spectrum, etc. In this chapter, only current signals are used as the input attribute.

Two practical problems must be overcome in the implementation of wavelets for fault detection in power systems:

- Adopting the DWT coefficients directly for fault detection requires large memory space and computing time. Thus, a feature vector with reduced quantities and without losing the property of the original signal needs to be chosen.
- The wavelet function, as well as the decomposition level, must be properly chosen to meet the tradeoff between the accuracy of detecting the fault type and the computing efficiency.

The extracted feature of the proposed arc fault detection is based upon Parseval's theorem, states that if the used wavelets form an orthonormal basis and satisfy the admissibility condition, then the energy of the original signal is equal to the energy in each of the expansion coefficients, that is [51, 76]:

$$\int |f(t)|^2 = \sum_{k=-\infty}^{\infty} |c(k)|^2 + \sum_{j=0}^{J} \sum_{k=-\infty}^{\infty} |d_j(k)|^2$$

The original signal is decomposed into J levels by the wavelet transform. The energy is partitioned in time by k and in scale by j in the wavelet domain. C is the approximated coefficients from the jth level of the wavelet transform, d is for the detail coefficients from the jth level of wavelet transform. The energy variation of the fault signal at different resolution levels is adopted as feature vectors. By employing the Parseval's

theorem, the information is condensed, thus reducing the data size and yielding a manageable feature vector. Thus, the extracted feature of the proposed detection method is based on the average signal power in different resolutions. Since the frame size is customizable without trading off between the frequency and time resolution, transient behaviors in the signal can still be captured at the desired frequency resolution.

The literature suggests that the arc signature is most prevalent from 40 kHz to 100 kHz [66]. Thus, with the sampling frequency of 1 MHz, the 3^{rd} level (62.5 kHz ~ 125 kHz) and the 4^{th} level (31.25 kHz ~ 62.5 kHz) of the decomposition is selected as extracted attributes.

4.3 Learning Algorithm – Support Vector Machines

Although previous studies have demonstrated that wavelet transform has significantly superior processing results with arc fault signal analysis than Fourier-based methods, when it comes to practical implementation, classification is still needed to set up a boundary that enables the DSP or the microcontroller to determine if an arc fault has occurred.

A classifier can be designed using a number of possible approaches. In practice, the choice of a classifier is a difficult problem and it is often based on which classifier(s) happen to be available, or best know, to the user.

Three different fundamental approaches to designing a classifier are identified in:

- Classifier design based on the concept of similarity. This is the simplest and the most intuitive approach to classifier design: patterns that are similar to the particular training samples are assigned to the corresponding class of the training sample. Once a good metric has been established to define similarity, patterns can be classified by template matching or the minimum distance classifier using a few prototypes per class. The most typical classification algorithm is 1-nearest neighbor (1-NN) classifier. There is no training needed. It also provides robust performance when sufficient training patterns are provided. However, in real-time hardware implementation, this type of classifier would need a significant memory space and the real-time classification would be slow since patterns would be assigned to the class of the nearest neighbor. 1-NN rule can be

conveniently used as a benchmark for all the other classifiers since it appears to provide a reasonable classification performance in most applications. Further, assuming Euclidean distance is used to find the nearest neighbor since the 1-NN rule does not require any user-specified parameters, its classification results are implementation independent.

- Classifier design based on the probabilistic approach.
    - The optimal Bayes decision rule assigns a pattern to the class with the maximum posterior probability. This rule can be modified to take into account costs associated with different types of misclassifications. For known class conditional densities, the Bayes decision rule gives the optimum classifier. However, it is not realistic to have the prior probabilities and the class-conditional densities at hand since the photovoltaic systems in the field operate differently due to the variation of insolation, the ambient and the panel temperature, etc. Also, components supplied by different companies may respond to the same environment change in diverse ways, and thus, give different performance.
    - Logistic regression, which is based on the maximum likelihood approach, treats the output of a sigmoid function at the probability that the label is equal to 1 for a new input. In logistic regression, every training sample has a certain influence on the resulting classifier. However, as shown in the later sections of this chapter,

the training samples distribute in a very broad range of the sample space. A huge of amount of them sit very far away from the other the samples of the other class. Intuitively, we do not wish these "obvious" samples to be as influential as the "not-as-obvious" samples.

    - Two well-known nonparametric decision rules, the k-nearest neighbor (k-NN) rule, and the Parzen classifier also fall into this category. Same as 1-NN, both these classifiers require the computation of the distances between a test sample and all the patterns in the training set, which not only need memory space to save all the training, but massive computation power is demanded every new test sample. The large memory space and the excessive computation are not desirable for any real-time hardware implementation that aims to keep the setup under a limited budget.

- Classifier design based on constructing geometric decision boundaries. While this approach depends on the chosen metric, sometimes classifiers of this type may approximate the Bayes classifier asymptotically. The driving force of the training procedure is the minimization of a criterion such as the classification error or the mean squared error (MSE) between the classifier output and the label of the original pattern. The famous perceptron, feed-forward neural networks (multilayer perceptrons), and support vector machines all belong to this genre.

A few popular supervised learning algorithms are listed in Table V [1].

Support vector classifiers were first introduced by Vapnik [74]. Primarily, it is a two-class classifier. Nowadays, support vector machines (SVMs) are believed to be one of the best "off-the-shelf" supervised learning algorithms. The main concept behind SVM is to find a hyperplane as shown in Figure 50 with a maximum margin between the two adjacent classes which helps bound the generalization error of the classification model. Figure 50 illustrates a two class problem where a linear separation is achieved using a straight line. The margin refers to the blank space around the decision boundary defined by the geometric distance to the nearest training patterns. These patterns are called the support vectors, which eventually define the classification function. The number of support vectors is minimized by maximizing the margin. In cases where data points are clustered so that linear separation is not possible, the data points can be mapped into feature space (higher dimensional space) where a linear separation is possible. This hyperplane which is linear in the mapped feature space will not be linear in its original input space [1].

The derivation of an SVM classifier is as follows:

Let n-dimensional inputs x_i (i = 1, 2, …, m, where m is the number of samples) belong to class-1 or class-2 and associated to labels y_i = 1 for class-1 and y_i = -1 for class-2, respectively. For linearly separable data, a hyperplane $f(x) = 0$ which separates the data can be determined

$$f(x) = \omega^T x + b = 0$$

Table V: Supervised learning algorithms [1]

Method	Property	Real-time Implementation
Nearest Mean Classifier	Assigns patterns to the nearest class mean	Almost no training needed; fast testing; scale (metric) dependent
K-Nearest Neighbor Rule	Assigns patterns to the majority class among K nearest neighbor using a performance optimized value K	No training needed; robust performance; slow testing; scale (metric) dependent
Bayes Plug-In	Assigns patterns to the class which has the maximum estimated posterior probability	Yields simple classifiers (linear or quadratic) for Gaussian distributions; sensitive to density
Logistic Regression	Maximum likelihood rule for logistic (sigmoidal) posterior probability	Linear classifier; optimal for a family of different distributions (Gaussian); suitable for mixed data types
Binary Tree	Finds a set of thresholds for a pattern-dependent sequence of features	Overtraining sensitive; needs pruning; fast testing
Feed-Forward Neural Network	Iterative MSE optimization of two or more layers of perceptron (neurons, iterative optimization of linear classifiers) using sigmoid transfer functions	Sensitive to training parameters; slow training; nonlinear classification function; overtraining sensitive; needs regularization
Support Vector Machines	Maximizes the margin between the classes by selecting a minimum number of support vectors	Scale (metric) dependent; slow training; nonlinear; overtraining insensitive; good generalization performance

where $\boldsymbol{\omega}$ is an n-dimensional vector and b is the intercept term. The two vectors determine the position of the separating hyperplane. This separating hyperplane satisfies the constraints $f(x_i) \geq 0$ if $y_i = 1$and $f(x_i) \leq -1$ if $y_i = -1$ and this results in the functional margin:

$$y_i f(x_i) = y_i(\omega^T x_i + b) \geq 1, \text{ for } i = 1, 2, \ldots, m$$

The separating hyperplane that creates the maximum distance between the plane and the nearest data is called the optimal separating hyperplane as shown in Figure 50. The geometric margin is found to be $1/||\omega||^2$. Considering noise with the slack variable ξ_i and error penalty C_i, the optimal hyperplane can be found by solving the following convex quadratic optimization problem:

$$\min_{\omega,b} \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{m}\xi_i$$

$$\text{s.t. } y_i(\omega^T x_i + b) \geq 1 - \xi_i,\ i = 1, \ldots, m$$

$$\xi_i \geq 0,\ i = 1, \ldots, m$$

Examples are now permitted to have functional margin in (3) less than 1, and if an example has functional margin $1 - \xi_i$ (with $\xi > 0$), the extra cost of the objective function would be $C\xi_i$. The parameter C controls the relative weighting between the twin goals of making the $||\omega||^2$ small and of ensuring that most examples have a functional margin at least 1.

Now, the Lagrangian can now be formed:

$$\max_{\alpha} W(\alpha) = \sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum y_i y_j \alpha_i \alpha_j \langle x_i, x_j \rangle$$

$$\text{s. t. } 0 \leq \alpha_i \leq C,\ i = 1, \ldots, m$$

$$\sum_{i=1}^{m}\alpha_i y_i = 0$$

α_i s are found by solving the Lagrangian duality problem. Now the decision boundary for a two-class problem derived from the support vector machines can be written

as follows using a kernel function $K(\vec{x}^{(i)}, \vec{x})$ of a new pattern $\vec{x}$ (to be classified) and a training pattern $\vec{x}^{(i)}$:

$$D(x) = \sum_{i=1}^{m} \alpha_i y^{(i)} K(\vec{x}^{(i)}, \vec{x}) + b$$

where α_i s are all zero except for the support vectors [1, 73, 74].

Comparing with the also widely adopted feed-forward neural networks, SVMs do not suffer from problems like local minima and overtraining. On the other hand, since SVMs with simpler kernels are proposed in this project, the developed SVM classifier can be parametric and easy to implement (in contrast to the non-parametric SVM classifiers constructed from more advanced kernel functions which consist a set of support vectors).

Detailed theoretical comparison between different supervised learning algorithms and optimal learning algorithm selection are beyond the scope of this dissertation.

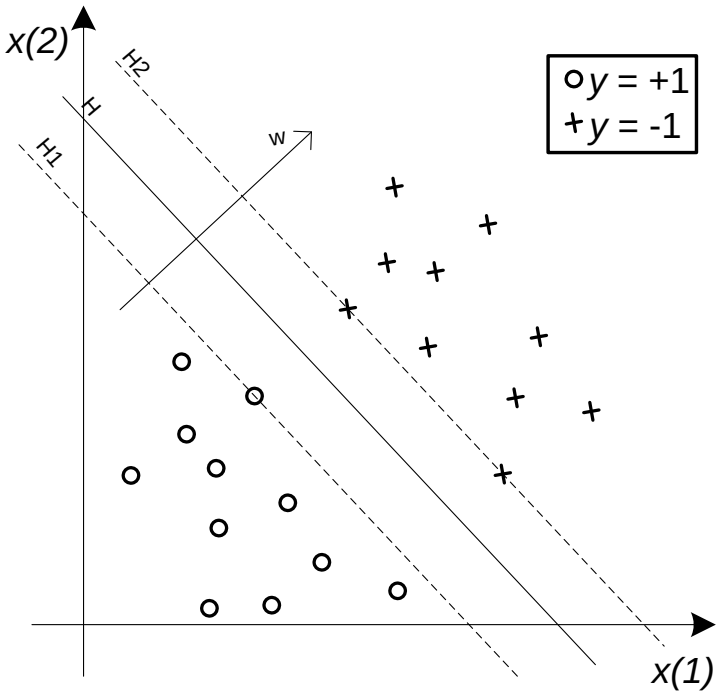


Figure 50: An SVM trained with samples from two classes. Samples on the margin are called the support vectors.

4.4 Hardware Implementation Strategy

Following the procedures mentioned above, the developed SVM will be implemented by dedicated hardware. A flowchart of the proposed implementation strategy is shown in Figure 51. On the left of the figure, it shows that the high-level procedure is divided into three steps: signal acquisition, feature extraction, and prediction & classification. In the middle, the figure reveals which hardware component each procedure is carried out. The specific implementations of each process are shown on the right.

An appropriate kernel function *K* need to be selected before applying the SVM algorithm. In its simplest form, we have $K(\vec{x}_i, \vec{x}) = <\vec{x}_i, \vec{x}>$, resulting in a linear classifier. For a 2-dimensional feature set:

$$D(x) = \sum_{i=1}^{m} \alpha_i y^{(i)} K(\vec{x}_i, \vec{x}) + b = \sum_{i=1}^{m} \alpha_i y^{(i)} (x_1^{(i)} x_1 + x_2^{(i)} x_2) + b = \omega_1 x_1 + \omega_2 x_2 + b$$

where $\omega_1 = \sum_{i=1}^{m} \alpha_i y^{(i)} x_1^{(i)}$, $\omega_2 = \sum_{i=1}^{m} \alpha_i y^{(i)} x_2^{(i)}$

As we can see here, the resulting decision boundary can be parametric without saving all the support vectors for kernel function calculation. For other kernels such as low-order polynomial kernels, the decision boundary function can also be straightforward with a few parameters. These parameters can be directly applied to the test data in real-time applications.

Despite high-quality models constructed by kernel SVMs, the use of kernel SVM in real-world application remains limited due to the high prediction cost. Linear SVM has

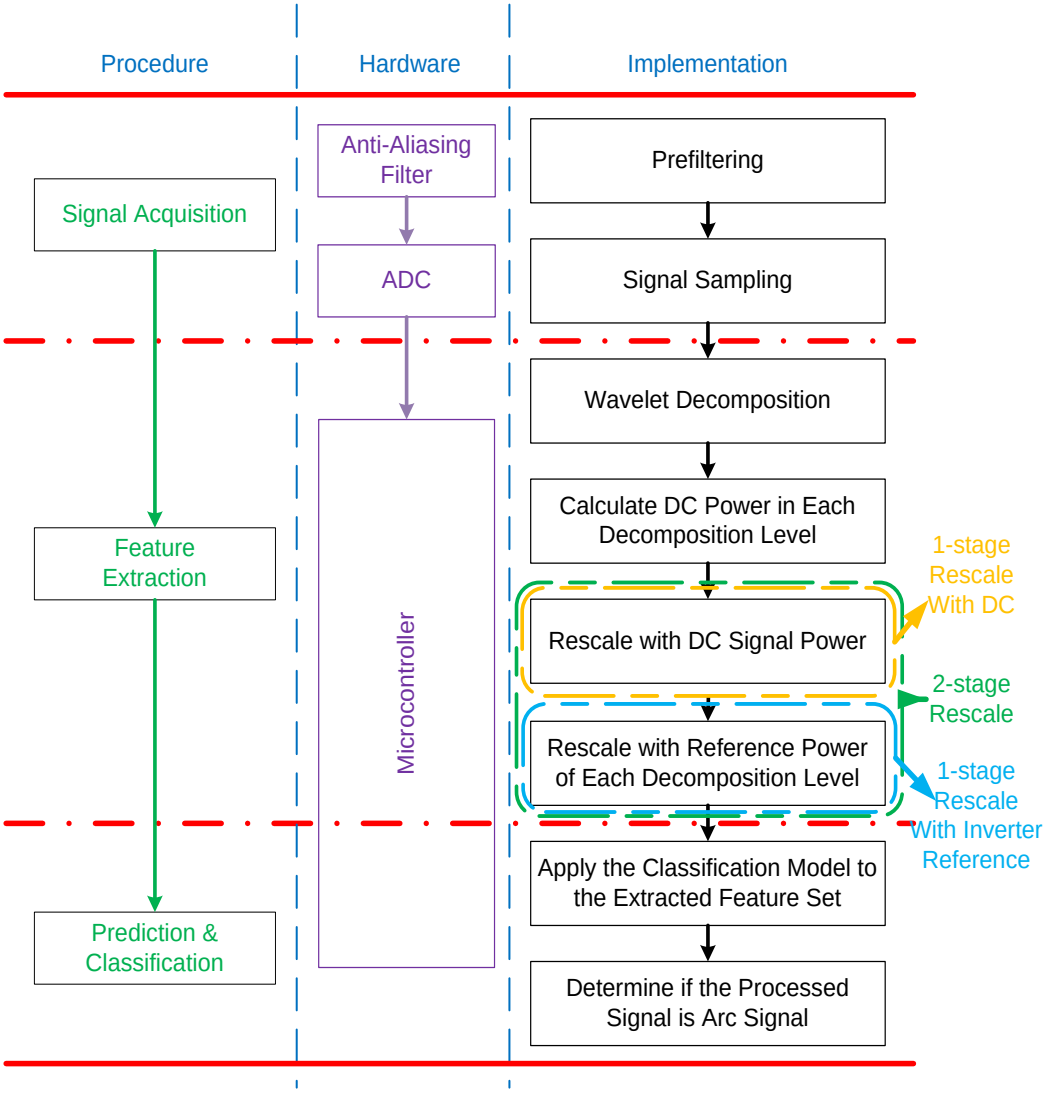


Figure 51: Work flow of the proposed hardware implementation strategy.

prediction complexity of *O(n)* with *n* as the number of input dimensions. Prediction complexity of kernel SVM depends on the choice of kernel and is typically proportional to the number of support vectors. Since the prediction is going to be implemented in real-time at a very high frequency on an embedded system (a DSP or a microcontroller), which has very limited memory space and computation speed compared with a powerful personal computer, linear SVM is first considered in this application.

The parameters of the developed classification boundary will be directly applied to the extracted feature vector without any further learning process in the embedded systems.

Now, let's take another look at the hardware implementation strategy of the DWT and classification. As illustrated in Figure 52, since the 3rd decomposition level only gives one output every eight samples and the discriminant function normally takes even less computation than 1 level of decomposition, the real-time classification using the DWT result does not require extra computation power from the DSP.

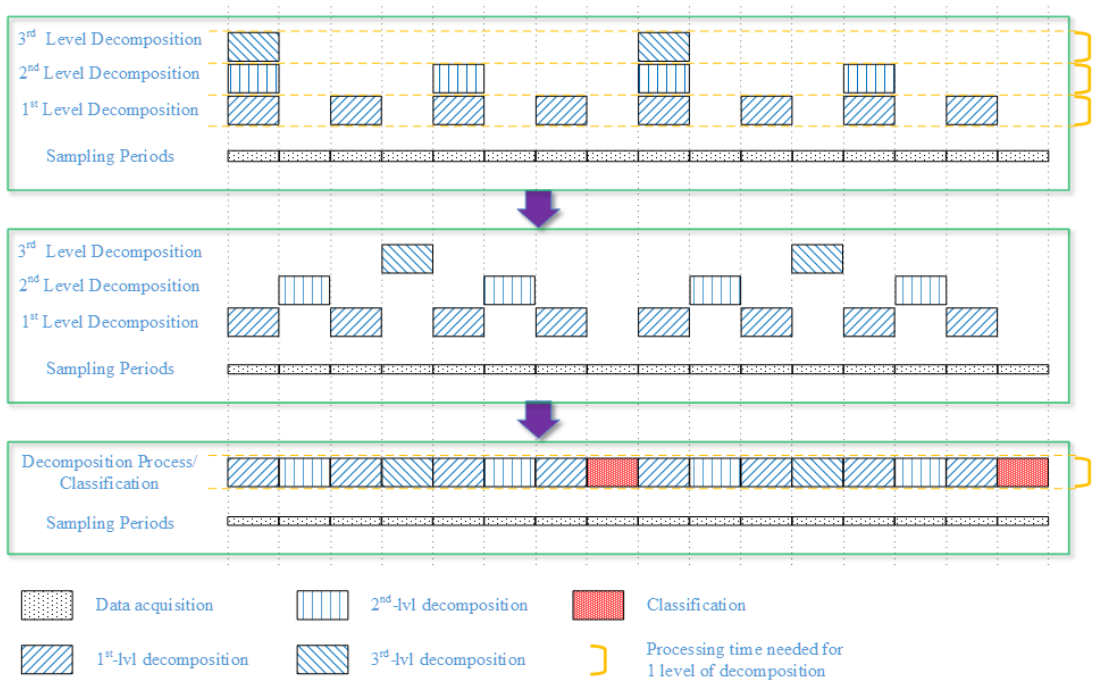


Figure 52: Hardware implementation of a 3-level DWT and classification.

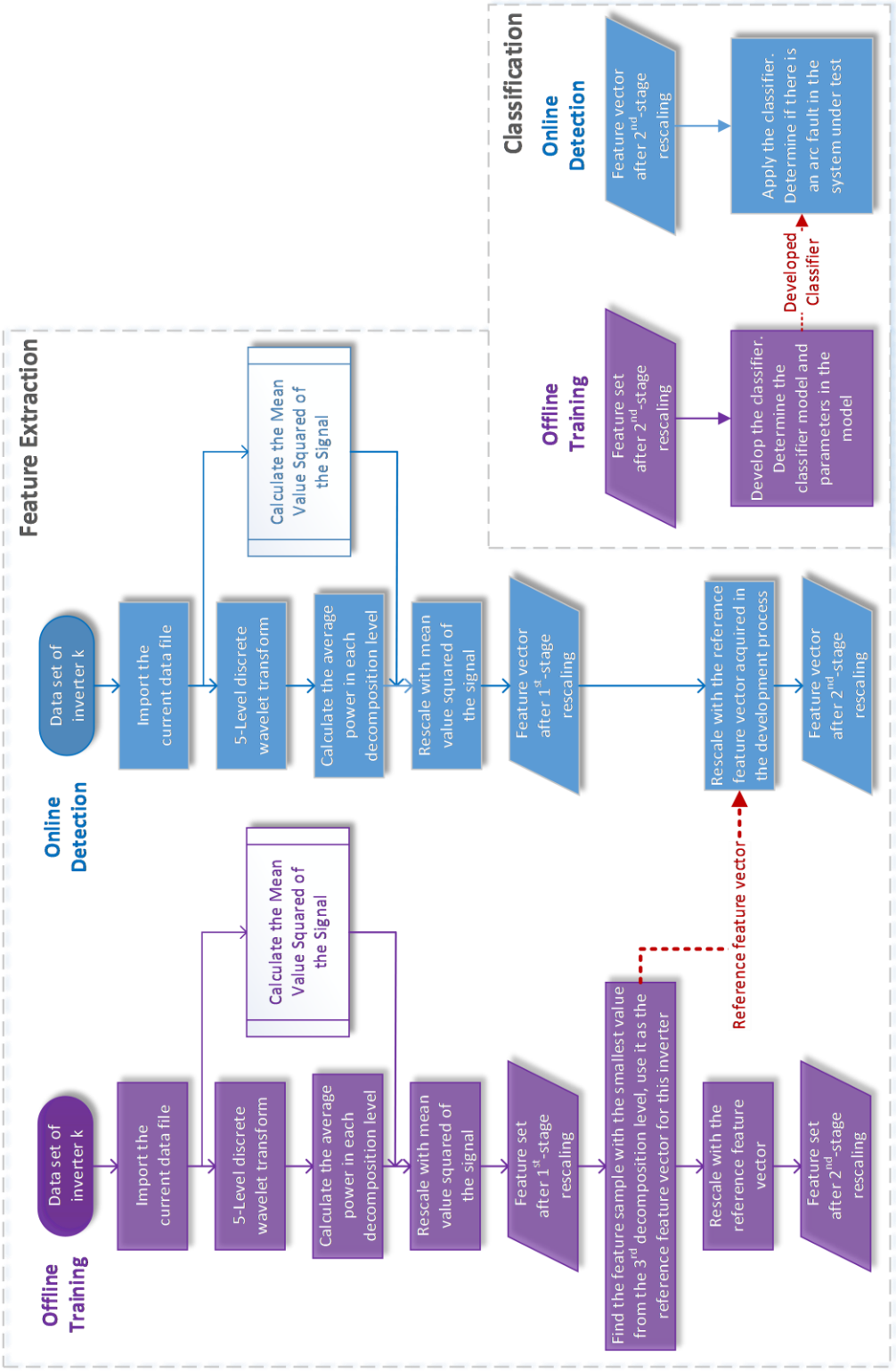


Figure 53: Offline model development vs. online detection.

4.5 Rescaling, and Cross-Validation with Linear Kernel

Since the range of values of raw data varies widely, in some machine learning algorithms, the learning process will not work properly without rescaling. Also, some optimization algorithms, such as gradient descent, converge much faster with properly rescaled feature set than the extracted raw feature.

The workflow of the entire feature extraction process is shown in Figure 53 and Figure 54. Figure 55 demonstrates the distribution of the extracted feature without normalization over the extracted attributes (average power of the 3rd and the 4th detail wavelet coefficients), with 1-stage rescaling, and with 2-stage rescaling.

Four different rescaling strategies are adopted in this paper:

1) No rescaling ➔ Figure 55(a);
2) Rescale by the power of the DC component in the current for each frame of signal ➔ Figure 55(b);
3) Use an inverter noise signal as the reference for the corresponding system setup. Rescale element-wise by the extracted attributes in each decomposition level of the reference signal ➔ Figure 55(c);
4) Rescale by both 1) and 2) ➔ Figure 55(d).

As shown in Figure 55(d), although some of the extracted feature samples of the positive class (inverter noise + arc fault) and the negative class (inverter noise) are clustered in similar areas, we should still be able to find a boundary between the two clusters to separate the two classes. However, samples of the two classes are clustered together for no rescaling and the two 1-stage rescaling feature sets. There is no obvious

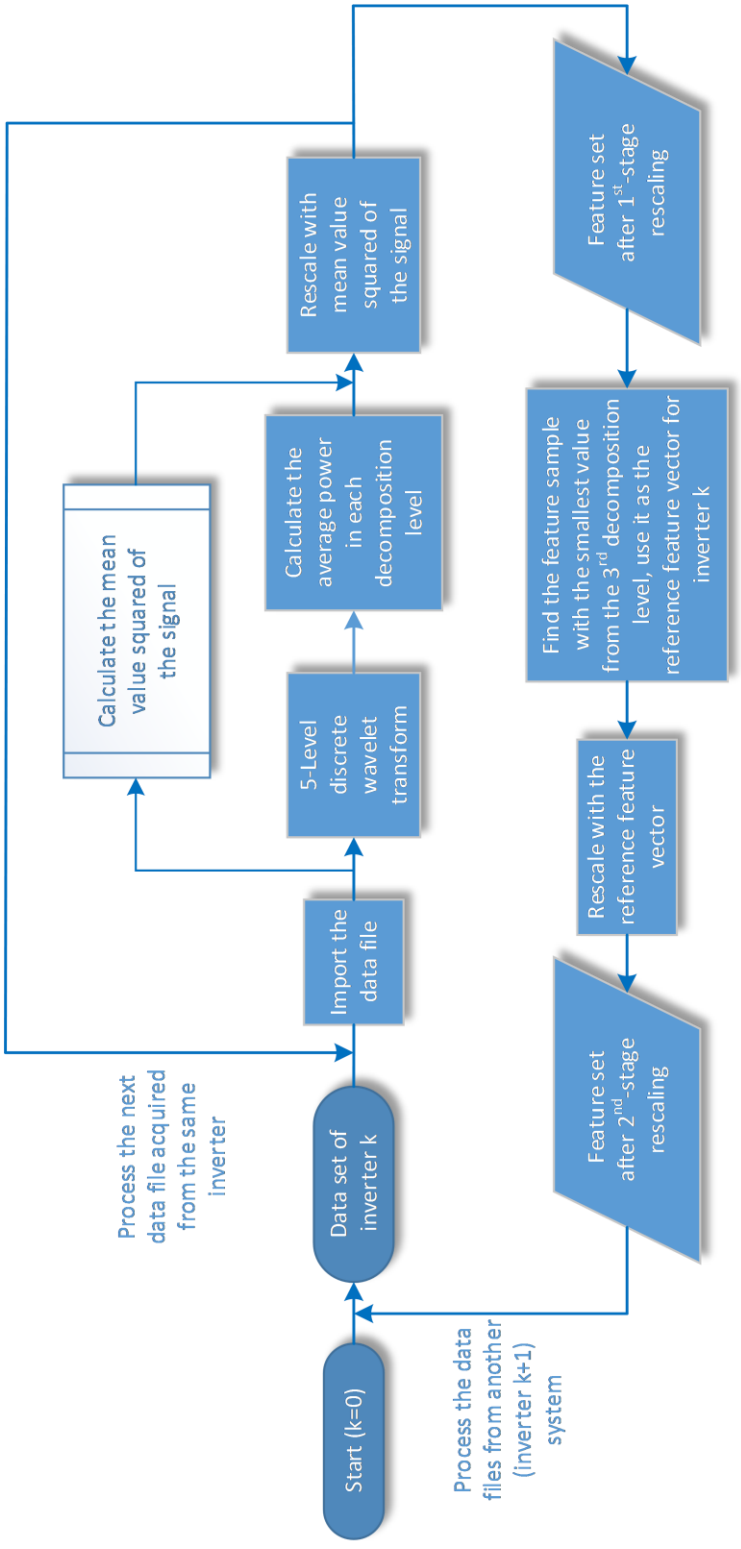


Figure 54: Work flow of the proposed hardware implementation strategy.

separation between them. Thus, the extracted features with 2-stage rescaling should be ideal for the application of a linear SVM classifier.

Four linear SVMs are derived using the training samples from the four rescaling strategies. The 10-fold validation accuracy is shown in Table VI. For the first two rescaling strategies, we get NaN for precision and 0 for recall. That is because the feature data set is significantly "skewed" (70 arc events and 475 non-arc events) and there is no clear boundary or grouping between the feature samples from the two distinct classes, in order to maximize the classification accuracy, the linear models predict that all new of the new samples as inverter noise (no sample is predicted as "positive") to maximize the accuracy of the prediction. Thus, while the general accuracy stays at about 0.87, the measurements of precision and recall truthfully reflect the derived model to be ineffective and unreliable. With precision and recall values taken into consideration, it is obvious that the combination of wavelet transform and 2-stage rescaling produces a very satisfying result. The hyperplane trained from the entire training feature set using linear SVM is plotted in Figure 56. The training accuracy is 99.63% with precision/recall rate at 0.9857/0.9857. It means the training data is linearly separable which also demonstrates the effectiveness of the feature extraction process.

Table VI: 10-fold cross validation result using SVM with linear kernel

Rescaling Strategy	Accuracy	Precision	Recall
0-Stage	0.8704	NaN	0.0000
1-Stage Mean Squared	0.8704	NaN	0.0000
1-Stage Inverter Noise	0.9101	0.8000	0.4000
2-Stage	0.9963	0.9857	0.9857

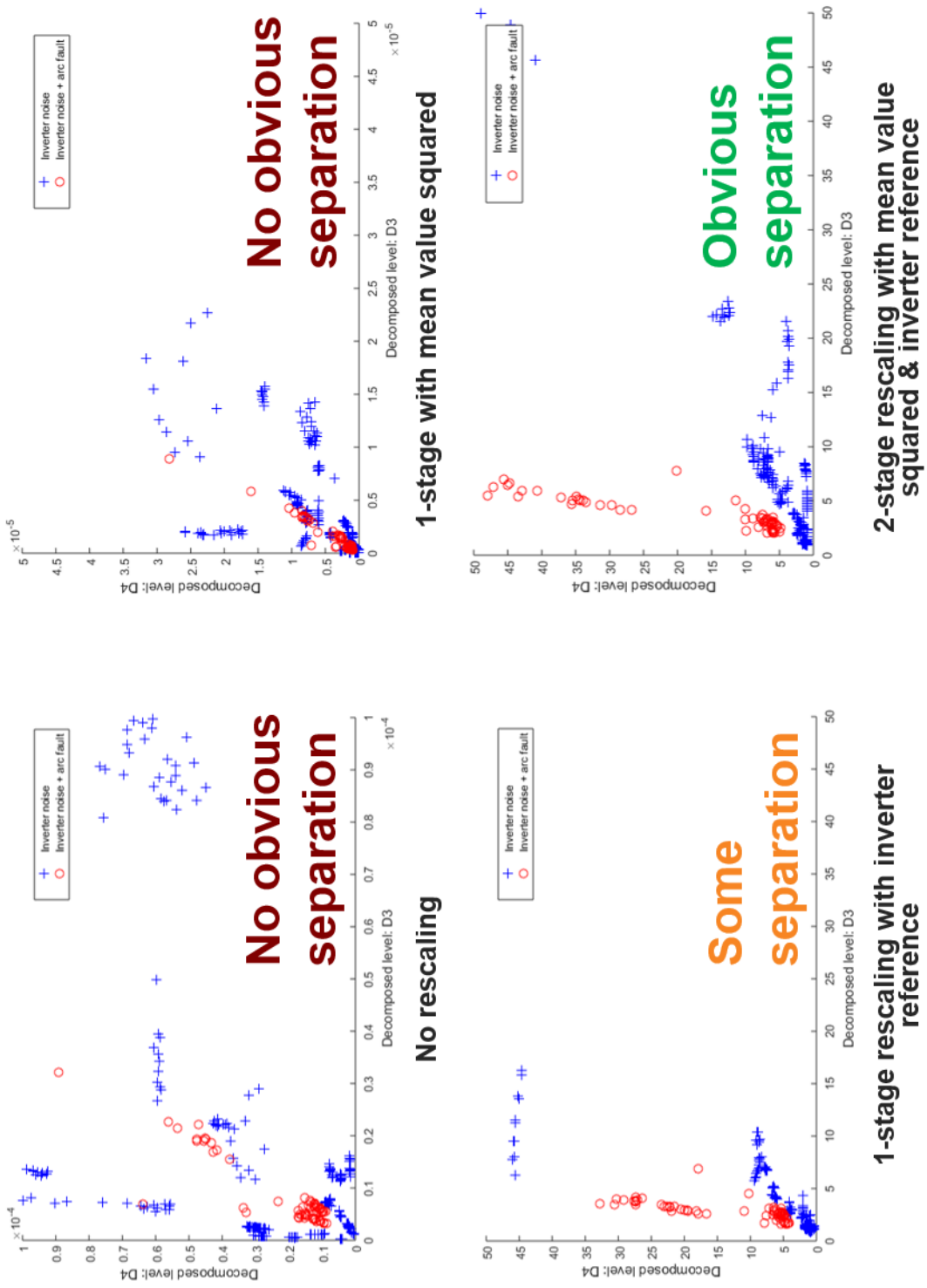


Figure 55: Simulated real-time implementation result.

An example of real-time implementation is simulated and demonstrated in Figure 57. The predicted “1” indicates that an arc fault is detected, and the predicted “0” suggests no arc fault has arisen in the system. The prediction result in Figure 57(a) shows there is no false tripping when the acquired signal only contains inverter noise. In the meantime, when an arc fault occurs with the presence of inverter noise at about 0.1s in Figure 57(b), the classifier detects the arc fault with high accuracy and fast response.

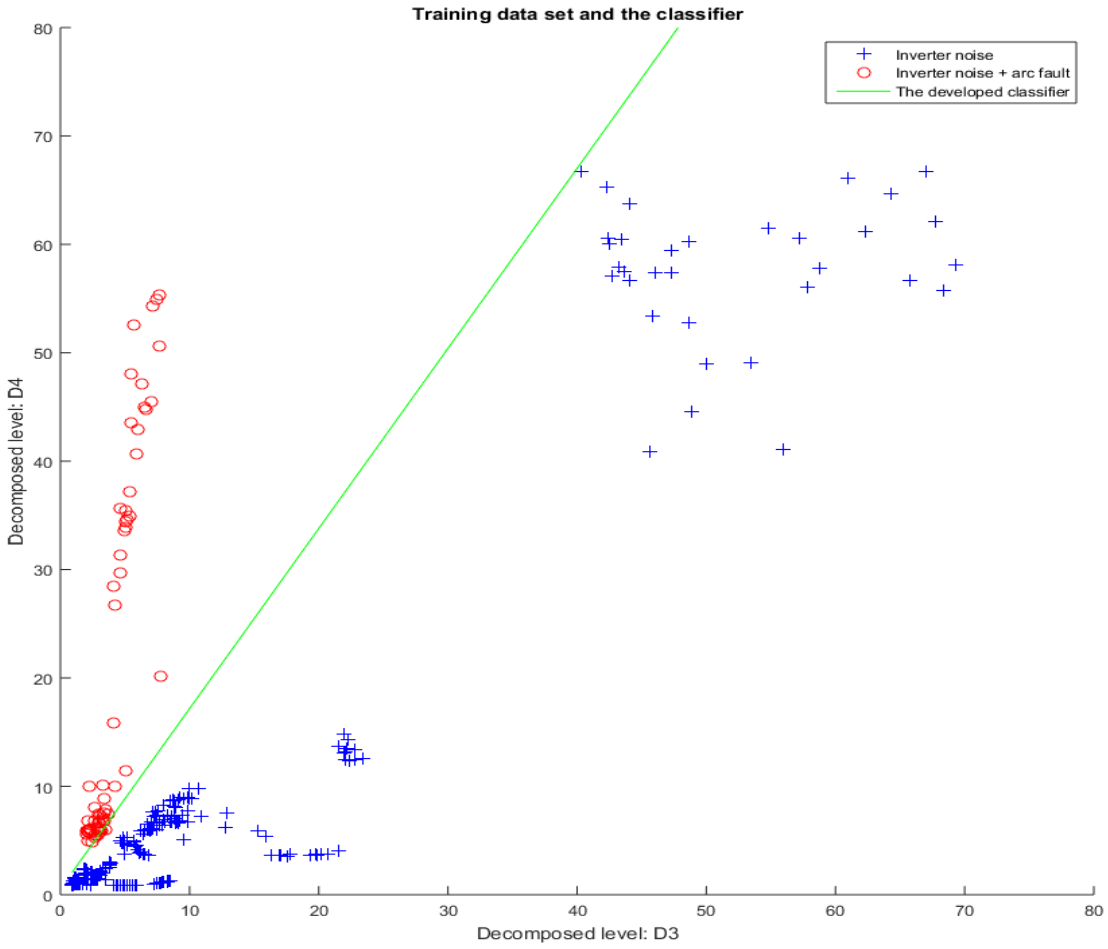
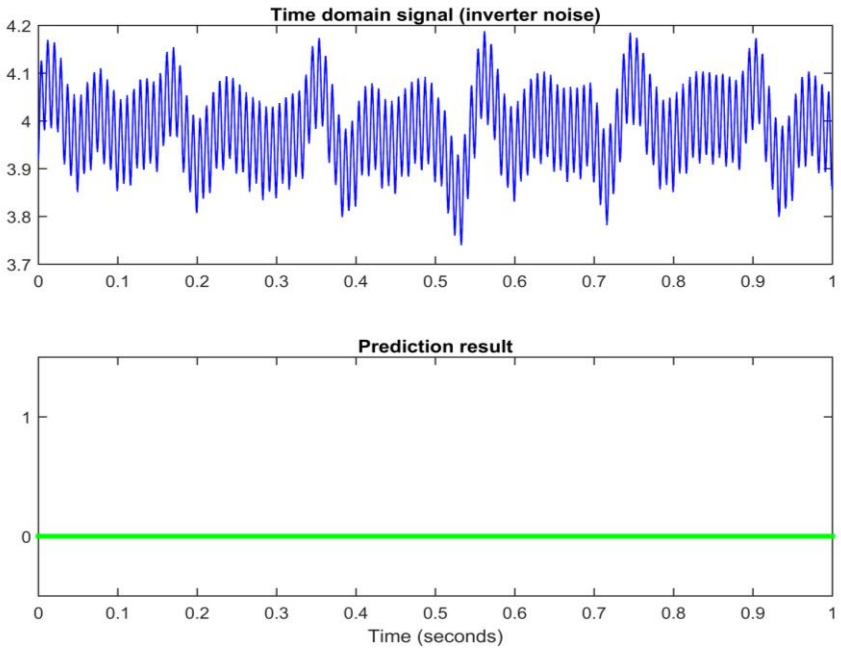
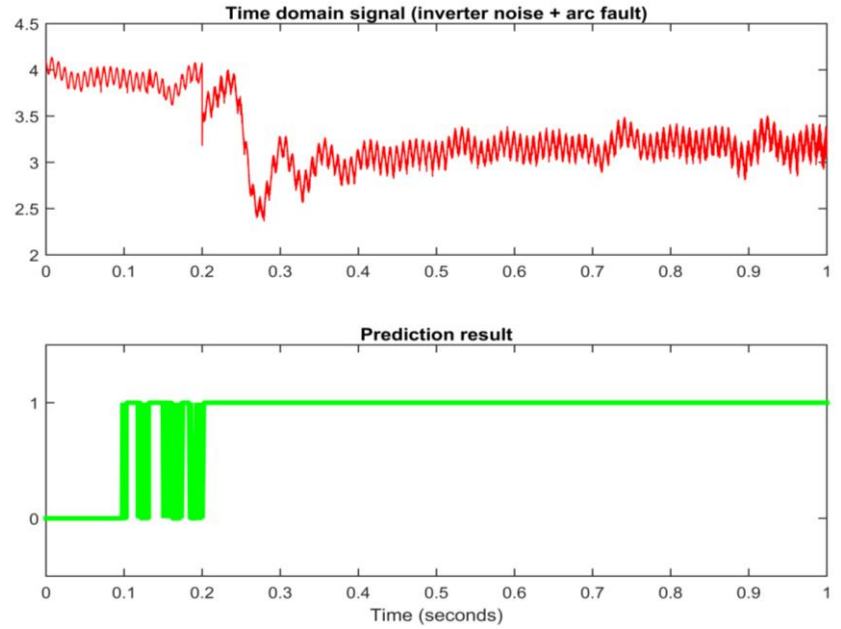


Figure 56: SVM trained with the entire training feature set.



(a) Prediction result of an inverter noise signal – no false tripping



(b) Prediction result of an inverter noise + arc fault signal
– fast and accurate detection

Figure 57: Simulated real-time implementation result.

4.6 Performance of SVM with Customized Feature Set

By observing the distribution of the feature samples on the *D*3-*D*4 plane (as shown in Figure 58 and Figure 59), it seems like a boundary function in the form of quadratic function, $D(x) = \omega_1 x_1^2 + \omega_2 x_1 + \omega_3 x_2 + b$, might fit the 2-stage rescaled data model better than the performance of the original feature set with the linear kernel.

Figure 59 shows the result of the boundary of a quadratic function derived by SVM. It can be easily seen that the quadratic function creates a discriminant boundary with larger margin and thus clearer separation between the two classes.

The triple criteria confirm the superiority of the quadratic function boundary over the linear boundary: accuracy = 100%, precision = 1, recall = 1. Further, there are only three support vectors for the quadratic boundary model, comparing with 14 support vectors for the linear model. It again verifies that the quadratic classifier fits the data model better and provides a more convincing classification.

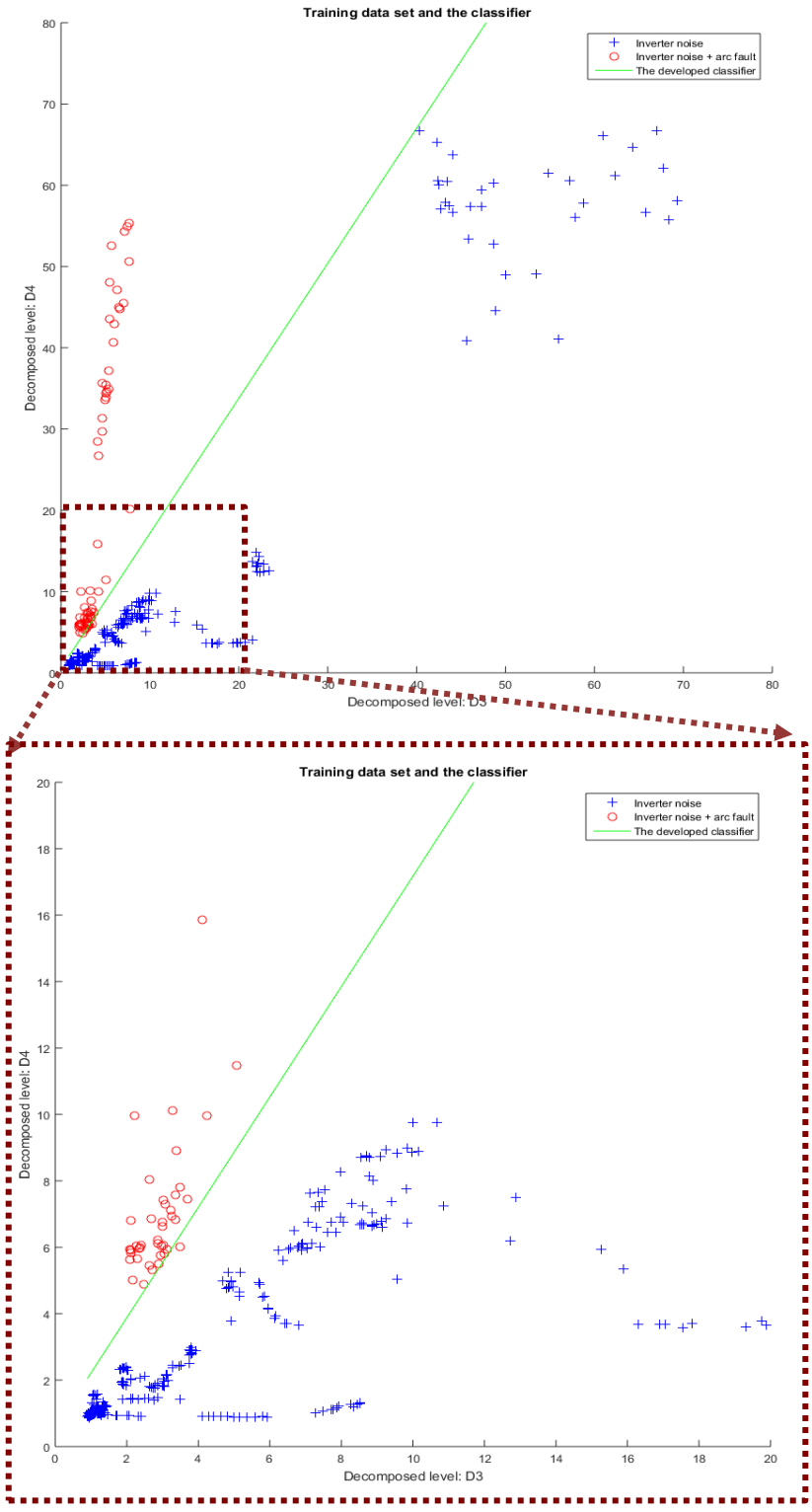


Figure 58: SVM trained with the entire training feature set using linear kernel.

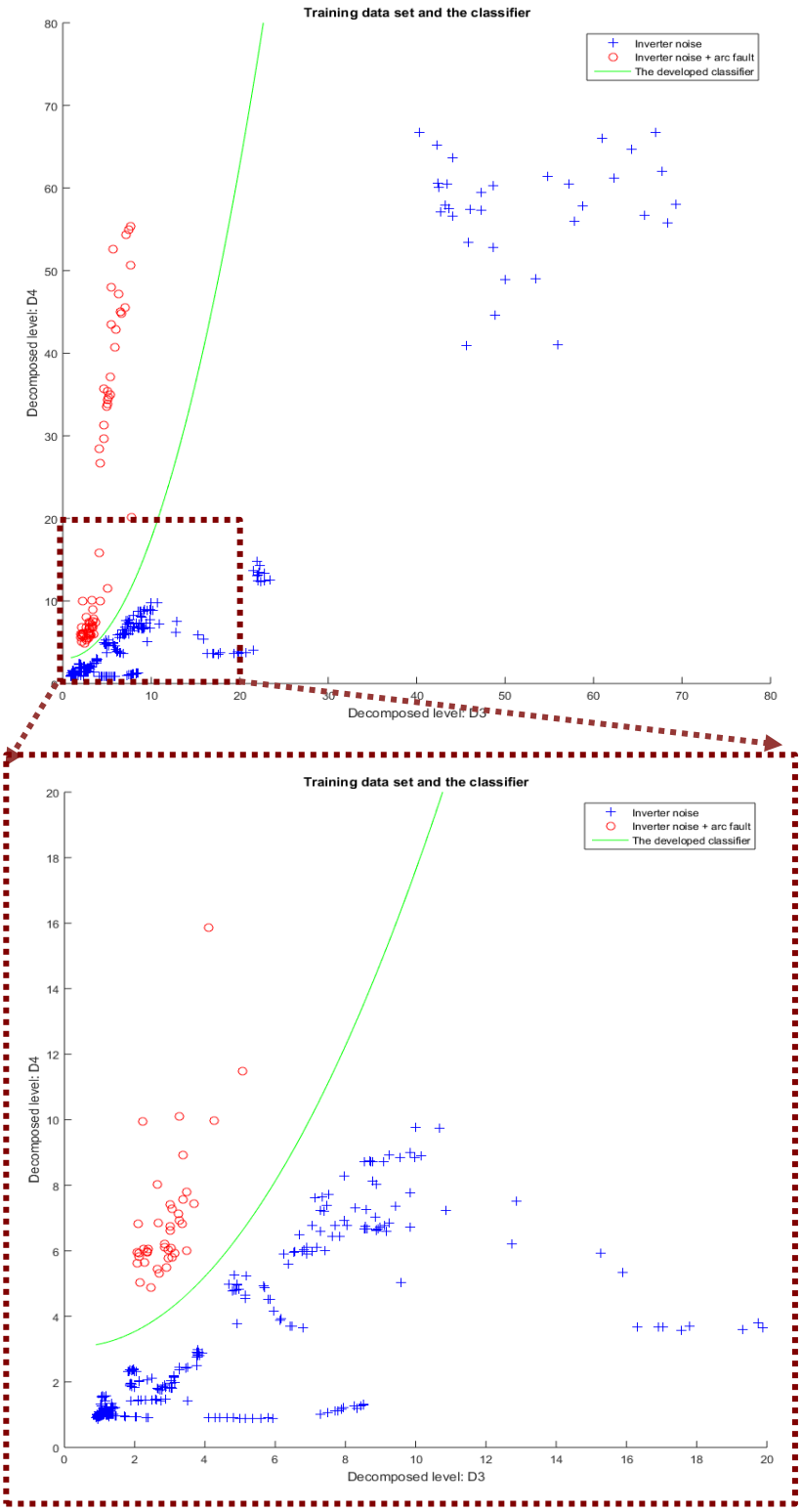


Figure 59: SVM trained with the entire semi-quadratic training feature set using linear kernel.

4.7 Exhaustive Search for the Optimal Feature Selection

Despite the curse of dimensionality mentioned in the first section of this chapter, a reduction in the number of features may lead to a loss in the discrimination power and thereby lower the accuracy of the resulting recognition system. Watanabe's famous ugly duckling theorem states that there is an unquantifiable number of shared properties between all objects, making any classification biased. It means that features have to be chosen carefully since it is possible to make two arbitrary patterns similar by encoding them with a sufficiently large number of redundant features.

Feature selection is about selecting (hopefully) the best subset of input feature set which leads to the smallest classification error or the optimum of the other criterion function of the classification (in this project, the paired value of precision and recall). The use of the classification criterion function makes feature selection procedures dependent on the sizes of the training, the method of feature extraction, and the specific classifier being used.

The most straightforward approach to the feature selection problem in this work would require 1) examining all $\binom{d}{m}$ possible subsets of size m from feature sets acquired from wavelet transform with different choice of mother wavelets, 2) deriving SVM model with different kinds of kernel selections, and 3) selecting the subset with the optimum value of the classification criterion. It has been shown that no nonexhaustive sequential feature selection procedure can be guaranteed to produce the optimal subset. The number of possible subsets grows combinatorially, making this exhaustive search impractical in

most problems. However, since our feature sets only have five dimensions from the 5-level wavelet decomposition, an exhaustive search is adopted in this project.

In our exhaustive search, feature sets with a list of mother wavelets (19 different wavelets) and different choice of dimensionalities (5-D, 4-D, 3-D, and 2-D), and SVM with a selection of kernel functions are tested. Linear, 2^{nd}-order polynomial, 3^{rd}-order polynomial, 4^{th}-order polynomial, and Gaussian kernel are used with feature sets of all listed dimensionalities. Quadratic feature sets are used with only 2-D feature sets with the linear kernel.

Catalogs containing the best result of their respective dimensionalities are listed below. The original name of the wavelets and the full result of the entire exhaustive search can be found in Appendix A.

4.7.1 Linear Kernel

D1, D2, D3, D4, and D5

	Accuracy	Precision	Recall
db1	0.9963	0.9857	0.9857
db2	0.9982	0.9859	1
db3	0.9963	0.9857	0.9857
db5	0.9982	0.9859	1
db9	0.9982	0.9859	1
db13	1	1	1
db19	0.9982	0.9859	1
coif1	0.9982	0.9859	1
coif2	0.9982	0.9859	1
coif3	0.9982	0.9859	1
coif4	0.9982	0.9859	1
coif5	0.9982	0.9859	1
sym2	0.9982	0.9859	1
sym3	0.9927	0.9714	0.9714
sym5	0.9982	0.9859	1
sym9	0.9982	0.9859	1
sym13	1	1	1
sym19	1	1	1
dmey	1	1	1

D1, D2, D4, and D5

	Accuracy	Precision	Recall
db1	0.9963	0.9857	0.9857
db2	1	1	1
db3	0.9982	1	0.9857
db5	1	1	1
db9	1	1	1
db13	1	1	1
db19	1	1	1
coif1	0.9982	1	0.9857
coif2	0.9982	1	0.9857
coif3	0.9982	0.9859	1
coif4	1	1	1
coif5	1	1	1
sym2	1	1	1
sym3	0.9982	1	0.9857
sym5	0.9982	0.9859	1
sym9	1	1	1
sym13	1	1	1
sym19	1	1	1
dmey	1	1	1

D1, D2, and D4

	Accuracy	Precision	Recall
db1	0.9963	0.9857	0.9857
db2	1	1	1
db3	1	1	1
db5	1	1	1
db9	1	1	1
db13	1	1	1
db19	1	1	1
coif1	1	1	1
coif2	1	1	1
coif3	1	1	1
coif4	1	1	1
coif5	1	1	1
sym2	1	1	1
sym3	1	1	1
sym5	1	1	1
sym9	0.9982	0.9859	1
sym13	1	1	1
sym19	1	1	1
dmey	0.9945	0.9718	0.9857

D2, and D4

	Accuracy	Precision	Recall
db1	0.9963	0.9857	0.9857
db2	1	1	1
db3	1	1	1
db5	0.9982	0.9859	1
db9	0.9982	0.9859	1
db13	0.9982	0.9859	1
db19	1	1	1
coif1	1	1	1
coif2	0.9982	0.9859	1
coif3	0.9982	0.9859	1
coif4	0.9982	0.9859	1
coif5	0.9982	0.9859	1
sym2	1	1	1
sym3	1	1	1
sym5	0.9982	0.9859	1
sym9	0.9982	0.9859	1
sym13	0.9982	0.9859	1
sym19	1	1	1
dmey	0.9927	0.9714	0.9714

Discussion from observing the classification result of SVM using linear kernel:

- Since the feature extraction process (which includes 5-level wavelet decomposition, the average power calculation, and the 2-stage rescaling) as made the work of classifier fairly “easy”, for any number of dimensionality more than 1-D. At least, one of the combinations gives satisfying accuracy, precision and recall numbers.
- Combinations containing *D4* seem to be consistently promising.
- 3-D combinations seem to perform the best throughout all selected wavelets.
- No obvious overfitting or underfitting.

4.7.2 Linear Kernel with Quadratic Feature Sets

D2, and D4

	Accuracy	Precision	Recall
db1	0.9982	1	0.9857
db2	1	1	1
db3	1	1	1
db5	1	1	1
db9	1	1	1
db13	1	1	1
db19	1	1	1
coif1	1	1	1
coif2	1	1	1
coif3	1	1	1
coif4	1	1	1
coif5	1	1	1
sym2	1	1	1
sym3	1	1	1
sym5	1	1	1
sym9	1	1	1
sym13	1	1	1
sym19	1	1	1
dmey	0.9982	0.9859	1

Discussion from observing the classification result of SVM using linear kernel with quadratic feature sets:

- It seems like the selected features, and the classifier fit the data model almost perfectly, which make the classification very easy and accurate.

4.7.3 2nd-Order Polynomial Kernel

D1, D2, D3, D4, and D5

	Accuracy	Precision	Recall
db1	0.2018	0.1174	0.8
db2	1	1	1
db3	0.7963	0.0816	0.0571
db5	0.8844	1	0.1
db9	0.8972	1	0.2
db13	0.9853	1	0.8857
db19	0.9853	1	0.8857
coif1	0.8606	0.4766	0.8714
coif2	0.8	0.1020	0.0714
coif3	0.8110	0.2203	0.1857
coif4	0.8844	1	0.1
coif5	0.9872	1	0.9
sym2	1	1	1
sym3	0.8037	0.1064	0.0714
sym5	0.8972	1	0.2
sym9	0.8844	1	0.1
sym13	0.9009	0.5755	0.8714
sym19	0.9835	0.9841	0.8857
dmey	0.8936	0.5556	0.8571

D1, D2, D3, and D4

	Accuracy	Precision	Recall
db1	0.9963	0.9857	0.9857
db2	0.9982	0.9859	1
db3	0.9982	0.9859	1
db5	0.9982	0.9859	1
db9	0.9982	0.9859	1
db13	0.9982	0.9859	1
db19	0.9982	0.9859	1
coif1	0.9982	0.9859	1
coif2	0.9982	0.9859	1
coif3	0.9982	0.9859	1
coif4	0.9982	0.9859	1
coif5	0.9982	0.9859	1
sym2	0.9982	0.9859	1
sym3	0.9982	0.9859	1
sym5	0.9982	0.9859	1
sym9	0.9982	0.9859	1
sym13	0.9982	0.9859	1
sym19	0.9982	0.9859	1
dmey	0.9982	0.9859	1

D1, D2, and D4

	Accuracy	Precision	Recall
db1	0.99633	0.985714	0.985714
db2	1	1	1
db3	1	1	1
db5	1	1	1
db9	1	1	1
db13	1	1	1
db19	1	1	1
coif1	1	1	1
coif2	1	1	1
coif3	1	1	1
coif4	1	1	1
coif5	1	1	1
sym2	1	1	1
sym3	1	1	1
sym5	1	1	1
sym9	1	1	1
sym13	1	1	1
sym19	0.998165	0.985915	1
dmey	0.99633	0.985714	0.985714

D2, and D4

	Accuracy	Precision	Recall
db1	0.9963	0.9857	0.9857
db2	1	1	1
db3	1	1	1
db5	0.9982	0.9859	1
db9	0.9982	0.9859	1
db13	0.9982	0.9859	1
db19	1	1	1
coif1	1	1	1
coif2	0.9982	0.9859	1
coif3	0.9982	0.9859	1
coif4	0.9982	0.9859	1
coif5	0.9982	0.9859	1
sym2	1	1	1
sym3	1	1	1
sym5	0.9982	0.9859	1
sym9	0.9982	0.9859	1
sym13	0.9982	0.9859	1
sym19	1	1	1
dmey	1	1	1

Discussion from observing the classification result of SVM using 2nd-order polynomial kernel:

- Combinations containing *D4* seem to be consistently promising.
- 3-D feature set seems to perform the best throughout all selected wavelets.
- Some of the 5-D feature sets with 2nd-order polynomial clearly overfit the data model.

4.7.4 3rd-Order Polynomial Kernel

D1, D2, D3, D4, and D5

	Accuracy	Precision	Recall
db1	0.9945	0.9855	0.9714
db2	0.7945	0.3765	0.9143
db3	0.5064	0.1759	0.7714
db5	0.9945	0.9718	0.9857
db9	0.9963	0.9857	0.9857
db13	0.9963	0.9857	0.9857
db19	0.9945	0.9718	0.9857
coif1	0.8257	0.4194	0.9286
coif2	0.9945	0.9855	0.9714
coif3	0.9927	0.9853	0.9571
coif4	0.9963	0.9857	0.9857
coif5	0.9963	0.9857	0.9857
sym2	0.7523	0.3367	0.9571
sym3	0.5321	0.1886	0.8
sym5	0.9927	0.9714	0.9714
sym9	0.9963	0.9857	0.9857
sym13	0.9963	0.9857	0.9857
sym19	0.9963	0.9857	0.9857
dmey	0.9908	0.9452	0.9857

D1, D2, D3, and D4

	Accuracy	Precision	Recall
db1	0.9927	0.9583	0.9857
db2	1	1	1
db3	0.9229	0.6346	0.9429
db5	0.9651	0.8312	0.9143
db9	0.9009	0.5769	0.8571
db13	0.8844	1	0.1
db19	0.9505	0.8209	0.7857
coif1	0.996	0.9857	0.9857
coif2	0.9725	0.8313	0.9857
coif3	0.9523	0.8235	0.8
coif4	0.8844	1	0.1
coif5	0.8844	1	0.1
sym2	0.9982	0.9859	1
sym3	0.9156	0.6132	0.9286
sym5	0.8807	0.5203	0.9143
sym9	0.9706	0.9821	0.7857
sym13	0.8954	1	0.1857
sym19	0.9615	0.9153	0.7714
dmey	0.9688	0.9492	0.8

D1, D2, and D3

	Accuracy	Precision	Recall
db1	0.9945	0.9718	0.9857
db2	0.9982	0.9859	1
db3	1	1	1
db5	0.996	0.9722	1
db9	0.9872	0.92	0.9857
db13	0.9908	0.9452	0.9857
db19	0.9908	0.9452	0.9857
coif1	0.9945	0.9589	1
coif2	0.9982	0.9859	1
coif3	0.9945	0.9589	1
coif4	0.9908	0.9452	0.9857
coif5	0.9890	0.9444	0.9714
sym2	0.9982	0.9859	1
sym3	0.996	0.9722	1
sym5	0.9982	0.9859	1
sym9	0.9908	0.9577	0.9714
sym13	0.9927	0.9714	0.9714
sym19	0.9872	0.9315	0.9714
dmey	0.9780	0.8918	0.9429

D2, and D4

	Accuracy	Precision	Recall
db1	0.9963	0.9857	0.9857
db2	0.9982	0.9859	1
db3	0.9963	0.9722	1
db5	0.9872	0.9846	0.9143
db9	0.9963	0.9857	0.9857
db13	0.9927	0.9459	1
db19	0.9945	0.9718	0.9857
coif1	0.9963	0.9857	0.9857
coif2	0.9853	0.9844	0.9
coif3	0.9853	0.9844	0.9
coif4	0.9963	0.9857	0.9857
coif5	0.9927	0.9583	0.9857
sym2	0.9982	0.9859	1
sym3	0.9835	0.9692	0.9
sym5	0.9872	0.9846	0.9143
sym9	0.9963	0.9857	0.9857
sym13	0.9945	0.9718	0.9857
sym19	0.9817	0.9688	0.8857
dmey	0.6147	0.2348	0.8857

Discussion from observing the classification result of SVM using 3rd-order polynomial kernel:

- 4-D, 5-D feature sets from DWT with most of the selected wavelets obviously overfit the model with the 3rd order polynomial kernel.
- Combination of *D*2 and *D*4 still performs fairly well
- SVM with 3rd-order polynomial kernel does not perform as well as that with 2nd-order polynomial kernel.

4.7.5 4th-Order Polynomial Kernel

D1, D2, D3, D4, and D5

	Accuracy	Precision	Recall
db1	0.9963	1	0.9714
db2	0.5376	0.1604	0.6143
db3	0.7468	0.25	0.4857
db5	0.9560	0.75	0.9857
db9	0.9596	0.7667	0.9857
db13	0.9780	0.8625	0.9857
db19	0.9780	0.8625	0.9857
coif1	0.5541	0.1554	0.5571
coif2	0.9486	0.7386	0.9285
coif3	0.9541	0.7419	0.9857
coif4	0.9651	0.7931	0.9857
coif5	0.9761	0.8519	0.9857
sym2	0.5431	0.1518	0.5571
sym3	0.739	0.26	0.5571
sym5	0.9578	0.7582	0.9857
sym9	0.9615	0.7753	0.9857
sym13	0.9761	0.8608	0.9714
sym19	0.9780	0.8625	0.9857
dmey	0.9780	0.8625	0.9857

D1, D2, D3, and D4

	Accuracy	Precision	Recall
db1	0.9725	0.839506	0.9714
db2	0.9945	0.971831	0.9857
db3	0.9890	0.944444	0.9714
db5	0.996	0.985714	0.9857
db9	0.9853	0.907895	0.9857
db13	0.9615	0.775281	0.9857
db19	0.9321	0.657143	0.9857
coif1	0.9963	0.985714	0.9857
coif2	0.9908	0.945205	0.9857
coif3	0.9890	0.932432	0.9857
coif4	0.9853	0.907895	0.9857
coif5	0.9413	0.69	0.9857
sym2	0.9963	0.985714	0.9857
sym3	0.9890	0.944444	0.9714
sym5	0.996	0.985714	0.9857
sym9	0.9743	0.841463	0.9857
sym13	0.9137	0.6	0.9857
sym19	0.9541	0.741935	0.9857
dmey	0.8550	0.451613	0.6

D1, D2, and D3

	Accuracy	Precision	Recall
db1	0.625688	0.24812	0.942857
db2	0.994495	0.958904	1
db3	0.998165	0.985915	1
db5	0.994495	0.958904	1
db9	0.992661	0.971429	0.971429
db13	0.990826	0.945205	0.985714
db19	0.990826	0.957746	0.971429
coif1	0.99633	0.972222	1
coif2	0.99633	0.972222	1
coif3	0.988991	0.932432	0.985714
coif4	0.992661	0.958333	0.985714
coif5	0.988991	0.944444	0.971429
sym2	0.99633	0.972222	1
sym3	0.99633	0.972222	1
sym5	0.998165	0.985915	1
sym9	0.988991	0.944444	0.971429
sym13	0.981651	0.894737	0.971429
sym19	0.987156	0.931507	0.971429
dmey	0.985321	0.942857	0.942857

D2, and D4

	Accuracy	Precision	Recall
db1	0.9982	1	0.9857
db2	0.9963	0.9857	0.9857
db3	0.9982	1	0.9857
db5	0.9963	0.9857	0.9857
db9	0.9798	0.8734	0.9857
db13	0.9890	0.9324	0.9857
db19	0.9945	0.9718	0.9857
coif1	0.9963	0.9857	0.9857
coif2	0.9890	0.9324	0.9857
coif3	0.9908	0.9452	0.9857
coif4	0.9798	0.8734	0.9857
coif5	0.9761	0.8519	0.9857
sym2	0.9963	0.9857	0.9857
sym3	0.9963	0.9857	0.9857
sym5	0.9963	0.9857	0.9857
sym9	0.9780	0.8625	0.9857
sym13	0.9780	0.8625	0.9857
sym19	0.9908	0.9452	0.9857
dmey	0.6330	0.2538	0.9571

Discussion from observing the classification result of SVM using 3rd-order polynomial kernel:

- 4-D, 5-D feature sets from DWT with most of the selected wavelets obviously overfit the model with 4th-order polynomial kernel.
- Combination of *D*2 and *D*4 still performs fairly well
- SVM with 4th-order polynomial kernel does not perform as well as that with 2nd-order and 3rd-polynomial kernel.

4.7.6 Discussion of the Exhaustive Search

Since 4th-order polynomial does not perform as well as the polynomial kernels with lower order, kernels which map the attributes onto an even higher dimensions are no longer considered because of 1) the higher risk of overfitting, 2) the computation complexity of the prediction process, and 3) the memory space needed to save a large number of support vectors.

In the appendix, SVMs with Gaussian kernel do not provide any improvement over the simpler low-order polynomial kernels.

In general, the 2-dimension and 3-dimension feature sets perform better than the feature sets with higher dimensions. And the best performance of the exhaustive search comes from the SVM with the linear kernel using the quadratic feature set of D2 and D4.

4.8 Summary of the Chapter

This chapter proposes a classification strategy for arc fault detection in photovoltaic systems by using discrete wavelet transform for feature extraction and support vector machines for classification. Since the developed classifier is designed for real-time DSP/MCU applications, the computation load involved in the classification and the memory space used for support vector storage are two major concerns. Thus, linear SVM was considered first to find linearly separable 2-dimensional feature sets.

It has been shown that the rescaling strategy of the feature extraction plays a significant role in the entire classifier development. In this work, a 2-stage rescaling strategy is found to be efficient to provide linear separation between the two classes for the extracted feature set to be close to linearly separable. The cross-validation results show that the proposed combination of DWT with db9, 2-stage rescaling, and SVM with linear kernel provides a very good classification performance in practice. The simulation of real-time implementation validates that the developed classifier is capable of detecting arc fault reliably in a real-time application with fast response and high accuracy.

By observing the distribution of the 2-D feature set, by mapping the feature set of *D*3 and *D*4 onto *D*3, *D*4, and *D*3^2 gives us a perfect classification accuracy. It is also proven that SVM with the linear kernel using 2-D semi-quadratic feature sets performs the best in the exhaustive search for the best combination of feature extraction, feature selection, and classifier model.

5. CONCLUSIONS

5.1 Contributions

The research presented in this dissertation has reviewed the existing PV arc fault detection approaches and their limitations. Investigated the frequency characteristics of the DC arc, proposed and developed a PV arc detection procedure by using wavelet transform for feature extraction and support vector machines for classification

In sections, a new approach for arc analysis in DC PV systems has been proposed based on wavelet transform. The fundamental feasibility of applying wavelet transform has been presented. A comparison between the Fourier transform method and the proposed wavelet transform method has been studied with both simulation analysis and experimental results.

The presence of switching harmonics and ambient electrical noise can mask the arc signal, making detection of an arc difficult. Fourier analysis is usually not able to discover transient signals and abrupt changes like sudden arc faults and arc flashes. If the duration of the arc flash lasts for a very short period of time in comparison with the sampling window of FFT, it is likely that the arc flashes will not be observable. However, the wavelet transform is extraordinarily effective in detecting the exact instant the signal changes. The results suggest that the wavelet transform approach is not only capable of analyzing arc fault in DC systems but that it also provides a more readily detectable signal and better performance than the FFT method.

In subsequent work, arc fault signals in the presence of inverter noise have been further studied by using waveforms synthesized from real-world PV system voltages and currents. These waveforms are comprised of superimposed arcing and inverter electrical noise at a user-specified arc-signal-to-noise ratio. The test results using the synthesized test signals coincide with preceding theoretical analysis.

Section 4 proposes a technique for arc fault detection in photovoltaic systems by using discrete wavelet transform for feature extraction and support vector machine for decision making. Since the developed classifier is designed for real-time DSP/MCU applications, the computation load involved in the classification and the memory space used for support vector storage are two major concerns. Thus, linear SVM was first considered in this paper.

It was shown that the rescaling strategy of the feature extraction plays a significant role in the entire classifier development. In this work, a 2-stage rescaling strategy is found to be efficient to provide linear separation between the two classes for the extracted feature set to be linearly separable. The cross-validation results show that the proposed combination of DWT, 2-stage rescaling, and linear SVM provides accurate prediction performance in practice. The simulation of real-time implementation validates that the developed classifier is capable of detecting arc fault reliably in a real-time application with fast response and high accuracy.

By observing the distribution of the 2-D feature set, by mapping the feature set of *D*3 and *D*4 onto *D*3, *D*4, and *D*3^2 gives us a perfect classification accuracy. It is also proven that SVM with the linear kernel using 2-D semi-quadratic feature sets performs

the best in the exhaustive search for the best combination of feature extraction, feature selection, and classifier model.

With the exhaustive feature selection search, it has been demonstrated that the optimal feature extraction and feature selection method is possible by evaluating the metrics pair: precision and recall.

5.2 Future Work

- The major obstacle from carrying this work further is to collect noise signatures from different arc-faults and inverters in all possible working conditions in order to build a universal database which truthfully represents the underlying distribution of real-world arcing and non-arcing events.
- Except for current measurement, other possible measurements (such as voltage, temperature, irradiance, etc.) can be incorporated and develop a more comprehensive feature set with a wide variety of features.
- The algorithm can be further developed into a multiclass classification algorithm which is capable of detecting and distinguishing all kinds of hazardous faults using the same methodology, such that the system would only need one versatile monitoring device that guarantees the robust operation of the system.
- Although the application system used in this work was a PV DC collection grid, the results of the arc fault analysis algorithm can generally be applied to any DC electricity distribution systems and DC microgrids.

REFERENCES

[1] A. K. Jain, R. P. W. Duin, and M. Jianchang, "Statistical pattern recognition: a review," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 22, pp. 4-37, 2000.

[2] "Trends 2015 in Photovoltaic Applications - Survey Report of Selected IEA Countries between 1992 and 2014," International Energy Agency - Photovoltaic Power Systems Programme (IEA - PVPS)Oct 23 2015.

[3] B. Brooks. (2015). *The heat is on: fault detection and fire prevention*. Available: http://solarprofessional.com/articles/operations-maintenance/the-heat-is-on-fault-detection-and-fire-prevention

[4] W. Bower, "Fire Safety Issues (Arc-fault Issues)," Sandia national Laboratories2010.

[5] D. C. T. L. Ran Fu, David Feldman, Kristen Ardani, Robert Margolis, "U.S. Solar Photovoltaic System Cost Benchmark: Q1 2016," National Renewable energy Laboratory (NREL) NREL/TP-6A20-66532, 2016.

[6] M. Woodhouse, R. Jones-Albertus, D. Feldman, R. Fu, K. Horowitz, D. Chung, *et al.*, "On the Path to SunShot: The Role of Advancements in Solar Photovoltaic Efficiency, Reliability, and Costs," National Renewable Energy Laboratory, U.S. Department of Energy, Golden, CO NREL/TP-6A20-65872, May 2016.

[7] M. K. Alam, F. Khan, J. Johnson, and J. Flicker, "A Comprehensive Review of Catastrophic Faults in PV Arrays: Types, Detection, and Mitigation Techniques," *IEEE Journal of Photovoltaics,* vol. PP, pp. 1-16, 2015.

[8] S. Harb, M. Kedia, Z. Haiyu, and R. S. Balog, "Microinverter and string inverter grid-connected photovoltaic system, A comprehensive study," in *39th IEEE Photovoltaic Specialists Conference (PVSC)*, 2013, pp. 2885-2890.

[9] J. Johnson, "Overview of Arc-Faults and Detection Challenges," Sandia National Laboratories, technical presentationFeb 2011.

[10] M. Rabla, E. Tisserand, P. Schweitzer, and J. Lezama, "Arc Fault Analysis and Localisation by Cross-Correlation in 270 V DC," in *59th IEEE Holm Conference on Electrical Contacts*, 2013, pp. 1-6.

[11] A. Lazkano, J. Ruiz, E. Aramendi, and L. A. Leturiondo, "Evaluation of a New Proposal for an Arcing Fault Detection Method Based on Wavelet Packet Analysis," *European Transactions on Electrical Power*, vol. 14, pp. 161 - 174, May/June 2004.

[12] N. I. Elkalashy, M. Lehtonen, H. A. Darwish, M. A. Izzularab, and A. M. I. Taalab, "Modeling and experimental verification of high impedance arcing fault in medium voltage networks," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 14, pp. 375-383, April 2007.

[13] G. D. Gregory and G. W. Scott, "The arc-fault circuit interrupter, an emerging product," in *IEEE Industrial and Commercial Power Systems Technical Conference*, 1998, pp. 48-55.

[14] C. C. Grant, "Fire Fighter Safety and Emergency Response for Solar Power Systems," Fire Protection Research Foundation2010.

[15] G. Yunmei, W. Li, W. Zhuoqi, and J. Binfeng, "Wavelet Packet Analysis Applied in Detection of Low-Voltage DC Arc Fault," presented at the 4th IEEE Conference on Industrial Electronics and Applications, ICIEA, 2009.

[16] F. M. Uriarte, A. L. Gattozzi, J. D. Herbst, H. B. Estes, T. J. Hotz, A. Kwasinski, *et al.*, "A DC Arc Model for Series Faults in Low Voltage Microgrids," *IEEE Transactions on Smart Grid*, vol. 3, pp. 2063-2070, 2012.

[17] T. Hazel. (2012) Never Take Safety for Granted: Safety As Main Design Criteria for All Electric Power Distribution Systems. *IEEE Industry Applications Magazine*. 33-41.

[18] L. Zhigang, G. Rong, L. Jun, and A. Huang, "A High-Efficiency PV Module-Integrated DC/DC Converter for PV Energy Harvest in FREEDM Systems," *IEEE Transactions on Power Electronics*, vol. 26, pp. 897-909, 2011.

[19] P. H. Schavemaker and L. V. D. Sluis, "The Arc Model Blockset," in *Preceeding of the Second Latest International Conference, EuroPES*, 2002.

[20] S. McCalmont, "Low Cost Arc Fault Detection and Protection for PV Systems," National Renewable Energy Laboratory NREL/SR-5200-60660, October 2013.

[21] "Anual Report 2015," International Energy Agencey - Photovoltaic Power Systems Programme (IEA-PVPS)May 13 2016.

[22] "Trends 2015 in Photovoltaic Applications - Executive Summary," International Energy Agency (IEA) - Photovoltaic Power Systems Programme (PVPS) IEA-PVPS T1-27:2015, October 23 2015.

[23] "SunShot Initiative - 2014 Peer Review Report," U.S. Deparment of Energy2014.

[24] "Peer Review Guide," U.S. Department of Energy, Office of Energy Efficiency and Renewable EnergyAugust 2004.

[25] "SunShot Vision Study," Solar Energy Technologies Office - U.S. Department of Energy, Washington, DC DOE/GO-10202103037, 2012.

[26] "SunShot Initiative - Tackling Challenges in Solar: 2014 Portfolio," Solar Energy Technologies Office - U.S. Department of Energy2014.

[27] R. Wiser, T. Mai, D. Millstein, J. Macknick, A. Carpenter, and S. Cohen, "On the Path to Sunshot: The Environmental and Public Health Benefits of Achieving High Penetration of Solar Energy in the United States," Lawrence Berkeley National Laboratory, National Renewable Energy Laboratory, Golden, CO NREL/TP-6A20-65628, LBNL-1004373, May 2016.

[28] M. Mehos, C. Turchi, J. Jorgenson, P. Denholm, C. Ho, and K. Armijo, "On the Path to SunShot: Advancing Concentrating Solar Power Technology, Performance, and Dispatchability," National Renewable Energy Laboratory, Sandia National Laboratories, Golden, CO NREL/TP-5500-65688, SAND2016-2237, May 2016.

[29] B. Palmintier, R. Broderick, B. Mather, M. Coddington, D. Baker, F. Ding, *et al.*, "On the Path to SunShot: Emerging Issues and Challenges in Integrating Solar with the Distribution System," National Renewable Energy Laboratory, Sandia National Laboratories, Massachusetts Institute of Technology, Golden, Co NREL/TP-5D00-65331, SAND2016-2524, May 2016.

[30] D. Feldman and M. Bolinger, "On the Path to SunShot: Emerging Opportunities and Challenges in Financing Solar," National Renewable Energy Laboratory, Lawrence Berkeley National Laboratory, Golden, CO NREL/TP-6A20-65638, May 2016.

[31] P. Denholm, K. Clark, and M. O'Connell, "On the Path to SunShot: Emerging Issues and Challenges in Integrating High Levels of Solar into the Electrical Generation and Transmission System," National Renewable Energy Laboratory, Golden, CO NREL/TP-6A20-65800, May 2016.

[32] D. Chung, K. Horowitz, and P. Kurup, "On the Path to SunShot: Emerging Opportunities and Challenges in U.S. Solar Manufacturing," Natonal Renewable Energy Laboratory, Golden, CO NREL/TP-7A40-65788, May 2016.

[33] G. Barbose, J. Miller, B. Sigrin, E. Reiter, K. Cory, J. McLaren, *et al.*, "On the Path to SunShot: Utility Regulatory and Business Model Reforms for Addressing the Financial Impacts of Distributed Soalr on Utilities," Lawrence Berkeley National Laboratory, National Renewable Energy Laboratory, Golden, CO NREL/TP-6A20-65670, LBNL-1004371, May 2016.

[34] H. Zhang, T. Chen, and W. Li, "Arc Fault Signatures Detection on Aircraft Wiring System," in *6th World Congress on Intelligent Control and Automation*, 2006, pp. 5548 - 5552.

[35] K. Koziy, B. Gou, and J. Aslakson, "A Low-Cost Power-Quality Meter With Series Arc-Fault Detection Capability for Smart Grid," *IEEE Transactions on Power Delivery,* pp. 1-1, April, 2013.

[36] W. David Chan Tat and Y. Xia, "A novel technique for high impedance fault identification," *IEEE Transactions on Power Delivery,* vol. 13, pp. 738-744, Jul 1998.

[37] C. E. Restrepo, "Arc Fault Detection and Discrimination Methods," in *53rd IEEE HOLM Conference on Electrical Contacts* 2007, pp. 115-122.

[38] F. M. Uriarte and V. Centeno, "High-impedance fault detection and localization in distribution feeders with microprocessor based devices," in *Proceedings of the 37th Annual North American Power Symposium*, 2005, pp. 219-224.

[39] "UL 1699 - Arc Fault Circuit-Interrupters," ed.

[40] R. F. Dvorak and K. B. Wong, "Arc Fault Circuit Interrupter System," U.S. Patent 7,253,637 B2, Aug 7, 2007.

[41] M. McMahon and K. R. Morgan, "Electrical wiring device with protective features," U.S. Patent 2012/0154972 A1, Jun 21, 2012.

[42] X. Zhou, J. J. Shea, J. C. Engel, K. L. Parker, and T. J. Miller, "Arc Fault Circuit Interrupter and Method of Parallel and Series Arc Fault Detection," U.S. Patent 7,558,033 B2, Jul. 7, 2007.

[43] *Underwriter Laboratories (UL) Subject 1699B - Outline of investigation for photovoltaic (PV) DC arc-fault circuit protection*, 2013.

[44] L. Wen-jun and L. Yuan-chun, "Arc Fault Detection Based on Wavelet Packet," in *4th International Conference on Machine Learning and Cybernetics*, 2005, pp. 1783 - 1788.

[45] K. M. Armijo, J. Johnson, M. Hibbs, and A. Fresquez, "Characterizing fire danger from low-power photovoltaic arc-faults," in *40th IEEE Photovoltaic Specialist Conference (PVSC)*, 2014, pp. 3384-3390.

[46] NFPA70, "National Electrical Code 2011," in *Article 690 – Solar Photovoltaic Systems*, ed, 2011.

[47] J. Johnson, K. M. Armijo, M. Avrutsky, D. Eizips, and S. Kondrashov, "Arc-fault unwanted tripping survey with UL 1699B-listed products," in *Photovoltaic Specialist Conference (PVSC), 2015 IEEE 42nd*, 2015, pp. 1-6.

[48] J. C. Engel, "Combination AFCIs: What they will and will not do," in *IEEE Industry Application Society Electrical Safety Workshop (ESW)*, 2012, pp. 1-18.

[49] F. Boico and C. Oberhauser. SolarMagic SM73201 DC Arc Detection Evaluation Board [Online]. Available: http://www.ti.com/lit/an/snoa564a/snoa564a.pdf

[50] Y. H. Gu and M. H. J. Bollen, "Time-frequency and time-scale domain analysis of voltage disturbances," *IEEE Transactions on Power Delivery*, vol. 15, pp. 1279-1284, 2000.

[51] S. V. Narasimhan, N. Basumallick, and S. Veena, *Introduction to Wavelet Transform: A Signal Processing Approach*, 1 ed.: Alpha Science Intl Ltd, 2011.

[52] Y. Xiu, J. Shengchang, L. Herrera, and W. Jin, "DC Arc Fault: Characteristic Study and Fault Recognition," in *1st International Conference on Electric Power Equipment - Switching Technology (ICEPE-ST)*, 2011, pp. 387-390.

[53] PYWAVELETS. *Wavelet Browser*. Available: http://wavelets.pybytes.com/

[54] M. Karimi, H. Mokhtari, and M. R. Iravani, "Wavelet Based On-Line Disturbance Detection for Power Quality Applications," *IEEE Transactions on Power Delivery*, vol. 15, pp. 1212 - 1220 Oct 2000.

[55] P. Chengzong and M. Kezunovic, "Fast Distance Relay Scheme for Detecting Symmetrical Fault During Power Swing," *IEEE Transactions on Power Delivery*, vol. 25, pp. 2205-2212, Oct 2010.

[56] W.Zhao, Y.H.S., and Y.Min, "Wavelet Analysis Based on Scheme for Fault Detection and Classification in Underground Cable Systems," *Electric Power System Research,* vol. 53, pp. 23-30, 5 January 2000.

[57] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, "Wavelet Toolbox - User's Guide ", ed, 2013.

[58] K. H. Kashyap and U. J. Shenoy, "Classification of Power System Faults Using Wavelet Transforms and Probabilistic Neural Networks," in *International Symposium on Circuits and Systems*, 2003, pp. 423 - 426.

[59] A. Jensen and A. l. Cour-Harbo, *Ripples in Mathematics: the Discrete Wavelet Transform*, Springer ed., 2001.

[60] W. Li, A. Monti, and F. Ponci, "Fault Detection and Classification in Medium Voltage DC Shipboard Power Systems With Wavelets and Artificial Neural Networks," *IEEE Transactions on Instrumentation and Measurement,* vol. PP, pp. 1-1, 2014.

[61] C. Parameswariah and M. Cox, "Frequency characteristics of wavelets," *IEEE Transactions on Power Delivery*, vol. 17, pp. 800-804, 2002.

[62] S. Mallat, *A Wavelet Tour of Signal Processing (Wavelet Analysis & Its Applications)*, 2 ed.: Academic Press, 1999.

[63] L. Zhang and P. Bao, "Edge detection by scale multiplication in wavelet domain," *Pattern Recognition Letters,* vol. 23, pp. 1771-1784, 12// 2002.

[64] J. L. Guardado, S. G. Maximov, E. Melgoza, J. L. Naredo, and P. Moreno, "An improved arc model before current zero based on the combined Mayr and Cassie arc models," *IEEE Transactions on Power Delivery,* vol. 20, pp. 138-142, 2005.

[65] P. H. Schavemaker, "Arc Model Blockset- User's Guide," D. U. o. Technology, Ed., ed, 2001.

[66] J. Johnson and K. Armijo, "Parametric study of PV arc-fault generation methods and analysis of conducted DC spectrum," in *40th IEEE Photovoltaic Specialist Conference (PVSC)*, 2014, pp. 3543-3548.

[67] Y. Xiu, L. Herrera, H. Yi, and W. Jin, "The Detection of DC Arc Fault: Experimental Study and Fault Recognition," in *27th Applied Power Electronics Conference and Exposition (APEC)*, 2012, pp. 1720-1727.

[68] F. Erhard, B. Schaller, and F. Berger, "Field test results of serial DC arc fault investigationson real photovoltaic systems," in *Power Engineering Conference (UPEC), 2014 49th International Universities*, 2014, pp. 1-6.

[69] W. Zhan and R. S. Balog, "Arc fault and flash detection in DC photovoltaic arrays using wavelets," in *39th IEEE Photovoltaic Specialists Conference (PVSC)*, 2013, pp. 1619-1624.

[70] J. Johnson, B. Pahl, C. Luebke, T. Pier, T. Miller, J. Strauch, *et al.*, "Photovoltaic DC Arc Fault Detector testing at Sandia National Laboratories," in *37th IEEE Photovoltaic Specialists Conference (PVSC)*, 2011, pp. 003614-003619.

[71] J. J. Kenneth M. Armijo, Richard K. Harrison, Kara E. Thomas, Michael Hibbs, Armando Frequez, "Quantifying Photovoltaic Fire Danger Reduction with Arc-Fault Circuit Interrupters," presented at the European PV Solar energy Conference and Exhibition, Amsterdam, Netherlands, 2014.

[72] T. M. Mitchell, *Machine Learning*, 1 ed.: McGraw-Hill Education, 1997.

[73] A. Ng. *Machine Learning - Stanford University*. Available: https://www.coursera.org/learn/machine-learning#

[74] V. N. Vapnik, *Statistical Learning Theory*, 1 ed. New York: John Wiley & Sons, 1998.

[75] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," presented at the Proceedings of the 23rd international conference on Machine learning, Pittsburgh, Pennsylvania, USA, 2006.

[76] K. K. Shukla and A. K. Tiwari, *Efficient Algorithms for Discrete Wavelet Transform: With Applications to Denoising and Fuzzy Inference Systems*: Springer, 2013.