Hunting Cyclic Energy Wasters

Jim Paris, John S. Donnal, Robert Cox, and Steven Leeb, *Fellow, IEEE*

***Abstract*—Smart grid and smart meter initiatives seek to enable energy providers and consumers to intelligently manage their energy needs through real-time monitoring, analysis, and control. An essential element of intelligent management and control is access to information. The nonintrusive load monitor (NILM) identifies the operation of electrical loads from an aggregate service, making installation inexpensive, and speeding data collation and analytics. Smart meters are likely, in the near future, to be tasked with finding energy waste without requiring unreasonable demands for communication bandwidth. This paper presents NilmManager and NilmDB; tools for finding electromechanical energy wasters with a minimum of network bandwidth.**

***Index Terms*—Building efficiency, cycling systems, fault detection and diagnostics, intelligent metering, smart grid.**

I. Introduction

THE U.S. DEPARTMENT of Energy has identified "sensing and measurement" as one of the "five fundamental technologies" essential for driving the creation of a "smart grid" [1]. Basic energy-scorekeeping can be accomplished with the sample rates and computation capabilities provided by conventional digital "smart meters." However, future smart metering and control scenarios are likely to demand considerably more data analytics than are currently provided. One of the largest roadblocks to effective analytics, particulary for finding energy waste, is the disparity of scale inherent in data collection, transmission, and processing, which limits both the sample rate and resolution with which power consumption information is collected. Several high profile attempts, such as Google PowerMeter and Microsoft Hohm, have failed to find a model for analysis and delivery of energy information that avoids chasing pennies of waste with dollars of information technology infrastructure.

In this paper, we deploy a new time-series database available in [10], nonintrusive load monitoring database (NilmDB), custom developed for tackling energy analytics problems, including energy scorekeeping, power quality monitoring, and critical system diagnostics. This paper focuses on diagnostics and the early detection of energy waste. A cloud-computing application, NilmManager, can be used to manipulate and visualize data stored by NilmDB. NilmManager can be configured to provide tools for identifying faults that waste energy in subtle, difficult to detect ways.

Electromechanical loads consume the majority of generated electrical energy, and they often operate "mission critical" processes in residential, commercial, and industrial cites. These loads have been the focus of intense analysis for fault-detection and diagnostics for decades in buildings and in transportation systems [2], [3]. Both deterministic and stochastic approaches to fault detection have been studied extensively; excellent surveys can be found in [4] and [5]. A variety of electromechanical loads, important both for their commercial and industrial utility and also because of their energy consumption, run cyclically. These systems operate under some form of closed loop control intended to regulate an environmental or operating set point. For example, a motor may run periodically to create liquid refrigerant in an HVAC system for cooling (temperature control), or to compress air for pneumatic tools (pressure control). In pathological situations, which arise all too often in the field, closed-loop control can lead to difficult-to-detect energy waste. That is, detecting subtle problems that do not result in complete system failure, but that rather result in gradual degradation of system performance, can be a unique challenge. This paper presents tools for analyzing subtle changes in the behavior of cycling systems, changes that degrade operation progressively over time. We refer to this process as "hunting" or searching for the subtle clues of gradually developing pathological conditions.

For example, vapor-compression refrigeration and air-conditioning systems typically employ a hysteretic control to operate a compressor cyclically, generating compressed working fluid in the condenser when cooling is needed. Unfortunately, a wide range of fault conditions can leave the system operating, but at reduced efficiency. Reference [6] conducted a survey of 6000 distinct faults over a six year period, indicating that many faults, including 82% of evaporator fouling faults and 86% of condenser fouling faults, failed to result in a "loss of comfort." This means that the faulted systems continued to operate at reduced efficiency, consuming extra electrical energy with no obvious sign of trouble to building occupants, who continue to observe an acceptable room temperature with easily ignored increases in electrical bills. In mission-critical environments, e.g., in oil refineries or warships, we have observed that closed-loop control in cycling systems can put an order of magnitude more wear on a system than would be normally expected, leading to early and expensive equipment failures that produce mission cripples.

Manuscript received September 2, 2013; revised February 3, 2014, April 12, 2014, and July 14, 2014; accepted August 12, 2014. Date of publication September 4, 2014; date of current version October 17, 2014. This work was supported in part by the BP-Massachusetts Institute of Technology (MIT) Research Alliance, in part by the Office of Naval Research Structural Acoustics Program, in part by the MIT Energy Initiative, and in part by the Grainger Foundation. Paper no. TSG-00717-2013.

J. Paris, J. Donnal, and S. Leeb are with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: sbleeb@mit.edu).

R. Cox is with the Department of Electrical Engineering and Computer Engineering, University of North Carolina at Charlotte, Charlotte, NC 28223 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2014.2348532

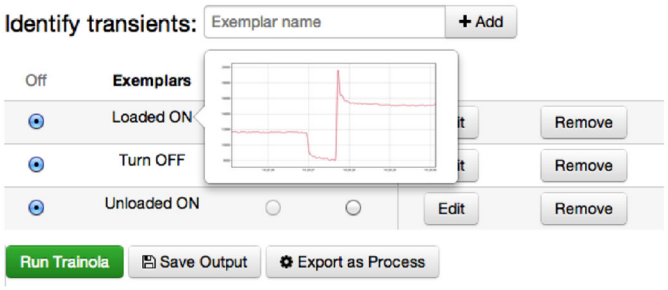


Fig. 1. Identifying load transients using the NilmManager web interface.

II. NILM and Cycling Systems

With the right monitoring tools, hunting cyclic energy wasters is fun and rewarding. The periodic actuation of a pump, vacuum, or compressor creates a cycling system in which electrical power usage follows a regular cycle based on the characteristics of the monitored mechanical variable (pressure, temperature, etc). Even non-electrical equipment, like boilers, can often be monitored using power signature techniques because most such systems include ancillary electrical devices like circulation fans which indicate equipment operation. Patterns of power consumption for electromechanical devices can be used to infer operating health of the cycling system plant and to detect a variety of fault conditions.

The nonintrusive load monitor (NILM) can serve as a platform for cycling system diagnostics with a minimum of installed sensors [7]–[9]. The NILM hardware rapidly samples the current and voltage in an electrical power system to calculate harmonic envelopes of power consumption. These harmonic envelopes can be used to identify when a given load turns on and off. Patterns and parameters of cycling system transients can be extracted to perform fault detection and diagnostics (FDD).

A typical modern NILM installation requires conventional computing capabilities, e.g., essentially any modest performance Linux system can host a NILM. The NilmDB and NilmManager are written in Gcc and with other common cross-platform tools, and can in principle be run on a variety of system hardware and common computing operating systems. However, any NILM capable of detecting line harmonics and performing parameter estimation might produce over 320MB of data per hour, recording 4 billion measurements per day. The recent development of an efficient distributed database system, NilmDB [10], for nonintrusive monitoring has changed the calculus of information access bandwidth to perform remote FDD on cycling systems. NilmDB stores data as streams in local storage which can be queried and extracted over an HTTP interface. NilmDB also stores decimated copies of each stream as distinct data streams tracking minima, maxima, and averages, so requesting a day of samples or a millisecond of samples returns a similar amount of data.

NilmManager, a cloud-based web application, provides access to all deployed installations of NilmDB at various field sites through a sophisticated GUI. Through NilmManager, users can quickly view power signature data from anywhere

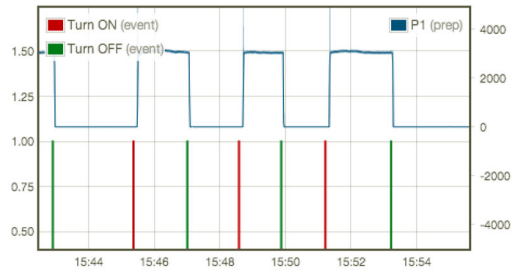


Fig. 2. Cycling events of a standard shop air compressor automatically detected by NilmManager. The turn on transient shows a characteristic step and inrush surge.

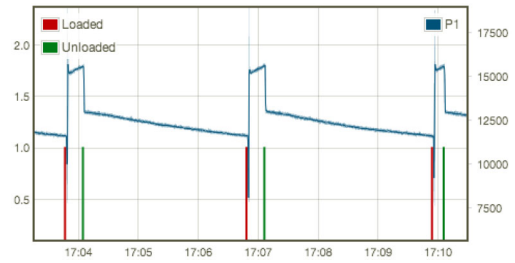


Fig. 3. Low pressure air compressor (LPAC) with load/unload control. Transients detected by NilmManager indicate change in operation state.

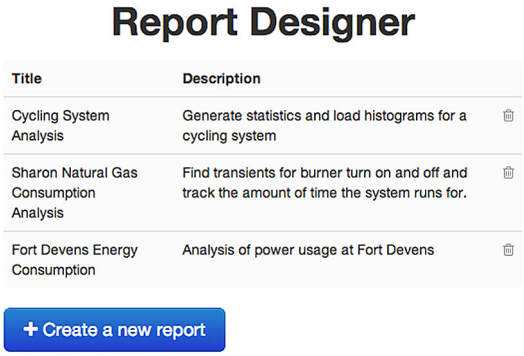


Fig. 4. Report designer.

in the world just as if the data were stored on their local machine. NilmManager further aids cycling system analysis by providing a graphic transient identification interface shown in Fig. 1. Users select a representative transient from the data and NilmManager saves the signature as an exemplar for the load. Systems with ON/OFF control have step waveform transients usually with some amount of inrush current as in Fig. 2. Systems with LOAD/UNLOAD control run continuously but consume different amounts of power depending on their state as in Fig. 3. Regardless of the type of load, once the characteristic transients are identified, the Manager issues commands to the remote NILM to automatically identify future transients of the load. This derived waveform, when processed with an appropriate model, provides a powerful diagnostic tool for cycling system analysis.

Most important, NilmManager provides a flexible remote interface for performing complex analytics, including support for fault detection and diagnostics, over the web. NilmManager avoids the need to transmit huge quantities of data from the on-site database to probe for pathologies. A user of NilmManager

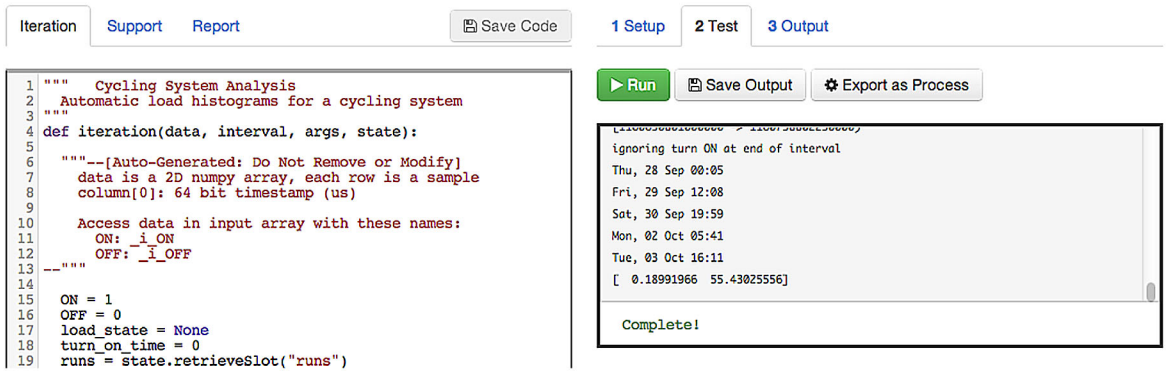


Fig. 5. Identifying load transients using the NilmManager web interface.

can create a request for a new report or analysis to be conducted using the "report designer" interface screen, accessible from any web browser, as shown in Fig. 4. Inside the "report designer" web application, the user can upload Python code for performing analysis using the remote computation capabilities of the NILM or nonintrusive smart meter. No huge quantities of data need ever be transmitted over the network. A small but powerful diagnostic code, custom tailored for a particular application, is designed and uploaded as shown in Fig. 5. Essentially any MATLAB-style mathematical analysis is available. Results from the analysis, as simple as an informative graph, table, or message, are all that need be transmitted back to a facilities owner or manager.

The next sections demonstrate the use of NilmManager for diagnostic applications.

III. Modeling Cycling Systems

Actuators in cycling systems can suffer from a wide range of problems, and nonintrusive electrical monitors can quickly and efficiently acquire the data needed to detect these problems without requiring the installation of additional sensors. Consider, for example, the vacuum pumps in a vacuum-assisted waste disposal system. These systems are commonly used for sewage collection aboard ships and aircraft and in many municipalities with flat terrain and high groundwater tables. Many cities, including New York, have recently started to deploy such systems for solid-waste collection since they reduce the emissions and costs associated with trucks. Fig. 6 includes data from an example installation designed to collect sewage aboard one of the U.S. Coast Guard's "famous"-class cutters. Under normal conditions, two vacuum pumps operate periodically and in alternation to create vacuum in a storage tank in the system, producing a typical pattern of pump run transients like that shown in Fig. 6(a). Over time, the pressure sensor in the storage tank can become clogged. As the sensor clogs, the controller loses the ability to detect the depressurization in the tank, and the vacuum pumps are run longer and longer, drawing down the tank pressure far below normal operating levels before triggering the faulty sensor, as shown in Fig. 6(b). A different pathology is illustrated in Fig. 6(c), in which one of the two vacuum pumps has developed a faulty seal. In this case, the controller operates the faulted pump, observes no change in tank pressure, and

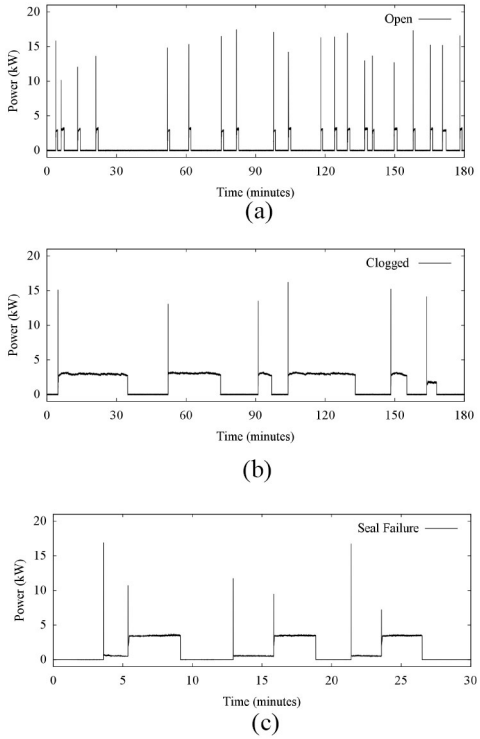


Fig. 6. Pump run transients for waste disposal system, showing normal operation and detected fault conditions. (a) Normal operation, showing typical pump run transients. (b) Operation with clogged pressure sensor. Pathological pump runs are distinguished by the extended pump run times in comparison to normal operation. (c) Operation with faulty vacuum seal. The faulty pump is distinguished by its reduced initial power draw in comparison to normal operation. The control system triggers a second pump to run in each instance after the first fails to perform as expected. The second pump consumes power at normal levels.

automatically operates the second pump immediately in seriatim to depressurize the tank. Both the sensor clog and the seal failure produce distinctive power consumption signatures that can essentially be reduced to a figure of merit or health "score" from a single cycle of operating data as compared to a baseline case.

Other pathologies can lead to individual pump runs that betray little or no presence of a fault. For example, a modest leak in the hundreds of feet of fittings and pipe that make up the system can result in pump runs with a conventional duration and shape of power consumption, but which occur

more frequently or with a different statistical distribution. In this case, nonintrusive power data can be used as base data for statistical identification of faults over longer windows of time. NilmDB and NilmManager make it easy to extract hours, days, or weeks of pump operation. This base data must be used with a carefully developed model of system behavior in order to distinguish true faults from nonpathological changes in use. Models for nonintrusive fault identification, like most FDD models, stem from an understanding of how the system behaves under normal operating conditions, and how system behavior is affected by environmental variables. System behavior in cycling systems is driven by sources or sinks of the controlled variable. For example, in a compressed-air system used to power various pneumatic tools, the actuator operating schedule is determined by the actions of the various tools following operator demand. The cycling of a residential heater, on the other hand, is affected by heat flow, and is impacted by system factors like insulation, outdoor temperature, and indoor occupancy. Typically, from the viewpoint of the actuator, external driving forces are unknown and control occurs based on the status of a state variable or derived quantity such as temperature, pressure, or fluid level. Algorithmic queuing theory provides a strong framework for system modeling of these cycling systems, a framework very useful for interpreting the data available from a NILM.

With appropriate models for the behavior of each of the individual sources and sinks, the complete response of the system can be predicted by simulation. If a leak is inserted into the system, for example, the effect will be to add a continuously active sink element. In order to determine the impact of the leak, it can be included into a model and simulated. An analysis of the model's predictions under faulty and nonfatty conditions indicates how a leak can be detected using only the operating schedule of the actuator. In general, the required analysis will differ from one class of systems to the next.

To illustrate the model-building procedure, the following three subsections present the analysis required for pneumatic systems such as those used for waste disposal and air-powered tools. This analysis is provided to illustrate the process, and any specific references to any one target system are intended only for ease of illustration. The field results presented in Section IV demonstrate how the procedure can be generalized in the diagnostic context without requiring exhaustive analysis.

A. Load Dynamics

To ensure instant availability and to provide for short periods of high demand, pneumatic systems typically have an air receiver or vacuum tank that is periodically charged by a compressor or pump. As loads draw from the pressurized reservoir, system pressure decreases. Once a certain low-pressure set point has been reached, the controller transmits a start command to the pump or compressor, causing the device to begin charging the system. Once the pressure has reached the predetermined high set point, the pump motor is either de-energized or the pumps are directed to enter a recirculation mode. In vacuum-assisted waste-disposal systems, there are two relevant types of vacuum reducing events. The first of these are

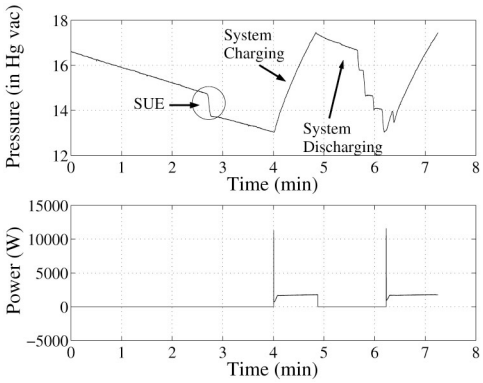


Fig. 7. Reservoir pressure and pump input power during several charge and discharge cycles in the vacuum-assisted waste-disposal system aboard SENECA. Note that the sharp drops in pressure are the direct results of system usage events (SUEs) and that the persistent background loss is the result of a small leak.

what are termed system usage events, and they result from the operation of typical system loads, such as drains. In general, these loads cause sharp drops in vacuum pressure. In addition, there are also leaks, which typically result in a persistent vacuum loss.

Fig. 7 shows both pressure and pump power in a representative cycling system aboard the USCGC SENECA. Any time that a system usage event (SUE) removes vacuum from the system, a sharp drop is observed in the measured vacuum pressure. Note that as the number of SUEs increases, the discharge period shortens and the number of pump runs increases. By comparison, the development of a leak, which causes the persistent loss shown in Fig. 7, also increases the number of pump runs and decreases the average length of a discharge period.

Assuming that the system can be accurately modeled using lumped-element approximations, the pressure loss due to a leak, which is denoted as $P_{\text{leak}}(t)$, is the solution to the first-order differential equation[1]

$$\frac{dP}{dt} = -cP \tag{1}$$

where c is a constant whose value depends upon the parameters of the system (i.e., capacity, etc.).

Assuming that N usage events have occurred during the current discharge period,[2] the reservoir pressure, $P(t)$ can be approximated as

$$P(t) = P_{\text{high}} - (\Delta P_1 + \cdots + \Delta P_N) + P_{\text{leak}}(t) \tag{2}$$

where P_{high} is the high pressure set point, ΔP_k is the pressure removed by the kth SUE, $P_{\text{leak}}(t)$ is a functional description of the leak, and $t = 0$ is defined to be the time at which the current discharge period began. In order for the pumps to energize, the total pressure loss must be sufficient to cause $P(t)$ to fall below the controller's low-pressure set point, P_{low}. Equivalently, the pumps will operate as soon as the total pressure loss, ΔP_{loss}, is greater than the difference between P_{high}

[1]For convenience, the terms "pressure" and "vacuum pressure" are used interchangeably in this section.

[2]As indicated in Fig. 7, the term "discharge period" refers to the interval when the pump is not operating.

and P_{low}. Thus, mathematically, the pump trigger condition can be written as

$$\begin{aligned} P_{\text{high}} - P_{\text{low}} &\leq \Delta P_{\text{loss}} \\ &\leq (\Delta P_1 + \cdots + \Delta P_N) + P_{\text{leak}}(t). \end{aligned} \tag{3}$$

B. System Usage Process

The rate at which individual usage events affect the system has a strong impact on the operating schedule of the pumps. In general, however, these events occur at random intervals, as human users typically operate the drains when needed.

The determination of the usage process in a cycling system amounts to the selection of an appropriate queueing model. These models originated in the early 1900s, when they were investigated by Erlang for purposes of developing automatic telephone exchanges. Just as in the case of the telephone problem, cycling systems have a community of users that both request service at random intervals and require service for random lengths of time. The selection of an appropriate model requires consideration of the statistics of the arrival process, the statistics of the service times, and the number of servers [11].

In the waste-disposal example, the system usage process is approximated as an M/D/∞ queue. The name of this standard queueing model, written in Kendall's notation, makes reference to the way in which the model addresses each of the three considerations listed previously. In this nomenclature, the first character describes the arrival process, the second character describes the service time distribution, and the third character lists the number of servers. In terms of the M/D/∞ queue, this means that the arrival process is Poisson, the service times are deterministic, and the number of servers is infinitely large [11]. Field data has shown that the Poisson process reasonably approximates the arrivals in a waste-disposal system. Second, given that usage events have been shown to occur almost instantaneously, one can approximate the distribution for the service times using one deterministic value, namely zero. Third, the claim that the number of servers approaches infinity is effectively true, at least as long as the system has enough toilets and drains to guarantee that patrons will rarely, if ever, have to wait.

In the special case that the service times are zero, the combined operating schedule of the individual pneumatic loads is governed exclusively by the statistics of the Poisson arrival process, which is denoted as $N(t)$. In this model, each new arrival increases the value of $N(t)$ to the next largest integer. Assuming that the arrival rate λ does not change as a function of time, this process can be defined as follows. The process starts at $t = 0$, i.e., $N(0) = 0$, and for any times t and s such that $t > s \geq 0$, the increment $N(t) - N(s)$ is independent of $N(\tau)$ for all $\tau \leq s$. This increment has the following Poisson distribution [12]:

$$\begin{aligned} \Pr\big[N(t) - N(s) = k | N(\tau), \tau \leq s\big] &= \Pr[N(t) - N(s) = k] \\ &= \frac{[\lambda(t-s)]^k e^{-\lambda(t-s)}}{k!}. \end{aligned}$$

Since $N(0) = 0$, this becomes

$$\Pr[N(t) = k] = \frac{(\lambda t)^k e^{-\lambda t}}{k!}. \tag{4}$$

Physically, (4) expresses the probability that $N(t)$ is equal to a given integer value.

For purposes of modeling the system usage process in a cycling system, it is important to be able to predict more than just the number of arrivals in a given interval. Specifically, it is also necessary to be able to predict the time between individual arrivals. To derive the distribution for this quantity, consider each interarrival period to be a random variable T. For a moment, focus on the time to the first arrival. The probability that this event occurs during the interval $[0, t]$ is a quantity known as the cumulative distribution function (CDF) of T. Thus, by definition, this function is

$$F_T(t) = \Pr[T \leq t]. \tag{5}$$

Using both (5) and the total probability law [13], (5) can be rewritten as follows:

$$\begin{aligned} F_T(t) &= \Pr[T \leq t] \\ &= 1 - \Pr[T > t] \\ &= 1 - \Pr[N(t) = 0] \\ &= 1 - e^{-\lambda t}. \end{aligned} \tag{6}$$

From the above result, it is possible to determine the probability distribution function (PDF) for T. This function, which is denoted as $f_T(t)$, is the derivative of the corresponding CDF, $F_T(t)$. Thus, the PDF is

$$\begin{aligned} f_T(t) &= \frac{dF_T(t)}{dt} \\ &= \lambda e^{-\lambda t}. \end{aligned} \tag{7}$$

The Poisson process is memoryless, meaning that the behavior of the arrivals after any time $t \geq 0$ is itself a Poisson process that is independent of the behavior prior to that time [13]. Thus, (7) is the PDF for the time between any two arrivals.

C. Ideal Behavior of the Pump Operating Schedule

In order to predict the operating schedule of the actuator in a cycling system, the usage and load models must be combined. In this case, it is convenient to begin by first considering the following set of idealized operating conditions.

1) Each SUE reduces the reservoir pressure by an amount ΔP.
2) System usage is a homogeneous Poisson process, i.e., λ is not a function of time.
3) The system usage process resets at the beginning of each discharge period.

In this situation, the control relation presented in (2) simplifies as follows:

$$\begin{aligned} P(t) &= P_{\text{high}} - (\Delta P_1 + \cdots + \Delta P_N) + P_{\text{leak}}(t) \\ &= P_{\text{high}} - N\Delta P + P_{\text{leak}}(t). \end{aligned} \tag{8}$$

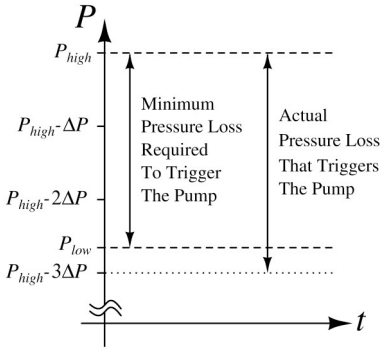


Fig. 8. Key pressure values in an example cycling system in which $N_{\max} = 3$. As shown, the pump will energize as soon as the reservoir pressure drops below P_{low}, which, in this system, can only happen after the arrival of the third SUE.

If no leaks are present in the system, then this result can be reduced even further. With $P_{\text{leak}}(t) = 0$

$$P(t) = P_{\text{high}} - N\Delta P. \tag{9}$$

Since ordinary usage is the only source of loss in (9), it is clear that the pump will energize as soon as $N\Delta P$ is large enough that $P(t) \leq P_{\text{low}}$. Given that each SUE reduces $P(t)$ by the same amount, (9) clearly implies that there is a fixed number of usage events that must transpire before the controller will command the pump to energize. For instance, in the example system used to generate Fig. 8, it is clear that the pump will begin to operate immediately following the occurrence of the third SUE. For convenience, we define the variable $N_{\max}$, which represents the maximum number of usage events that can occur during any single discharge period. As implied by Fig. 8, the value of $N_{\max}$ is simply the smallest integer that guarantees the validity of the inequality

$$P_{\text{low}} \geq P_{\text{high}} - N\Delta P. \tag{10}$$

Solving, it is found that

$$N_{\max} = \left\lceil \frac{P_{\text{high}} - P_{\text{low}}}{\Delta P} \right\rceil \tag{11}$$

where $\lceil x \rceil$ is the ceiling of x.

When operating under the simplified conditions considered here, it is possible to predict the distribution for the time T_p that elapses during any individual discharge period. Given the fact that the number of usage events impacting the system is a fixed quantity, T_p is simply the sum of the $N_{\max}$ interarrival times, that is

$$T_p = T_1 + T_2 + T_3 + \cdots + T_{N_{\max}}. \tag{12}$$

Because T_p is the sum of a fixed number of random variables, its PDF is given by the relation

$$f_{T_p}(t) = f_1(t) * f_2(t) * f_3(t) * \cdots * f_{N_{\max}}(t) \tag{13}$$

where $*$ is the convolution operator [13]. Given the interarrival model presented in (7), the PDF for T_p is

$$f_{T_p}(t) = \lambda^{N_{\max}} \frac{t^{N_{\max}-1} e^{-\lambda t}}{(N_{\max}-1)!}. \tag{14}$$

TABLE I
PARAMETERS USED IN THE IDEAL SYSTEM SIMULATION

Parameter	Value
P_{high}	17.5 inHg
P_{low}	13.5 inHg
λ	30 hr^{-1}
ΔP	1.1 inHg

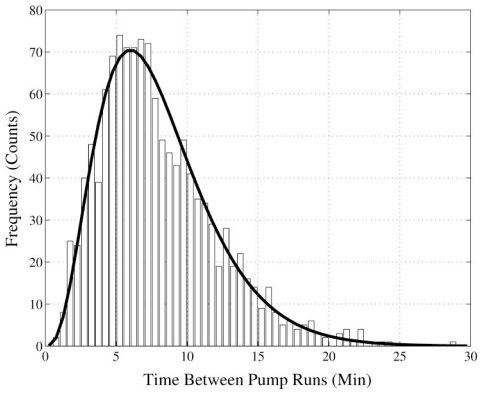


Fig. 9. Expected (solid line) and simulated (bars) frequency distributions for the discharge time, T_p, under fault-free conditions. The simulation time was 1 week.

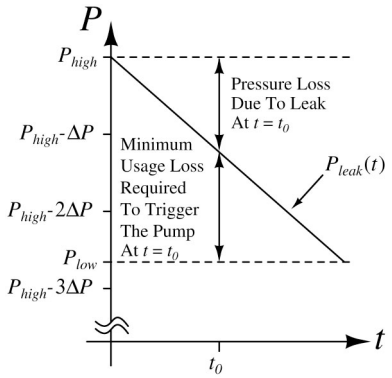


Fig. 10. Leak-induced pressure loss versus time in an example system with $N_{\max} = 3$. Note that the distance between $P_{\text{leak}}(t)$ and P_{low} is the amount of usage-induced pressure loss that would be needed in order to cause the pump to energize at time t.

In general, this two-parameter distribution is known as the Erlang PDF of order k [13]. In (14), $k = N_{\max}$. Note that in the special case that $k = 1$, the Erlang PDF reduces to the exponential distribution.

In order to demonstrate the idealized behavior described above, a simulation was designed and executed in Simulink [8], [9], [14], [15]. In the simulation, the user can set the values of all of the relevant system parameters (i.e., ΔP, P_{high}, P_{low}) as well as the value of the Poisson parameter, λ. Table I lists the values used to generate the simulated results presented graphically in Fig. 9. Also shown in that figure is the frequency distribution predicted by (14).

In the event that the ideal system develops a leak, its behavior will depart from that predicted above. A simple graphical

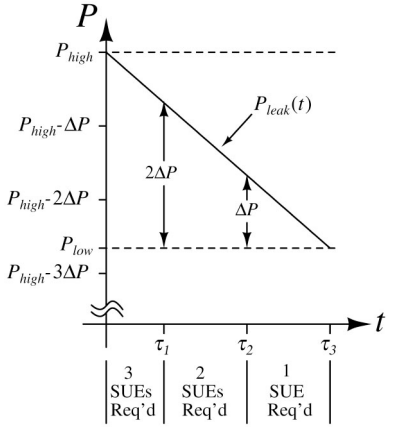


Fig. 11. Leak-induced pressure loss versus time in an example system with $N_{\max} = 3$. Note that at time τ_1, the vertical distance between $P_{\text{leak}}(t)$ and P_{low} is exactly $2\Delta P$. As a result, only 2 SUEs are required to trigger the pumps at that time.

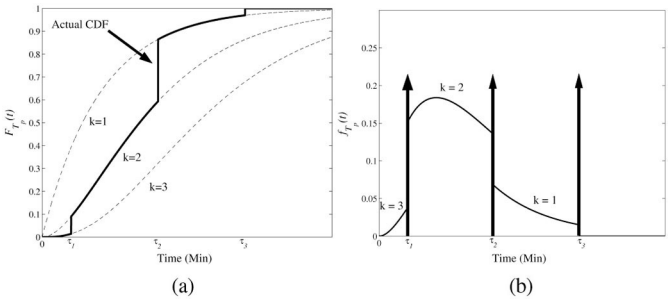


Fig. 12. Theoretical distributions for T_p when a leak exists in the ideal system. (a) CDF. (b) PDF, $f_{T_p}(t)$.

explanation for this departure is presented in Fig. 10. As shown, the leak causes the reservoir pressure to decrease continuously. As a result, the number of SUEs required to initiate pump operation becomes dependent upon the time that has elapsed since the beginning of the discharge period. In the example system used to generate Fig. 10, for instance, it is clear that the only way to force the pumps to reenergize during the first few minutes of the discharge period is for three usage events to occur in relatively short succession. If that does not happen, however, the leak will continue to reduce the pressure in the reservoir. Eventually, leak-induced loss will be large enough that only two SUEs are needed to initiate pump operation. For example, at the time t_0 shown in Fig. 10, it is clear that the pressure loss resulting from two SUEs will be sufficient to cause the pumps to energize.

With certain simplifying assumptions, a more rigorous foundation can be provided for the leak behavior described above. Specifically, for a lumped element system, the time dependence of the leak-induced pressure loss can be described mathematically using the solution to (1). If the controller is programmed so that it keeps the pressure within a relatively narrow range, then the exponential solution to (1) can be represented using the following first-order Taylor series expansion:

$$P_{\text{leak}}(t) = P_{\text{high}} e^{-ct} \approx P_{\text{high}} - \alpha_{\text{leak}} t \tag{15}$$

where $\alpha_{\text{leak}} = cP_{\text{high}}$. For a system that obeys these assumptions, there are exactly $N_{\max}$ times at which the required amount of usage-induced pressure loss decreases. The procedure used to identify these times is outlined graphically in Fig. 11. As an example, consider what happens if the discharge period lasts until time τ_1. At that exact instant, the leak has reduced the pressure to the point that the pumps will operate if the usage-induced loss is exactly $(N_{\max} - 1)\,\Delta P$. Prior to that time, however, only the occurrence of $N_{\max}$ SUEs could have forced the system pressure to fall below P_{low}. A similar procedure can be used to determine τ_2, the time at which the required usage loss drops from $(N_{\max} - 1)\,\Delta P$ to $(N_{\max} - 2)\,\Delta P$. In general, the values of the times τ_k are given by the formula

$$\tau_k = \frac{P_{\text{high}} - P_{\text{low}} - (N_{\max} - k)\,\Delta P}{\alpha_{\text{leak}}} \quad \text{for } k = 1, 2, 3, \ldots, N_{\max}. \tag{16}$$

Clearly, the behavior described above has an effect on the distribution $f_{T_p}(t)$. In particular, leak conditions cause the PDF to change at the times τ_k. For $0 < t < \tau_1$, the only possible set of events that can elicit pump operation is the arrival of $N_{\max}$ SUEs. Thus, the PDF in this region is still given by (14). Recall, however, that the pumps will automatically operate at $t = \tau_1$ if exactly $N_{\max} - 1$ SUEs impact the system during the interval $0 < t < \tau_1$. Thus, there is a certain fixed probability that the pumps will operate at that time, and it is equal to the probability that the system experienced at least $N_{\max} - 1$ but not $N_{\max}$ SUEs prior to τ_1. To determine the value of this probability, we must employ the Erlang CDF of order k, which is defined as

$$F(t; k, \lambda) = \begin{cases} \frac{\lambda^k}{(k-1)!} \int_0^t x^{k-1} e^{-\lambda x} dx, & t \geq 0 \\ 0, & t < 0. \end{cases} \tag{17}$$

For a given value of k, this equation expresses the probability that at least k arrivals occurred prior to time t. Thus, from the above arguments, the probability that the pump runs at time τ_1 can be expressed using the relation

$$F(\tau_1; N_{\max} - 1, \lambda) - F(\tau_1; N_{\max}, \lambda). \tag{18}$$

Between the times τ_1 and τ_2, only $N_{\max} - 1$ SUEs are required to initiate a pump run. The PDF in this region will again be Erlang, but it will be of a reduced order. Specifically, this section of the PDF is the Erlang distribution of order $N_{\max} - 1$, that is

$$f_{T_p}(t) = \lambda^{N_{\max}-1} \frac{t^{N_{\max}-2} e^{-\lambda t}}{(N_{\max} - 2)!}, \qquad \text{for } \tau_1 < t < \tau_2. \tag{19}$$

In order to generalize the procedure presented above, it is useful to consider the complete CDF of the variable T_p. From the above arguments, it is clear that for the times $0 < t < \tau_1$, the CDF is Erlang with $k = N_{\max}$. After $t = \tau_1$, the CDF is still Erlang, but its order is reduced by one. As a result, there is a step change in the overall CDF at $t = \tau_1$, and the height of this change is given by (18). This behavior is illustrated

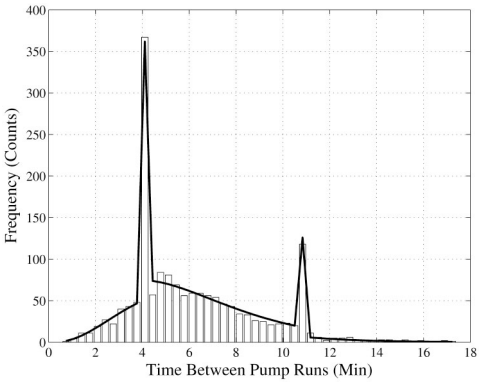


Fig. 13. Expected (solid line) and simulated (bars) frequency distributions for the discharge time, T_p, in the idealized system with a 10 inHg/hr leak rate. The simulation time was one week.

graphically in Fig. 12(a). As shown, the CDF has similar step changes at each of the τ_k, as the number of required SUEs, and thus the order of the required Erlang, decreases by one. Furthermore, the height of the "jump" at each τ_k corresponds to the probability of observing a pump run at that time.

The complete PDF can be derived by differentiating the CDF shown in Fig. 12(a). In between the τ_k, this is a straightforward operation. At the boundaries, however, the CDF has step changes, meaning that the derivative at those locations is an impulse whose area corresponds to the height of the change [16]. Thus, the complete PDF for T_p during leak conditions is

$$\begin{aligned} f_{T_p}(t) = {} & f\left(t; N_{\max}, \lambda\right)\left[u(t) - u\left(t - \tau_1\right)\right] \\ & + \sum_{i=2}^{N_{\max}} f\left(t; \alpha, \lambda\right)\left[u(t - \tau_{i-1}) - u\left(t - \tau_i\right)\right] \\ & + \sum_{i=1}^{N_{\max}-1} \left[F\left(t; \alpha, \lambda\right) - F\left(t; \alpha + 1, \lambda\right)\right]\delta\left(t - \tau_i\right) \\ & + \left[1 - F\left(t; 1, \lambda\right)\right]\delta\left(t - \tau_{N_{\max}}\right) \end{aligned}$$

where

$$\alpha = N_{\max} - i.$$

For demonstration purposes, Fig. 12(b) shows the PDF that corresponds to the CDF presented in Fig. 12(a). Fig. 13 displays a simulated frequency distribution that was obtained by inserting a small leak into the ideal system characterized by the parameters presented in Table I. The actual leak rate in this case is 10 inHg/hr. Note that the additional probability of observing a pump run at the times τ_k causes several large "pulses" to appear. For comparison purposes, Fig. 13 also shows the predicted frequency distribution.

IV. Results

This section considers two specific real-world cycling systems in which the modeling techniques described above have been used to detect faults. The first is the collection, hold, and transfer (CHT) system responsible for removing waste aboard the USCGC SENECA. The second is a compressed-air system operating a shop tool.

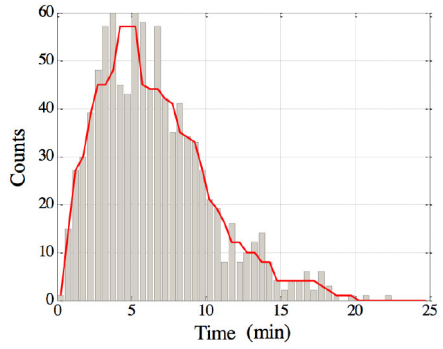


Fig. 14. Normal operation pattern of CHT system onboard USCGC SENECA.

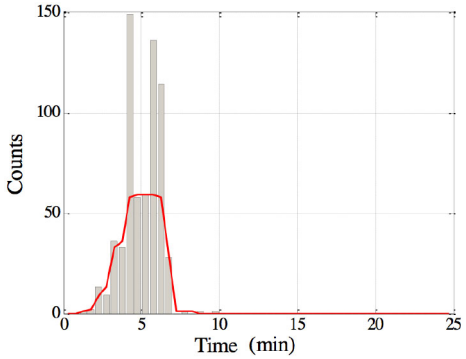


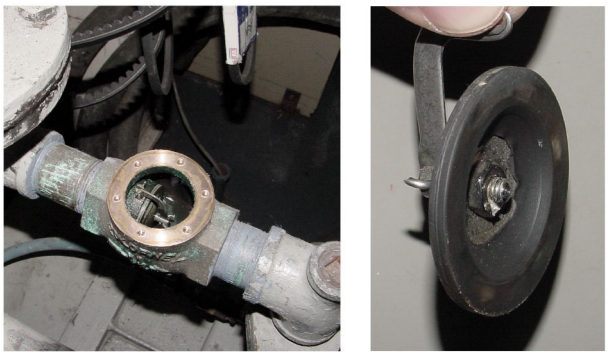
Fig. 15. Abnormal operation pattern indicative of leak onboard USCGC SENECA.

A. Coast Guard CHT Diagnostics

A NILM system was installed on board the USCGC SENECA to monitor the operation of the CHT system. The system was first characterized in a training phase to provide a baseline model. The resulting histogram is shown in Fig. 14, which is fully consistent with the nonfaulty behavior predicted by the model presented above. The vertical bars in the histogram show the counts or number of instances of the time delay indicated on the horizontal axis between pump runs. A fit-line is also illustrated to show a numerical function fit to this observed frequency distribution. During a portion of the install period, the distribution changed as the result of a large leak. The crew was not aware of the leak since no alarms or other warnings existed to alert them to the problem. Fig. 15 shows the histogram generated by the leaking CHT system. This histogram includes data taken over a 72 h period. Note the presence of the large discontinuities as anticipated by the model described above.

After alerting the crew to the suspected leak, maintenance technicians discovered the two check valves at the suction of the vacuum pumps to be faulty. The check valves are meant to shut after the vacuum pump de-energizes and maintain the vacuum in the system. Disassembly of the valves revealed pitted faces and loose components. Fig. 16(a) and (b) shows the condition of the valve and internals as it was disassembled. The valves were beyond repair and had to be replaced. After replacement, the distribution returned to normal.

A second fault-detection example from the CHT system is illustrated in Fig. 17. The length of each vacuum pump run



(a) (b)

Fig. 16. Check valve removed from the SENECAs waste-disposal system following the detection of a leak by a NILM. Note the uneven wear on the rubber face. (a) Check valve fixture. (b) Valve face.

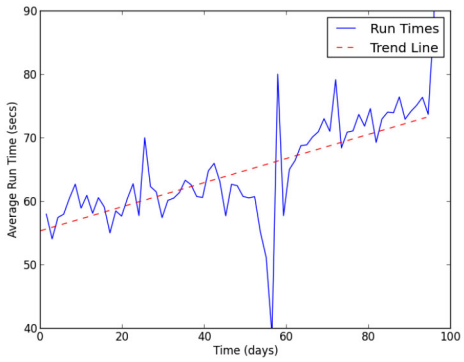


Fig. 17. Trend line for vacuum sensor failure.

can be tracked and trended using NilmManager. In the CHT system, a sensor port can clog, causing the system controller to command longer and longer pump runs, with excessive wear on the pumps and substantially increased energy consumption, as the controller creates excessively low pressures in response to a muted measurement through the sensor clog. Fig. 6(a) and (b) illustrates this lengthening of the pump runs. Over the course of a three-month period, this impending failure resulted in an increased energy cost of 30% over the baseline system operation. The gradual failure essentially adds a one horsepower parasitic load onto the shipboard power system. The trend line, extracted by a NilmManager report over the course of an underway cruise, is shown in Fig. 17, clearly illustrating the progress of the fault and offering a visual telltale, allowing the engineering duty officer to hunt for the pathology as early as possible.

B. Shop Air Compressor Diagnostics

Air compressors are commonly used to operate a variety of tools in machine shops. To model the effect of air tool usage, a computer-actuated valve periodically released pressure from the hydraulic line of an example compressor. The system generated usage patterns based on a statistical model using an exponential distribution ($\lambda = 5$ min) for interarrival times and normally distributed service times ($\mu = 4$ s, $\sigma = 1$ s). The system ran intermittently over the course of three weeks producing 122 GB of current and voltage data. At the conclusion of the experiment NilmManager was used to automate

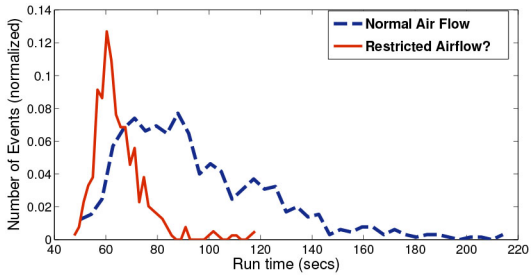


Fig. 18. Diagnostic histograms suggesting faulty valve operation. Wider distributions indicate longer compressor run times.

transient detection, which it completed in under 5 min. During the three week field test, no faults in operation were detected; however, after generating the operational histograms it became clear that the electronic valve suffered a partial breakdown partway through the experiment. Fig. 18 shows normalized histograms for normal compressor operation during the first two weeks and abnormal operation during the last week. The computer-driven actuator operated on the same statistical model throughout the experiment so the number and duration of "tool usages" is approximately the same for both histograms. This indicates that the valve itself periodically failed to open or only opened partially during the later part of the experiment.

The behavior observed in Fig. 18 is consistent with the modeling procedure described in Section III. In this case, the faulty valve causes a change in one of the primary system parameters, namely the load dynamics. The resulting change in the distribution is expected, although not in the same way predicted for the constant leak described in Section III. Although one could develop a more complete analytical model to describe the behavior observed here, it would be challenging and not necessarily rewarding. What is important to note is that one can use the basic modeling procedure to understand normal operating conditions and can then explain deviations using appropriate physical arguments. This is important, as it suggests that the model can be easily adapted as needed to the specifics of a given system without the need for excessive analysis by a subject matter expert.

V. Conclusion

Cycling systems, which represent a significant portion of energy consumption in residential, commercial, industrial, and military environments, commonly have energy-wasting faults that are difficult to detect since underlying system performance does not change significantly. This paper has demonstrated that nonintrusive electrical monitoring can be used to detect such faults in these systems without requiring the installation of additional sensors. By developing an appropriate model that considers both the physics of the target system and the statistics of its usage, one can develop powerful diagnostics that lend themselves to widespread adoption within the smart grid framework.

Acknowledgment

The authors would like to thank the U.S. Coast Guard for providing them access.

REFERENCES

[1] U.S. Department of Energy. (2008, Oct.). *The Smart Grid: An Introduction* [Online]. Available: http://energy.gov/sites/prod/files/oeprod/DocumentsandMedia/DOE_Book_Single_pages.pdf

[2] J. A. Crossman, H. Guo, Y. L. Murphey, and J. Cardillo, "Automotive signal fault diagnostics—Part I: Signal fault analysis, signal segmentation, feature extraction and quasi-optimal feature selection," *IEEE Trans. Veh. Technol.*, vol. 52, no. 4, pp. 1063–1075, Jul. 2003.

[3] J. Schmalzel, F. Figueroa, J. Morris, S. Mandayam, and R. Polikar, "An architecture for intelligent systems based on smart sensors," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 4, pp. 1612–1616, Aug. 2005.

[4] R. Isermann, *Fault-Diagnosis Systems*. Berlin, Germany: Springer-Verlag, 2006.

[5] G. Rizzoni, A. Soliman, and K. Passino, "A survey of automotive diagnostic equipment and procedures," in *Proc. Int. Congr. Expo.*, Detroit, MI, USA, 1993, pp. 1–9.

[6] M. Breuker and J. Braun, "Common faults and their impacts for rooftop air conditioners," *Int. J. HVAC+R Res.*, vol. 4, no. 3, pp. 303–318, Jun. 1998.

[7] C. Laughman *et al.*, "Power signature analysis," *IEEE Power Energy Mag.*, vol. 1, no. 2, pp. 55–63, Mar. 2003.

[8] J. Paris, "A framework for non-intrusive load monitoring and diagnostics," M.Eng. thesis, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, Feb. 2006.

[9] T. DeNucci *et al.*, "Diagnostic indicators for shipboard systems using non-intrusive load monitoring," in *Proc. 1st IEEE Elect. Ship Technol. Symp. (ESTS)*, Philadelphia, PA, USA, Jul. 2005, pp. 413–420.

[10] J. Paris, "A comprehensive system for non-intrusive load monitoring and diagnostics," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, Sep. 2013.

[11] H. C. Tijms, *A First Course in Stochastic Models*. Hoboken, NJ, USA: Wiley, 2003.

[12] A. S. Willsky, G. W. Wornell, and J. H. Shapiro, "Stochastic processes, detection and estimation," MIT Course 6.432 Course Notes, Feb. 2001.

[13] D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to Probability*. Belmont, MA, USA: Athena Scientific, 2002.

[14] J. Mosman, "Evaluation of non-intrusive load monitoring for shipboard cycling system diagnostics," M.S thesis, Dept. Mech. Eng., Massachusetts Inst. Technol., Cambridge, MA, USA, Jun. 2006.

[15] R. W. Cox, S. B. Leeb, S. R. Shaw, and L. K. Norford, "Transient event detection for nonintrusive load monitoring and demand side management using voltage distortion," in *Proc. 21st Appl. Power Electron. Conf. (APEC)*, Dallas, TX, USA, Mar. 2006, pp. 19–23.

[16] A. W. Drake, *Fundamentals of Applied Probability Theory*. New York, NY, USA: McGraw-Hill, 1967.

Jim Paris received the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2013.

John S. Donnal received the B.S. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 2007, and the M.S. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2013, where he is currently pursuing the Ph.D. degree.

His current research interests include nonintrusive load monitoring synthesis, energy harvesting, and communications systems.

Robert Cox received the doctoral degree from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2006.

He has been a Faculty Member with the Department of Electrical Engineering and Computer Engineering, University of North Carolina at Charlotte, Charlotte, NC, USA, since 2006.

Steven Leeb (F'07) received the Ph.D. degree from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 1993.

He has been a Faculty Member with the Department of Electrical Engineering and Computer Science, MIT, since 1993, where he holds a joint appointment with the Department of Mechanical Engineering. His current research interests include the development of signal processing algorithms for energy and real-time control applications.