ELSEVIER

Energy and Buildings 24 (1996) 51–64

ENERGY AND BUILDINGS

Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms

Leslie K. Norford, Steven B. Leeb
Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 8 September 1993; revised 17 September 1995

Abstract

Increased interest in energy scorekeeping, load forecasting and improved control of electricity-consuming equipment has focused attention on the instrumentation required to obtain the desired data. Work performed by the authors and other researchers has shown that individual loads can be detected and separated from rapid sampling of power at a single point serving a number of pieces of equipment, for example the electrical service entrance for an entire house or all of the central space-conditioning equipment in a commercial building. This technique has worked well in tests in houses but faces more difficult challenges in commercial buildings. We present our results for this centralized or non-intrusive load monitoring technique, applied to the space-conditioning equipment in an office/laboratory building in which equipment start-up and shut-down was centrally observed and analyzed on the basis of changes in steady-state power. We further describe our enhanced technique for distinguishing loads by matching start-up transients to known patterns, and present laboratory tests of fully automated detection hardware and software.

Keywords: Electrical load monitoring; Commercial buildings; Load-detection algorithms

1. Background

This paper focuses on metering techniques for disaggregating individual electrical loads in individual buildings. Excluded from the scope of the study are engineering models or statistical methods for separating loads, by annual energy use or load shape, in large numbers of buildings grouped by classes. The purpose of this paper is to identify the need for and methods of obtaining disaggregated load data, to present two techniques for detecting the start-up of individual loads without submetering, and to compare these two methods.

Knowledge of electricity consumption and time of use in individual buildings is vital to consumers and electric utilities. This information, typically provided by a revenue meter at the point of electrical service to a building or group of buildings, is the basis for billing and payments. Correlations of consumption and peak load with time on a monthly scale can, if extended over periods of similar weather, reveal trends, up or down, that indicate in a rough way the impact of or need for conservation activity or load control. The same monthly data, available from billing records, can be correlated with outdoor temperature to separate baseload activity from weather-dependent consumption, particularly electric heating or cooling [1].

A time scale on the order of 15–60 min, as can be obtained from utility load-recording meters, improves the temporal resolution beyond that afforded by daily or monthly totals of electricity usage and provides some information about the scheduling of electricity-consuming devices, particularly in commercial buildings where schedules are typically very regular. Major changes in electrical power are readily associated with the switching of large banks of lights, ventilation fans or chillers [2]. Irregularities in known patterns are often clues about faulty equipment operation, including unnecessary operation, failure to operate when needed, or gross changes in power when operating. However, it is difficult to take the next step with 15–60 min whole-building data and accurately partition it into major components to determine daily disaggregated load shapes, even if equipment operating schedules are fixed. Variable loads are a major problem. If, for example, there is no on-site chiller (or it is winter) and if tenant loads are nearly constant (with day-to-night variation allowed), ventilation equipment can then be distinguished from tenant loads. Any variation in total electricity consumption is ascribed to the response of the ventilation system to changing thermal loads. If a chiller is present and chiller power as well as fan power varies throughout the day, load separation solely on the basis of the metered data becomes unfeasible.

0378-7788/96/$15.00
SSDI 0378-7788(95)00958-Z

For residential buildings, major electrical loads switch on and off frequently and often irregularly. Water heating elements are thermostatically controlled and cycle to maintain tank temperature. Clothes dryers are used whenever the occupant desires. Refrigerator compressors cycle on and off in response to temperatures within the refrigerator. Air conditioner compressors work in a similar fashion, responding to the thermostat's control signal. Most loads have historically been nearly constant when equipment is in operation, although recent application of variable-speed motor drives to air-conditioning compressor motors breaks this pattern. When motors or heaters are either on or off at a constant power level, with no other states, the load disaggregation problem is one of sorting out, on a relatively fine time scale, a series of step changes in the total electricity usage. The 15–60 min time scale typical of load recording meters is too coarse, because multiple events can occur within that period.

Utilities have compensated for the limitations of typical single-point load recording meters by opting for improved spatial rather than temporal resolution. That is, submeters have been installed to definitively isolate loads. In the residential sector, submeters have been used to measure consumption of water heaters and air conditioners. In commercial buildings, submeters are used to obtain data for case studies of particular buildings, often over periods of days to months, to measure changes in electrical power before and after conservation or load control programs. Submeters are also usefully employed to determine part-load performance of equipment with variable electrical power draw and thereby improve models of equipment performance. Further, submeters are the source of data needed to validate engineering or statistical models used to estimate disaggregated load shapes for classes of buildings [3]. The capital cost of submeters, the expense of their installation, their intrusive nature (by definition on the customer side of the revenue meter, therefore requiring access to customer houses for installation, maintenance and removal) and, finally, the cost of transferring the data from the meters to a centralized facility and then performing the analysis combine to throw considerable weight against this approach.

Both modeling and metering have become more difficult in recent years as equipment part-load performance improves and power drawn by the equipment more closely tracks the demand for some sort of service; it is easier to model an on-off device than one drawing variable power. This trend is true across the board, as service includes chilled water (affecting the chiller), airflow (fans), light (due to occupancy and daylight sensors) and, recently, computation or printing (with office equipment switching to a low-power state when not in use).

This paper describes an alternative to electrical submetering, that of measuring electrical power at a single or small number of points and assigning changes in power to individual pieces of equipment. It begins with a brief description of work down by others to pioneer this technique and apply it to residential buildings. It then distinguishes commercial buildings, the subject of our research, on the basis of load detection, measurement and identification. Next is a presentation of our work to date to detect and measure loads in a commercial building by use of algorithms based on the essence of the approach used in residential buildings, that of assessing the difference in power between two steady-state levels, separated by start-up or shut-down events. The paper then describes our approach to load detection based on rapid sampling and analysis of the shape of a start-up transient. Finally, this approach is compared to the steady-state approach and conclusions are offered.

2. Non-intrusive load monitoring in residential buildings

For houses, the on–off nature of most loads suggests that a close look, temporally, at building-total electrical power can compete favorably with submeters. This view of the metering problem led researchers to develop a low-cost, microprocessor-based recorder that samples the whole-building electrical service at relatively rapid intervals. This metering approach is called non-intrusive because the meter does not cross the customer–utility boundary. The prototype non-intrusive load monitor (NILM) described in detail in Ref. [4] consisted, logically, of five steps: power measurement, detection of on or off events, clustering of similar events, matching of on and off events over time, and equipment identification. Real and reactive power was calculated from measurements of current and voltage at one-second intervals. Steady power was defined by three or more samples falling within an empirically defined tolerance; when an appliance switched on or off, power samples changed and a new steady power level was established. The difference between the two steady power levels defined an event. These events, characterized by changes in real and reactive power and a time stamp, were clustered; that is, events within an empirically established tolerance of real and reactive power were considered to be associated with one or more pieces of equipment with the same characteristics. Start-up and shut-down events for simple types of equipment yield clusters of identical magnitude but opposite sign. For a given pair of clusters, a time series of on and off events was constructed. Finally, appliance identification was made by comparing powers with known characteristics of typical appliances. The meter was subjected to a limited field test and its output compared against submetered data in four houses, using data sets of 1–2 weeks in duration. Results were typically very good for most appliances and ranged from excellent for water heaters, where the difference in energy consumption as measured by submeters and estimated by the NILM differed by more than 10% in only one of four houses, to poor for electric ranges, where rapidly flickering heating elements were often not detected by the NILM [5].

Characteristics of the meter include the following.

(i) Easy installation at the monitoring site. Unlike submetering, the NILM requires a single set of electrical ties.

(ii) Automatic load identification. The NILM can automatically identify simple 'two-state' appliances in a target building without the need for a load survey. Some caution is needed here because the NILM can potentially be fooled by abnormal equipment performance and not match clustered electrical measurements with an appliance found in its library. The NILM, as a load survey instrument, can provide building-specific information useful for conditional demand analyses of classes of buildings, described briefly in Ref. [3].

(iii) Potential for on-site data analysis. The prototype meter, however, did not perform the time-matching and identification, which were done off site on the basis of clusters of data sorted by the NILM.

Continued research on the steady-state detection algorithm used in the NILM has extended its application to machines that have more that two states of operation [4]. Refrigerators, for example, have defrosters as well as compressors, and the two-state NILM can only detect the dominant state.

The steady-state detection approach can also be extended, but as yet has not been, to include the harmonic content of the electrical current, making the NILM a potentially important platform for power quality monitoring, particularly in commercial buildings. Many loads, such as computers and other office equipment, gas discharge lighting fixtures, and adjustable-speed motor drives, can draw distorted, non-sinusoidal current waveforms. By correlating changes in harmonic content with the operation of specific equipment, the NILM could track down power-quality offenders.

3. Challenges of extending the NILM to the commercial sector

The thermal loads and HVAC equipment sizes in a storefront may differ little from those of a house, although usage patterns are typically distinct. On the other hand, the types of equipment that generate power quality problems are more likely found in commercial buildings, making power-quality monitoring a more attractive feature of a NILM that is applied to commercial buildings. The similarities and differences between these two classes of buildings can be evaluated in terms of the NILM's three chief functions: load detection, load measurement, load identification.

3.1. Load detection

Detecting loads on the basis of changes in steady-state power makes the residential NILM susceptible to confusion if two loads start up at nearly the same time. That is, there must be a discernible steady-state power level between the changes due to each of the two loads; if not, the NILM will detect a single load with real and reactive power equal to the sum of the two components. This sum will typically not cluster with any other device. When the two devices turn off, their individual powers are properly clustered. The time series will be missing start-up signals for the two devices but the NILM software can then interpret the isolated start-up power as the sum of the two shut-down events and therefore as the best explanation of the data missing from the clusters of start-up events. Problems arise as the number of loads increases; for example, the sum of the two loads may naturally cluster with an entirely different load.

This problem is mitigated by the availability of other information, namely control signals that are sent from a building automation system to turn on or off individual pieces of equipment. A NILM brought into a commercial building and attached via a communications line to the building automation system, or even incorporated into such a system, is more intrusive but more intelligent. It will know when equipment turns on and off and can therefore identify overlaps. For equipment not controlled by the building automation system, the steady-state detection algorithm is more suspect.

3.2. Load measurement

Commercial buildings are more likely to have electrical loads that vary smoothly over time, rather than undergoing one or more discrete changes in state that can be approximated by steps. Variable-speed motor drives are perhaps the best known example, but even fixed-speed motors will draw varying power when connected to centrifugal machinery that experiences variable hydraulic loads. Dimmable lighting fixtures are another example. In these cases, a primary problem is that start-up and shut-down power may well differ, making it difficult to associate these transitions with the same device unless there is information about control signals or device output (pressure, flow, light level). More subtly, variable hydraulic loads make it difficult to pin down the start-up power of pumps and fans; as will be seen, the start-up transient is prolonged and may be masked by other changes in electrical power.

3.3. Load identification

In the residential NILM, loads are identified by characteristic values of real and reactive power. In commercial buildings, substantial efforts are made to reduce reactive power and make loads appear to be primarily resistive. Fluorescent lamp fixtures driven by power-factor corrected electronic ballasts will have essentially no reactive power in steady state and cannot be distinguished from a heater or coffee pot of the same magnitude. In this case, load identification is prone to failure if based solely on steady-state load characteristics.

If the NILM is attached to a building automation system, its task changes from identification based on power, as is done in the current implementation of the residential NILM, to an analysis of power based on identification. With the equipment identity known, abnormal power readings are no longer unidentified because they do not fall within a cluster of typical power levels but instead can be evaluated as being caused by equipment faults. So, the steady-state detection algorithm would appear in principle to be viable for commercial-building HVAC equipment controlled by an automation system. Equipment not controlled by an automation system includes tenant office electronics, lights, heaters and distributed ventilation fans. Power-factor correction is increasingly prevalent on office electronics and fluorescent lamp ballasts and identification based on steady-state power may fail.

4. Detection of on–off transitions, measurement of power levels, and fault detection in commercial HVAC equipment, using steady-state load detection

A NILM attached to the electrical system serving HVAC equipment cannot only provide data for utilities and owners interested in load management but can also inform the HVAC control system that equipment has indeed responded to on–off control signals. Such would not be the case if the device were inadvertently placed under manual control or if there were an electrical fault. Traditionally, a control system relying on electrical confirmation of response to on or off signals has made use of a single current transducer per device. We have assessed whether a NILM could perform the same task, without need for individual current transducers, and whether the steady-state load detection scheme embodied in the residential NILM is appropriate for a different environment. More information about our tests is found in Ref. [6]. Here we update and summarize the analysis in Ref. [6]; in particular, we provide information about application of the residential NILM's step-change algorithm for 3 and 5 kW tolerances.

We installed a watt transducer (and not a fully developed, stand-alone NILM) on the 480 V 3-phase electrical service that provides power for HVAC equipment in two campus buildings. The transducer's output was limited to real electrical power. Work in the residential sector has shown reactive power to be very useful and indeed reactive power is computed in the transient-event detector to be described later in this paper. However, real power was sufficient to assess many of the major aspects of the functionality of a NILM in a commercial building. The equipment in the test building included two identical 500 ton centrifugal chillers and associated chilled water (50 hp) and condenser water (40 hp) pumps; two large supply fans with adjustable-speed drives (125 and 100 hp); and a number of smaller pumps and fans; maximum total power for the system was about 1000 kW. One-second average samples of the watt transducer's output were stored on a portable computer. Fig. 1 and Fig. 2 show typical data, including large-amplitude oscillations that will be discussed later and on–off transitions of HVAC equipment.

These data can be used to answer an important question about the statistical validity of power measurements made via non-intrusive monitoring. If the magnitude of the unexplained electrical power variations measured at the HVAC electrical service distribution panel or other central location is large, then only larger changes in power levels are statistically significant. Standard deviations during periods of nearly steady power were about 5 kW. This remarkably small number, less than 1% of the total, indicates that pumps and fans of moderate or large size should be detected readily, while small return and exhaust fans would not be found easily. For a signal with a 5 kW standard deviation, the minimum difference in mean power before and after an event that will reject the hypothesis that there is no change at the 95% confidence level drops from 9.8 kW when the mean is calculated from five power measurements to 5.3 kW for a sample size of 10 and 2.7 kW when the sample size is increased to 30. Of course, more samples require longer steady periods before and after a piece of equipment turns on or off.

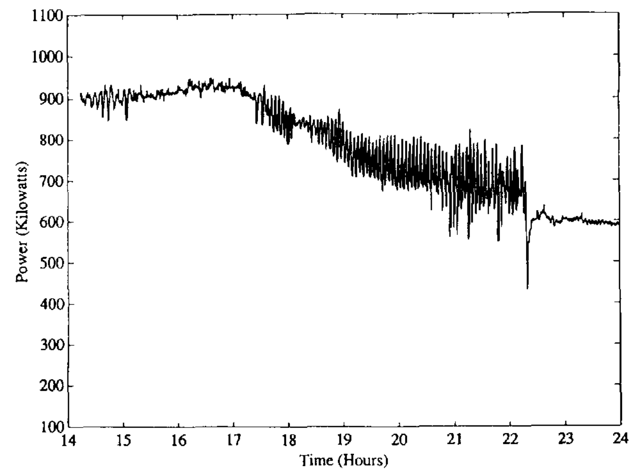

Fig. 1. Electric power at the building HVAC service entrance. A poorly tuned chiller controller operating under low-load conditions caused the large power oscillations.

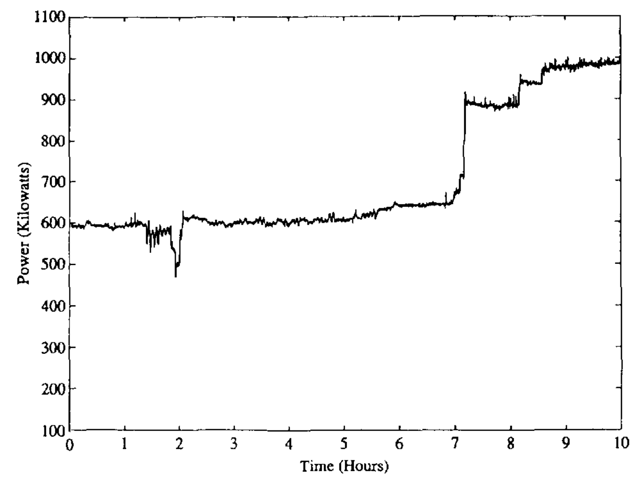

Fig. 2. Electric power at the building HVAC service entrance. A chiller and associated pumps were turned on at 7 a.m.

4.1. Pumps

Fig. 3 illustrates four on and off transitions for a 50 hp condenser-water pump, which was tested at a time when the

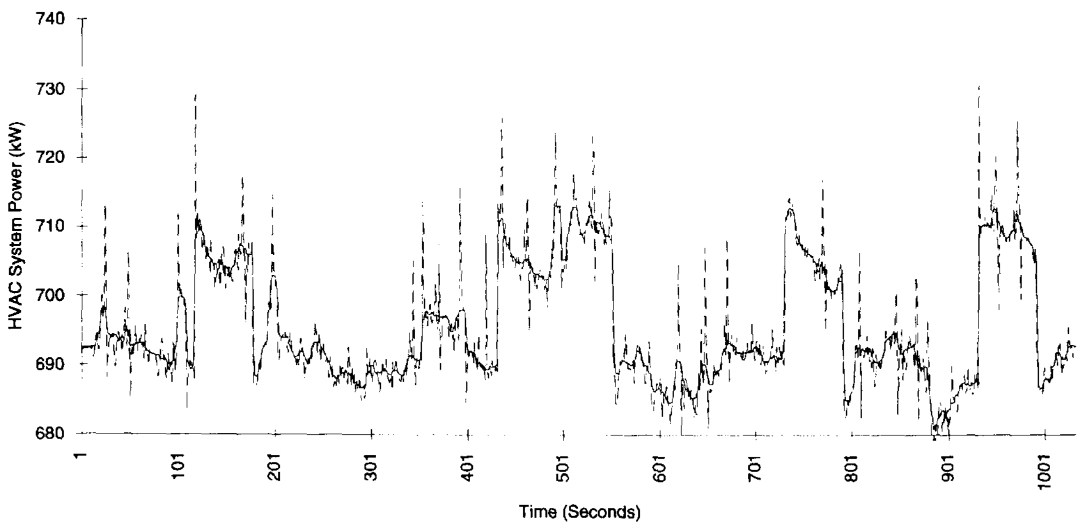

Fig. 3. Changes in electric power at the HVAC service entrance due to switching a condenser-water pump on and off. Noise in the raw data, shown as a dashed line, is removed by the median filter.

other condenser pump was in constant operation. The increase in system power when the pump is running is large relative to the 5 kW standard deviation and periods of pump operation can be visually discerned. However, the figure shows that the pump motor is more complicated than a simple two-state device with a motor start-up power surge and therefore more complicated to analyze.

(i) After start-up, there is, typically, a long period of slowly decreasing power, during which fluid pressures in the water loop are approaching equilibrium. The pump just started must reach a balance with the pump already running, at which time each pump is drawing less power than either would if running alone. This pattern is not consistent and is masked by small increases in system power.

(ii) After the test pump is turned off, the remaining pump, still in operation, is subjected to a larger load and its power increases. There is only a narrow window in time — about 5 s — during which to establish the power change of the test pump at shut-down; a longer averaging period will be affected by the increased power of the second pump. We have observed similar behavior with the two chillers.

(iii) There are periodic spikes as large as 20 kW magnitude, in many cases lasting just a single one-second sampling interval. Similar surges in power were observed with a submeter attached to a variable-speed drive fan and may be due to variable-speed-drive controllers responding to set-point changes.

How a NILM should respond to electrical impulses depends on an assessment of the cause of the impulses and the value of computing the energy associated with them. In our case, we ascribed the impulses to the operation of the variable-speed-drive controllers and sought to either ignore them or screen them out. In other cases, the impulses could be due to electrical heaters and would be of interest. Impulses trigger the edge detector used in the residential NILM, which then waits for a new period of near steady power. If they are not of interest they create a computational burden because they are treated as significant events; a NILM operating at one-second intervals would compute a near-zero change in power before and after the impulse. If they are considered of interest, the NILM sampling interval would need to be shortened to establish a steady power level at the peak of the impulse. The sampling speed must be chosen with care, as a faster speed makes the steady-state detection algorithm more likely to find spurious steady-state powers induced by other devices during a prolonged start-up transient.

A median filter is one technique that has worked well in practice to eliminate narrow spikes. Well known in the signal and image processing fields for several decades (see summary in Ref. [7]), the median filter operates by sliding a symmetrically placed window over a vector of input data. At any time, the output of the filter is the median value of the data in the window, which advances one point at a time along the stream of input data. A median filter can eliminate the spikes in the data we have recorded, which are very narrow, without sacrificing significant features of the data, which are of longer duration. We applied a window of 11 points. It is important to note that this filter differs from linear filters, which pass signals of specified frequency and are therefore poorly suited to distinguish impulses and edges, both of which have similar frequency content.

Fig. 3 shows the remarkable results of the median filter as well as the raw data. With the spikes removed from the raw data, it is now possible to consider algorithms for determining the magnitude of the changes in power. Fig. 4 shows step changes pulled out of the data by the residential NILM's algorithm, with steady state defined by 3 and 5 kW tolerances, for the first portion of the data. The performance of this algorithm merits several comments.

(i) The algorithm found the four pump tests, the first of which is shown in Fig. 4, but also reacted to numerous non-events, even after the spikes were removed. There is no apparent benefit from decreasing the tolerance, which will only produce more spurious events and make it more difficult to

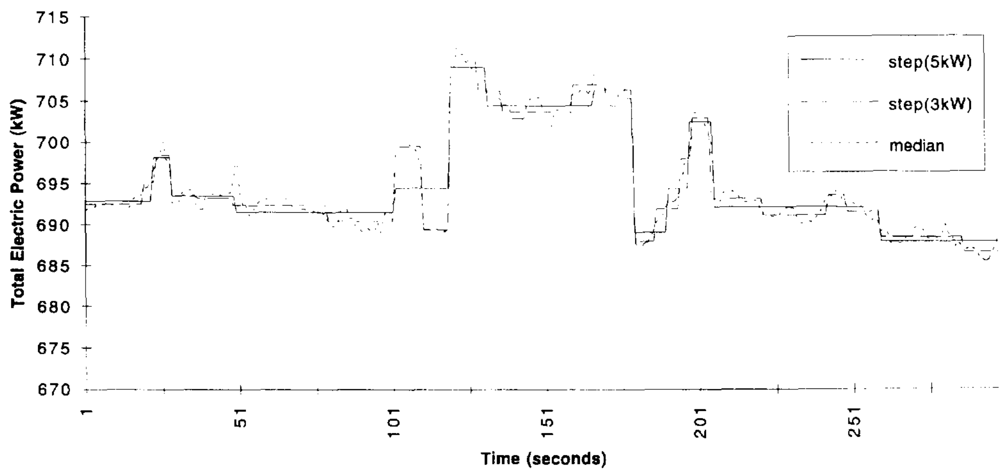

Fig. 4. Condenser pump operation, as shown by the median filter alone and by successive application of the median filter and the residential NILM step-change detection algorithm with 3 and 5 kW triggers.

establish a steady state immediately after the pump is turned on. The spurious events can be blocked by employing the algorithm only when the control signal is sent to a piece of equipment. In the absence of a control signal, the algorithm can be augmented by a transient-detection scheme, described later in this paper, that seeks to match the electrical data with templates of known devices and therefore blocks unknown and presumably uninteresting transients.

(ii) The determination of the increase in power when the pump was started is influenced by the tolerance parameter. With a 3 kW tolerance, the step changes closely track the data produced by the median filter. The average power for the four start-ups was 21.5 kW with a tight standard deviation of 1.1 kW. The power readings are sufficiently reproducible to be useful for confirmation of pump operation and for fault detection (where statistically significant changes in power over time constitute a fault). A wider tolerance is unsatisfactory; Fig. 4 shows that the 5 kW tolerance causes the pump start-up power to be substantially underestimated, by 5.8 kW.

(iii) The step-change algorithm approximates the gradual changes in power after the initial peak as a series of steps, the number of which varies across the four tests and also varies with tolerance. Without a reproducible pattern, it is impossible to model pump power as a finite-state machine. The problem with forcing the shape of the pump power curve into a series of steps can be circumvented by considering the transient to be a shape to be compared with a template, but the lack of reproducibility across the four tests argues for considerable care with this approach, which is the subject of planned future work.

(iv) The step-change algorithm with 3 kW tolerance yields an average shut-down power of 19.7 kW with a standard deviation of 2.5 kW. The power dropped, on average, 1.8 kW from the initial peak to the shut-down value. The algorithm cannot track this gradual decrease in power as it occurs but the peak start-up power and the shut-down power can still be associated with the same piece of equipment due to the control signals. The shut-down power, assumed to represent steady state because the pumps run at fixed speed and there are no throttling devices to produce a variable pressure, can be multiplied by runtime to estimate energy consumption. Use of the start-up power for energy consumption calculations will yield an overestimate and is not recommended.

To return to the three characteristics of the pump signal that complicate load detection and measurement, we conclude that shut-down power is more representative of system operation that start-up power, that small changes in power for one pump when another is switched on or off require a narrow time window to detect and may be insignificantly small, and that filtering can remove power spikes that are deemed to be uninteresting.

4.2. Fans

We recorded total HVAC electrical power at times when a 125 hp supply fan was turned off and on twice, followed by one off–on cycle for a 100 hp fan. Both fan motors are controlled by variable-speed drives (VSDs). The raw data are shown in Fig. 5, while Fig. 6 shows filtered electrical power. There were no in-rush power spikes at start-up, because the VSD electronics include a soft-start feature. The fans did not interact with other fans and, unlike the pump data, there was no evidence of increase in power shortly after shut-down. Power levels dropped after start-up, when the frequency established by the VSD was reduced from its initial value of 60 Hz to that required by the fan to maintain supply duct static pressure at set point.

The ramping of power at start-up to its maximum value took place over about 30 s and was approximated by the step-change algorithm as a series of steps, the number of which depended on the tolerance parameter. It is not possible to treat the transition as a single step with any reasonable tolerance and the start-up must therefore be modeled as a pattern of increasing steps, ending when power drops, or as a continuously increasing shape terminated by a drop in power. The two choices are similar, blurring the distinction between the

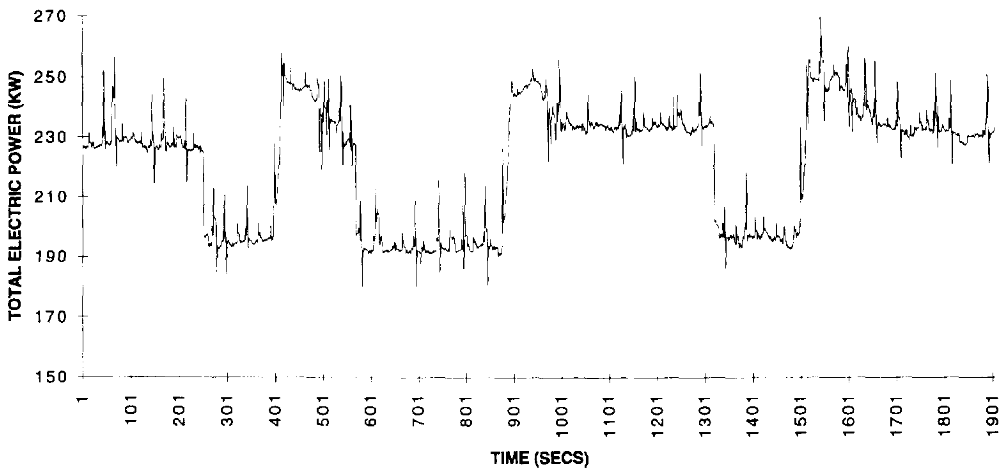

Fig. 5. Start and stop transitions for two large supply fans.

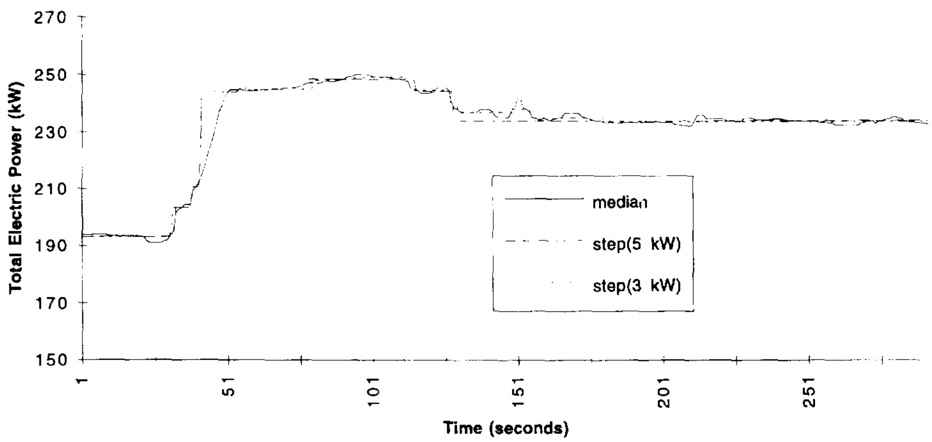

Fig. 6. Start and stop transitions as clarified by the median filter alone and by successive application of the median filter and the residential NILM step-change detection algorithm.

steady-state, step-change algorithm and the transient pattern matching algorithm to be described later.

Like the pump power, fan power drops after the initial peak. Unlike pump power, fan power at shut-down cannot be taken as a steady-state value, because pressure across the fans and flow through them both vary. Tracking fan power between start-up and shut-down is a difficult problem that will be briefly noted later. It is essential to follow the start-up transient to the short-term steady value, when the fan motor has slowed to the point dictated by the pressure controller, if there is any hope of accurately estimating fan energy use. Fig. 6 indicates that the steady-state power can be easily defined when the entire HVAC electrical service is steady over a time period of about 100 s. The speed-control signal provided to the VSDs would more precisely define steady state, but gaining access to this information complicates the load monitor.

4.3. Fault detection based on one-second power samples

Optimal control strategies have been developed by Braun et al. [8] and detection of power deviations from optimal conditions has been explored by Pape et al. [9]. The amount of information required for optimization has deterred its acceptance by industry and building owners. A NILM offers a lower cost, somewhat less informative, but still powerful approach by providing a basis for identifying what is clearly not correct, even if it is not possible to establish how to achieve what is optimal. Our monitoring of the HVAC electrical service entrance revealed two types of equipment faults that exemplify the application of the NILM to fault detection and diagnosis.

4.4. High controller gains

Fig. 1 shows power oscillations with a peak-to-peak amplitude, about 150 kW, that are so large as to be uniquely linked to the chillers. Oscillations started when the total power dropped and stopped when the power rose or, when one of the two chillers was turned off. A fast Fourier transform of the data showed a strong spike at the same frequency used by the chilled-water temperature controller. The data led us to conclude that oscillations occurred at times when the chillers were lightly loaded and were due to poorly tuned controller gains. With control gain too large, chiller power varies excessively with small changes in control input.

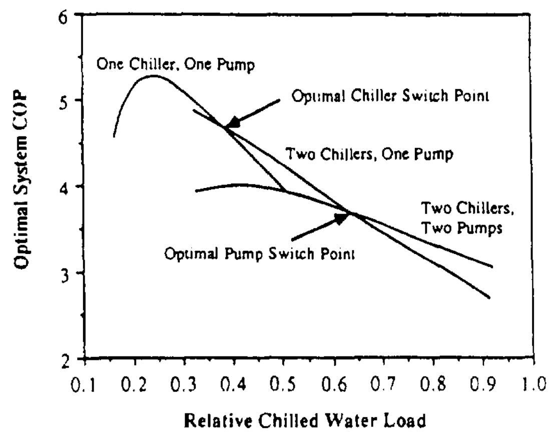

Fig. 7. Optimal switch points for a two-chiller, two-pump system, taken from Ref. [8]. Optimal switch points result in no change in power, while suboptimal switching shows power changes. A non-intrusive load monitor can reveal such changes.

4.5. Switching multiple chillers

Ref. [8] showed that optimal switch points can be defined as producing no discontinuity in system coefficient of performance (COP), as shown in Fig. 7. That is, if a second chiller is turned on at too low a cooling load, COP will drop and power will rise. Power will drop if it is turned on at too high a cooling load. Similarly, power will fall if the second chiller is turned off too late, that is, when the cooling load has dropped below the optimal switch point.

Our data in Fig. 1 show that the second chiller was turned off after the optimal cooling load switch point had been passed. Mean power dropped by about 100 kW, indicating that the combined power for both chillers exceeded the optimum by that amount for some period of time leading up to shutting down the second chiller. This information, if detected by a NILM, can guide plant operators toward more efficient plant operation.

A type of chiller staging not observed experimentally in this study concerns when to turn on the first chiller. At issue is how large a cooling load can be met by the ventilation system bringing in 100% outdoor air, and how much fan power is required. Traditional practice has maintained the supply air temperature at a fixed value, forcing the chiller on when the outside temperature approaches this set point (with a small decrement due to temperature rise across the fan). Fan power will therefore stay the same immediately after the chiller is turned on and the chiller will be running at relatively low load. Alternatively, the supply air temperature could be allowed to float upward, with the chiller turned on when the fan is running at maximum load or (less likely) the increase in fan power exceeds the power drawn by the chiller. The latter case is exactly the same as the problem of staging the second chiller. In either case, the NILM and the fan speed-control signal can be used to detect suboptimal performance.

5. Approach to transient pattern recognition

The transient electrical signal from motor-driven pumps and fans, when measured at one-second intervals, reflects the interaction of the equipment with piping systems and controls. It is clearly valuable in determining when equipment operation has reached a steady-power level and can also be used for fault detection. We have formalized the process of associating observed transients with particular pieces of equipment by developing a prototype transient-event detector which can operate on multiple time scales. At some time scale characteristic of a given piece or equipment, the transient behavior is intimately related to the physical task that the equipment performs. The turn-on transients associated with a fluorescent lamp and an induction motor, for example, are distinct because the physical tasks of igniting an illuminating arc and accelerating a rotor are fundamentally different. Further, the turn-on transients of induction motors have a characteristic shape which dilates or contracts in both magnitude and time as a function of the size of the motor. Transient profiles tend not to be eliminated even in loads which employ active waveshaping or power factor correction. These repeatedly observable turn-on transient profiles are suitable for identifying specific load classes and screening out uninteresting electrical information. We will now discuss the classification of segments of patterns from start-up signals and, briefly, the operation of the classification scheme at multiple time scales needed to detect small and large motor transients.

5.1. Pattern segments

Transient patterns of real and reactive power and higher harmonic content are derived from current and voltage samples averaged over at least one period of voltage variation at the fundamental frequency (60 Hz in the US). While work reported in this paper does not consider harmonic analysis, our second-generation NILM hardware is capable of detecting higher harmonics as an aid in load analysis. Attempting to identify complete transients is an undesirable approach because it severely cripples the ability of the NILM to separate overlapping events. Instead, a start-up signal is considered to be a time series of segments, some with substantial variation and others essentially steady. The transient event detector only searches for the significantly varying segments, denoted as v-sections; during the relatively quiescent periods the detector can respond to v-sections from other pieces of equipment. These v-sections have characteristic shapes associated with them, as will be seen.

During a training phase, either before installation or on-site, the event detector employs a change-of-mean detector [10] to segment a transient representative of a class of loads, which might be induction motors or lamps with rapid-start ballasts. This segmentation process delineates the set of v-sections that will represent a particular transient shape in the input data stream of any member of the class of equipment. The trace in Fig. 8, for example, shows the measured varia-

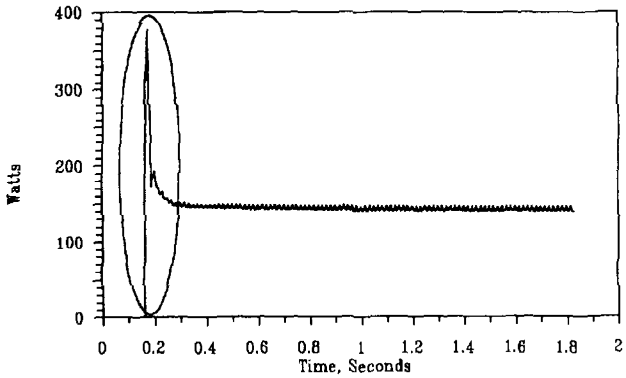

Fig. 8. Measured instant-start lamp-bank real-power transient, with v-sections noted.

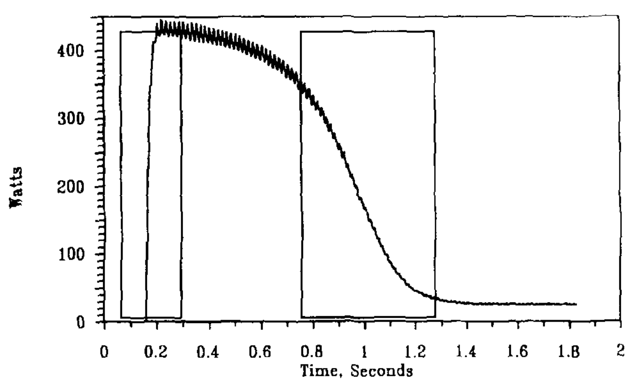

Fig. 9. Measured induction-motor real-power transient, with v-sections noted.

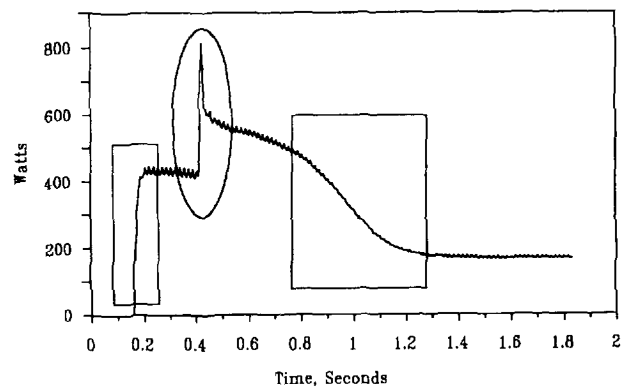

Fig. 10. Acceptable overlap of v-sections between lamps and a motor.

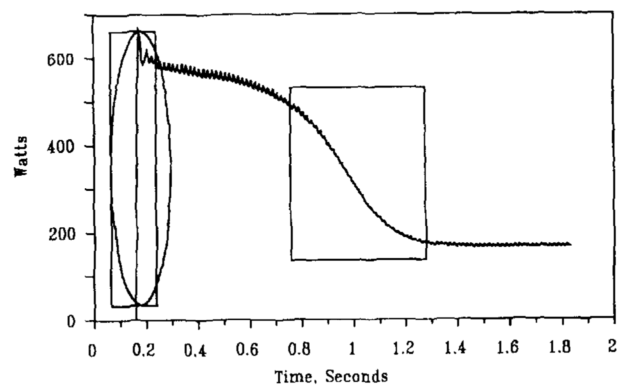

Fig. 11. Intractable overlap of v-sections.

tion in real power during the turn-on transient of an instant-start fluorescent lamp bank. Fig. 9 shows measured real power on one phase during the turn-on transient of a three-phase induction motor. The locations of the v-sections in the two waveforms, as computed by a change-of-mean detector, are approximately indicated by the ellipse in Fig. 8 and the rectangles in Fig. 9. In practice, a more complicated set of v-sections, which would include such other data streams as reactive power or higher harmonics, would be used to represent the transient profile of a load.

A complete transient identification is made by searching for a precise time pattern of v-sections. As long as each of the v-section shapes does not overlap the v-section of another device, the event detector will be able to identify the patterns of v-sections and therefore the transients. For example, the overlap of the two transients from the induction motor and the instant-start lamp bank shown in Fig. 10 is tractable because all of the v-sections for both transients are separated in time and therefore distinguishable. The overlap condition in Fig. 11 would not generally be tractable, since the instant-start v-section and the first induction motor v-section overlap severely. Since some degree of overlap is tolerable, the v-section set recognition technique will generally operate successfully in an environment with a higher rate of event generation than would a detector searching for whole, undisturbed transient shapes.

Classification of individual v-sections in the input data streams is determined by a set of pattern discriminator functions. These functions are used to compute a distance metric that locates a particular input vector in a region of a space of known transient templates. Because a v-section may appear on top of a variably large static or quasi-static level created by the operation of other loads, the discrimination process focuses on only the 'a.c.' or varying component of the v-section.

The prototype event detector employs a transversal or matched filter as a pattern discriminator, although other possibilities could be used and are discussed in Ref. [11]. Each v-section is positively identified by checking the outputs of two different transversal filters [12]. The first transversal filter scans an input data stream for a particular shape. The output of the shape transversal filter is the inner product of the a.c.-coupled and amplitude normalized template vector t_{ac} and the data vector x_{ac}. An output of unity indicates a perfect match between the template vector and the input data. Noise and slight variation in the repeatability of the v-sections will make a perfect match unlikely. In a practical system some degree of imperfection will be tolerated and any inner product within a certain tolerance of unity will constitute a match. The second filter checks the magnitude of a data segment that matches the shape of a template v-section, to ensure that a small wiggle or noise pattern that is fortuitously close in shape to a v-section template is not mistaken for an actual v-section.

Planned field tests will guide the selection of filter parameters. These tests will assess repeatability of start-up transients and their detectability. It can be argued that noisy field environments may require relaxed pattern matching tolerances or make tuning a very difficult exercise. The logical limit of this situation would be to approximate the v-sections as a composite of steps and lose the detail associated with their shape. In this case, the algorithms we have developed still have two powerful features: first, the separation of a start-up event into a series of components, necessary to sort out overlapping events, and second the ability to work over multiple time scales, to be discussed next.

5.2. Multiple time scales

Loads in a particular class which span a wide power range often exhibit transient profiles that are identical in shape but scaled in amplitude and duration. The transversal filter is suitable for identifying transient shapes over a narrowly defined time scale. The prototype event detector employs a tree-structured decomposition to search efficiently over many time scales with the transversal filters [11]. The use of the tree-structured decomposition is inspired by recent signal-processing applications of sub-band coding [13] and the discrete-time wavelet transform [14,15]. The ability of the prototype to work at multiple time scales is a significant advance, making it possible not only to detect large and small motors but also, in principle, the combined electrical–mechanical–hydraulic dynamics that characterized the lengthy pump and fan transients.

6. Prototype test equipment

The prototype event detector consists of three components: an analog preprocessor, a digital-signal processing card, and a personal computer. The event detector monitors the voltage and current waveforms on a three-phase electrical service that powers a collection of loads representative of important load classes in typical, medium to large size commercial and industrial buildings. It is used to identify the turn-on time and type of the various loads, as would be required of an NILM not connected to a building automation system, and it has no a priori knowledge of the operating schedule for the loads.

The NILM can be effectively tested in a controlled environment if that environment accurately represents the equipment found in a commercial building. In a laboratory, it is not reasonable to install equipment of the same power requirements as are found in an entire building, but it is possible to maintain the relative magnitudes of key types of equipment. The laboratory setting includes several motors and lamps. Were it possible to fully load the motors, this list of equipment would support motor:fluorescent light ratios of 0.2–10, making the laboratory mock-up match buildings with floor areas of 100–1000 m^2. With the motors normally unloaded, the ratios drop, although the start-up transients remain large due to motor in-rush current. Four loads were selected for inclusion in the initial tests of the NILM: two twin-tube instant-start fluorescent lamps with electronic ballasts, four twin-tube rapid-start fluorescent lamp fixtures with electronic ballasts, a 3-phase 1/4 horsepower induction motor, and a 3-phase 1/3 horsepower induction motor. The electrical hookup to the loads is routed through an electronically switched circuit breaker panel that activates loads with flexibility in relative timing. The pattern templates for the loads were captured during a one-time 'walk through' of the test stand. However, no data at all were collected from the large motor. Because the large and small induction motors are members of the same load class, a single transient template, appropriately scaled in amplitude and duration, was expected to prove satisfactory for identifying both motors.

7. Prototype performance

Figs. 12–16 show screen prints from the PC running the NILMscope user interface software during five of the exper-

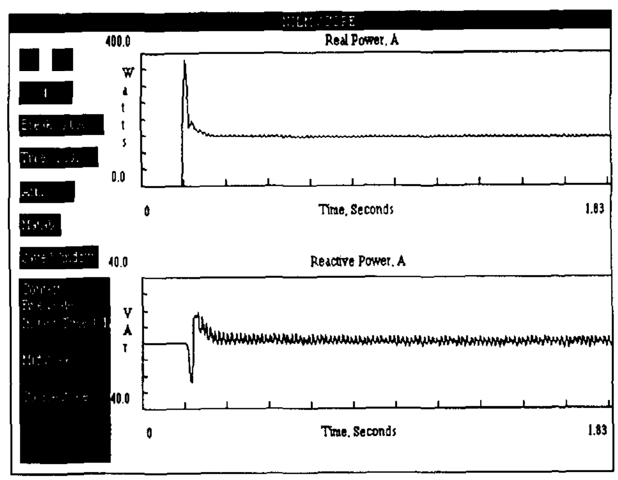

Fig. 12. Start-up transient of instant-start fluorescent lamps, as identified by the NILM.

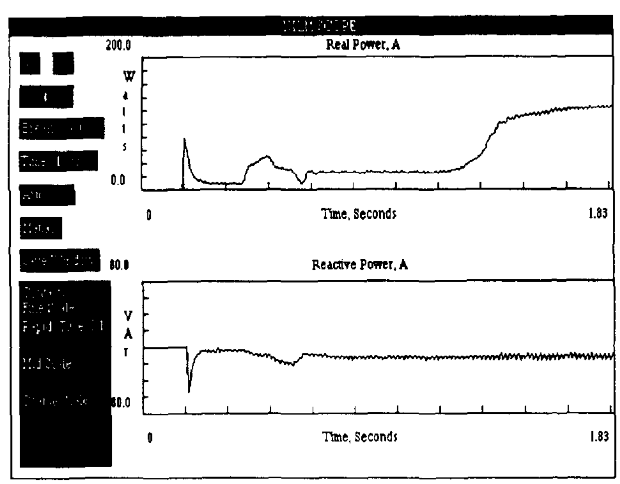

Fig. 13. Start-up transient of rapid-start fluorescent lamps, as identified by the NILM.

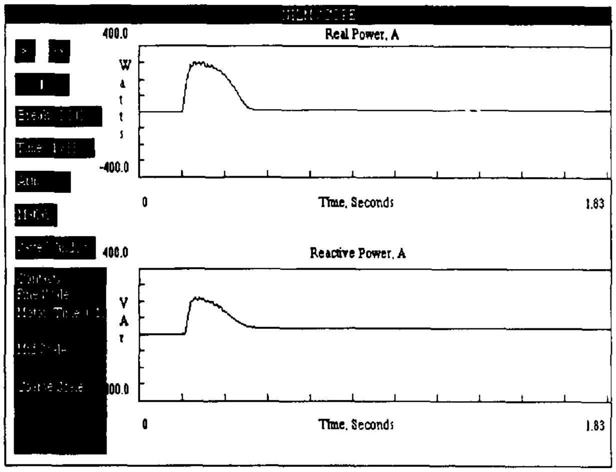

Fig. 14. Start-up transient of small induction motor, as identified by the NILM.

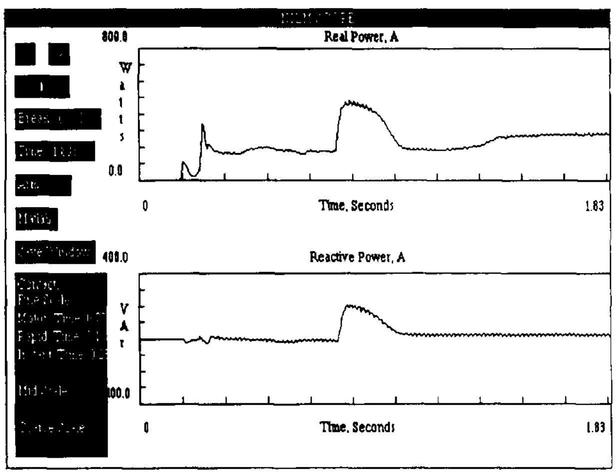

Fig. 15. Overlapping start-up transients of instant- and rapid-start fluorescent lamps and small induction motor, as identified by the NILM.

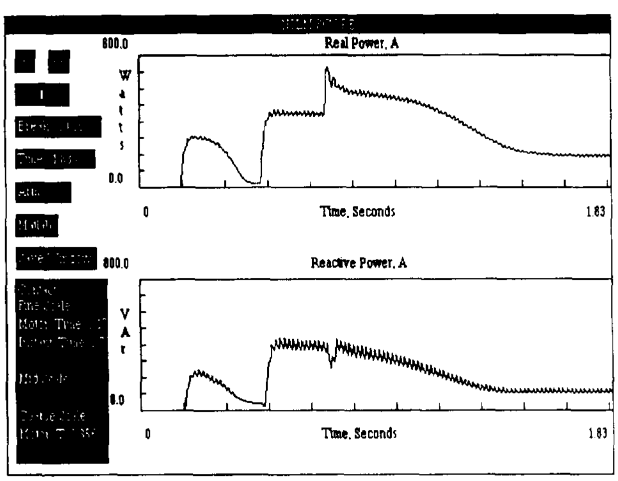

Fig. 16. Overlapping start-up transients of small and large induction motors and instant-start fluorescent lamps, as identified by the NILM.

iments conducted with the test stand. The graph windows in each figure show estimates of the envelopes of real and reactive power on one phase of the three-phase service. The lower left-hand corner of each screen identifies transient events by name, time of occurrence, and fine, medium or coarse time scale. Any event that was identified at the initial, highest sampling rate, i.e. at the first stage in the tree-structured decomposition, will be listed directly under the heading 'fine scale' in the contact window. By design, it is anticipated that events associated with the small motor and both lamp banks will be listed as fine scale events when they appear, assuming that the event detector functions properly. Events found at the next two coder stages in the tree-structured decomposition will be listed under the headings 'mid scale' or 'coarse scale'. For example, any events recognized by the properly working detector which are caused by the turn-on of the large induction motor should appear under the coarse heading.

The three tests shown in Figs. 12–14 record the performance of the prototype when challenged individually with the turn-on events of the instant- and rapid-start lamp banks and the small induction motor, respectively. In each case, the observed event has been properly identified in the contact window. Fig. 15 shows an example where both lamp banks and the small motor turn on so that all three transient events overlap. No key v-sections overlap with each other. All three events are correctly recorded at the finest time scale in the contact window.

In the final experiment, shown in Fig. 16, the small induction motor turns on and completes its transient, followed by the turn-on transient of the large induction motor and the instant-start lamp bank. Again, no key v-sections overlap with each other and all of the events are correctly identified at the appropriate time scales. Recall that the template for the turn-on transients of motors was generated from a single example of the small motor only. Nevertheless, the event detector correctly classified both the small and large motors.

8. Comparison of steady-state and transient load detection

Transient load detection takes more advantage of information available at equipment start-up than steady-state load detection, at the cost of more signal processing. We have shown that it is capable of resolving overlapping start-up events, a major advantage in separating loads when there is no additional identifying information. While building automation systems provide such information they are by no means in universal use and typically do not control most lighting circuits and tenant equipment.

The multiple time scale feature of the transient load detection algorithm, which could be applied as well within a steady-state approach, is an important weapon in discriminating significant events from those considered to be unimportant. We have shown that the steady-state approach was triggered by electrical impulses that we associated with variable-speed motor drives. We were focusing on motor start-up events and in effect had made an a priori judgment that these electrical impulses were not of interest. However, these impulses have a small but non-zero energy associated with them, which might in some cases be considered important. Also, similar impulses could be due to thermostated heaters operating in a variety of types of equipment, including oil-sump heaters in chillers, residential electrical ranges, and heaters in copiers and printers. A steady-state NILM with a sufficiently fast sampling rate could analyze the energy consumption associated with the impulse. However, the faster sampling rate would also make the NILM more likely to react to disturbances that appear as near-steady conditions during a prolonged motor start-up transient; in these cases, the start-up transient would be split by the steady-state NILM into more than one component. The transient detector, by design, operates at multiple time scales, avoiding the dilemma of choosing a single scale. It should be noted that the multiple-time-scale feature of the transient detector could be applied as well to a steady-state algorithm.

The steady-state algorithm triggers on all departures from steady power, clusters changes in power, and later identifies known changes. The transient detector puts the discrimination step first: only known transients are retained for clustering by power level or other analysis technique, while others are ignored. The potential risk of up-front discrimination is that noise could corrupt the start-up transient of a piece of equipment that should be monitored, to the point where a match is not made. In addition, up-front discrimination means the NILM may not discover unanticipated but potentially interesting events. Field tests, planned but yet to be performed, will help determine the effectiveness of the transient detector.

Transient start-up signals provide information not only about equipment identity but also about equipment health. Fault detection on the basis of equipment mechanical and electrical dynamics has been investigated for specific applications, including a centrifugal pump driven by a direct-current motor [16]. The structure of the model of the physical system, the number of model parameters and the amount of electrical noise affect the accuracy to which the parameters can be determined and their sensitivity to faults. Frequency analysis of motor electrical current is also a route to detecting motor bearing and rotor problems [17]. Transient detection, potentially, can provide a basis for identifying the same problems now found via frequency analysis. If further research shows this to be true, the transient detector offers significant added value.

But transient load detection itself is not a complete load monitoring system. The transient detection algorithm is intended to serve as the front end for an overall package that would keep track of device operation and energy consumption. To that end, both start-up and shut-down electrical events must be considered. At start-up, depending on the electrical environment and the goals of the analysis, the NILM could use a steady-state detector, the transient pattern matching algorithm, or a hybrid. The hybrid approach, which we are considering, would use a change-of-mean detector as a front-end for the transient pattern matching algorithm; the change-of-mean detector, somewhat similar in function to that used in the steady-state NILM, would reduce the computational burden on the transient pattern matching algorithm in situations where there were few events by triggering it only when there was a significant change in electrical power. At shut-down, transient signals are typically not available because a meter on the power side of an on–off switch sees only an abrupt decrease in real and reactive power. That is, the shut-down signal is a change in level, with no dynamics, and such signals must be reasonably combined with transient start-up information.

Consider the job of sorting out two devices. There are four cases of interest, based on the device type and magnitude of steady-state real and reactive power. The first and second cases concern devices of dissimilar electrical power which may or may not be of the same type. In these two cases the devices can be distinguished on the basis of magnitude alone. Start-up and shut-down cause step changes equal in magnitude but opposite in sign. No transient start-up information is needed if the start-up signals are well separated in time. Third, if the power magnitudes are the same and the device types are also the same (as would be revealed by the transient detector), it is possible to accurately tally the total energy for the device type. For example, if both are lights, start-ups and shut-downs may be assigned to either device without compromising the calculation of lighting energy. Last, if the devices have similar power magnitudes but are of different type, the start-up transients, which distinguish the devices, cannot be uniquely associated with shut-down level changes and energy calculations cannot be made.

Finally, a problem not completely solved within the framework of either steady-state or transient load detection concerns non-intrusive measurement of electrical power at times between start-up and shut-down for equipment that draws a varying amount of power. In these cases, it is necessary to correlate power with one or more variables that can be measured for each piece of equipment. For example, the power drawn by the motor powering a centrifugal fan or pump depends on the pressure rise across the fan or pump and the flow rate, along with the efficiencies of the device, the motor and the adjustable-speed drive, if present. If values were available for all variables, power could be predicted with precision. Power can also be correlated with a reduced set of variables during a test cycle performed when other loads are steady. We have correlated the power drawn by a centrifugal fan that was driven by a VSD with a single variable, either airflow or the VSD control signal [18]. There was no measure of pressure across the fan or of the more readily obtained pressure read by the set point controller downstream of the fan. We improved the correlation by including the pressure at the set point controller [19]. The first correlation has proved capable of accurately separating the power of two sets of fans but neither has been tested in an environment that includes more equipment.

9. Conclusions

The potential of reduced-cost data acquisition has motivated development of centralized or non-intrusive electrical load monitoring. For residential buildings, previous research based on steady-state detection algorithms [4] has yielded a device that is now being commercially developed. Our work, distinct in its approach and application, has focused on commercial buildings. To date, we have shown that electrical loads from space-conditioning equipment in commercial buildings can be detected centrally on the basis of appropriate filtering and changes in steady-state power. This approach, tested with manual data analysis in a single large institutional building, has subsequently been upgraded and been fully automated [20]. To distinguish equipment with near-simultaneous start-ups or different types of devices with similar real power levels, we have developed prototype hardware and software that analyze start-up transients that reveal the essential physics of the equipment. This prototype transient detection NILM performed well in laboratory tests and an upgraded version will next be tested in the field. Further enhancements include a more efficient parallel processing architecture to permit the NILM to compare the measured electrical signal with a number of patterns for individual devices.

The value of a NILM can be assessed by reviewing a sequence of tasks it will be asked to perform, including simple detection of a start-up event, analysis of that event, measurement of start-up power, detection of power oscillations while equipment is operating, detection of shut-down events, and estimation of energy consumption.

Start-up detection. In commercial buildings, detection of start-up events can answer an important question: did a pump, fan or chiller turn on in response to a control signal? We have shown in a single commercial building that steady-state load detection algorithms are capable of performing this task with a resolution adequate for finding all major loads. A NILM would in principle eliminate the need for current transducers or mechanical sensors at individual devices to perform this function.

Start-up analysis. If the NILM can accurately analyze the start-up transient (a subject of future research) it cannot only identify the type of equipment but also detect deterioration of the device. Such analysis would be enhanced by knowing a priori the identity of the device, as would be the case for major HVAC equipment controlled by an energy management system. Today, chillers are sufficiently expensive and electronics sufficiently powerful that on-line diagnostics are commonly provided; for example, temperature sensors are used to monitor motor bearings. For smaller fan and pump motors, periodic examination of motor health by portable instrumentation can in principle be performed but typically is not, given the expense. The NILM would be able to perform this task if motor diagnostics prove successful.

Quantification of start-up power. This step is more difficult than might be expected, particularly for devices that exhibit prolonged start-up transients. We have shown that steady-state algorithms appear adequate for relatively rapid start-up of a simple induction motor driving a pump but appear less successful for the slower start-up of devices controlled by variable-speed drives. In the latter case, the transient pattern matching algorithm may be of special advantage. Knowledge of start-up power provides a starting point for quantifying device energy consumption. The NILM can also provide an indication of device performance by comparing start-up power for a known piece of equipment against an expected range of powers. Device deterioration and unexpected equipment loads can in principle be detected in this way, depending on the resolution of the NILM and the magnitude of the fault. We have shown that analysis of start-up power can also be used to detect HVAC controller faults associated with switching multiple chillers or pumps.

Detection of power oscillations. One specific result of the work described in this paper was the unexpected detection of a poorly tuned chiller controller that created large power oscillations. While most of the power of the NILM comes from its ability to detect start-up or shut-down events, changes in operating power are also of interest. Another example of this is large changes in power when a fan with a variable-speed drive is boosted from a low-speed condition at night to a higher speed during occupied hours. This ramp transient can appear as a step over a sufficiently long sampling period [20] and can be analyzed as such.

Shut-down detection. Shut-down events are characterized by an abrupt decrease in electrical power and can be detected with steady-state algorithms. This has been amply demonstrated in residential buildings and we have also demonstrated its efficacy in a single institutional building. Detection of shut-down events is of course a necessary step in estimating electrical consumption over an operating period. But it can also be used to detect abrupt electrical or mechanical faults that would cause a piece of equipment to shut down abnormally before an energy management system send out a scheduled shut-down signal.

Calculation of energy consumption. For a device that operates at constant power, detection and matching of start-up and shut-down events, with an associated time for each, is adequate to compute energy use [4]. However, in commercial buildings, many motor-driven pieces of equipment do not operate at constant power; the list includes induction motors subject to variable loads, with or without variable-speed drives. Not only is it more difficult to match start-up and shut-down events, but it is necessary to interpolate the power between these events. The matching is made easy for equipment controlled by an energy management system, for which control signals are available. Power interpolation can be down by simply making a linear fit between start-up and shut-down, by using shut-down power alone in cases where a significant and short-lived start-up load is expected, or by attempting to track the operation of a variable-speed drive via the speed control signal.

Acknowledgements

The authors gratefully acknowledge financial support and technical advice provided by the Electric Power Research Institute and Johnson Controls, the guidance and support of James Kirtley and Richard Tabors, and the assistance of Roger Hill, Kurt Levens, Ruel Little and Nicholas Mabey.

References

[1] M.F. Fels, PRISM: an introduction, *Energy Build.*, *9* (1 and 2) (1986) 5–18.

[2] J. Haberl, D. Bronson, S. Hinchey and D. O'Neal, Graphical tools to help calibrate the DOE-2 simulation program to non-weather dependent measured loads, *ASHRAE J.*, *35* (1) (1993).

[3] J.H. Eto, H. Akbari, R.G. Pratt and S.D. Braithwait, End-use load shape data application, estimation, and collection, *Proc. ACEEE 1990 Summer Study on Energy Efficiency in Buildings*, Vol. 10, American Council for an Energy-Efficient Economy, Washington, DC, 1990, p. 39.

[4] G.W. Hart, Nonintrusive appliance load monitoring, *Proc. IEEE*, *80* (12) (1992) 1870–1891.

[5] Tabors Caramanis and Associates, Nonintrusive load monitoring, *Draft Final Rep.*, Electric Power Research Institute, Palo Alto, CA, 1992.

[6] L. Norford and N. Mabey, Non-intrusive electric load monitoring in commercial buildings, *Proc. Eighth Symp. Improving Building Systems in Hot and Humid Climates, Dallas, TX, 1992.*

[7] W.C. Karl, S.B. Leeb, L.A. Jones, J.L. Kirtley, Jr. and G.C. Verghese, Applications of rank-based filters in power electronics, *IEEE Trans. Power Electron.*, *7* (3) (1992) 437–444.

[8] J.E. Braun, S.A. Klein, J.W. Mitchell and W.A. Beckman, Applications of optimal control to chilled water systems without storage, *ASHRAE Trans.*, *95* (1) (1989) 663–675.

[9] F.L.F. Pape, J.W. Mitchell and W.A. Beckman, Optimal control and fault detection in heating, ventilating, and air-conditioning systems, *ASHRAE Trans.*, *97* (1) (1991) 729–735.

[10] M. Basseville and A. Benveniste, *Detection of Abrupt Changes in Signals and Dynamical Systems*, Springer, Berlin, 1980.

[11] S.B. Leeb, A conjoint pattern recognition approach to nonintrusive load monitoring, *Ph.D. Dissertation*, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, 1993.

[12] S.B. Leeb, S.R. Shaw and J.L. Kirtley, Jr., Transient event detection in spectral envelope estimates for nonintrusive load monitoring, *IEEE Trans. Power Delivery*, *10* (3) (1995) 1200–1210.

[13] P.P. Vaidyanathan, Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial, *Proc. IEEE*, *78* (1) (1990) 56–93.

[14] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. Pattern Anal. Mach. Intelligence*, *11* (7) (1989) 674–693.

[15] Y. Meyer, Orthonormal wavelets, *Wavelets: Time–Frequency Methods and Phase Space, Proc. Int. Conf., Marseille, France, 1987*, Springer, Berlin, pp. 21–37.

[16] R. Isermann, Process fault diagnosis based on dynamic models and parameter estimation methods, in R. Patton, P. Frank and R. Clark (eds.), *Fault Diagnosis in Dynamic Systems; Theory and Applications*. Prentice Hall, New York, 1989.

[17] P.J. Tavner and J. Penman, *Condition Monitoring of Electrical Machines*, Research Studies Press, Exeter, UK, 1987.

[18] L. Norford and R. Little, Fault detection and load monitoring in ventilation systems, *ASHRAE Trans.*, *99* (1) (1993) 590–602.

[19] D.M. Lorenzetti and L.K. Norford, Pressure setpoint control of adjustable speed fans, *ASME J. Solar Energy Eng.*, *116* (Aug.) (1993) 158–163.

[20] R.O. Hill, Applied change of mean detection techniques for HVAC fault detection and diagnosis and power monitoring, *M.Sc. Thesis in Building Technology*, Massachusetts Institute of Technology, Cambridge, MA, 1995.