OPLD: Towards improved non-intrusive office plug load disaggregation

Balaji Kalluri*, Sekhar Kondepudi*, Kua Harn Wei*, Tham Kwok Wai*, Andreas Kamilaris^
*Department of Building, National University of Singapore, Singapore
^Department of Computer Science, University of Cyprus, Cyprus
Email: {balajikalluri, sekhar.kondepudi, bdgkuahw, bdgtkw}@nus.edu.sg, kami@cs.ucy.ac.cy

Abstract*—Practical energy auditing in offices poses several challenges unlike homes e.g. physically large space, diverse energy appliances types, several appliance instances and occupancy non-obstructiveness. However, improved energy-auditing measures using predictive analytics can benefit energy savings and reduce building operational costs. A review of some publicly available energy datasets such as AMPds, BLUED, ECO, REDD etc. is presented to help understand practical limitations in relying and applying them alike for office plug load audits. A possible approach to predict miscellaneous electrical plug loads (MELs) is proposed using Office Plug Load Dataset (OPLD) based on empirical characteristics and measurements of MELs devices. This work in progress study is one of the first attempts to characterize office desktop appliances across multiple states through a very large experimental dataset. The dataset might be effective in identifying individual appliances & its states in aggregate signature. This can find promising application in improving our understanding on office MELs and thus disaggregating them from single-point measurement.

***Keywords*—offfices, MELs, characterization, OPLD, energy audits.**

I. Introduction

Year after year starting from 2011 public energy datasets have been made public in order to progress research by applying information science to solve energy related issues intelligently in the context of buildings, especially homes. Examples of such datasets are REDD (2011) [1], BLUED (2012) [2], AMPds (2013) [3] and ECO (2014) [4] besides few others. The idea that this paper proposes also stems from similar problem but from a different context (i.e. offices), thus the name ***o****ffice* ***p****lug* ***l****oad* ***d****ataset* (OPLD).

Miscellaneous Electrical Loads (MELs) refers to all electrical loads except the conventional heating, cooling and lighting load in a building. It account for more than 20% of primary energy used in commercial buildings, and this load component is projected to increase to nearly 40% in the next 20 years [5]. It was also found that commercial MELs consume more energy than any traditional building load [6]. The evidence of such a trend was also observed in a tropical city e.g. Singapore, with combined office and miscellaneous plug loads accounting for about 25% of the overall building load. It is also apparent that both enterprise and commercial building stock across the globe are on the rise. A recent report from Building Construction Authority of Singapore indicate that the tropical city is no different in this context with office building stock ranks high (~52%) in gross foot area among all building category. Motivated by these opportunities several studies have been conducted by our team within university's enterprise-like buildings. One such study identified several best practices in the use of office *Information and Communication Technologies* (ICT) appliances rationally [7]. While another six-month study in a shared open-plan office setting found occupants and their interaction with personal desktop PCs play significant role in curtailing energy wastes [8]. For results of such studies to be more meaningful similar study has to be carried out across many such offices. Therefore, the notion of scalability of office plug load audits is the core idea of this paper.

The challenge with non-intrusive plug load audit in offices and similar workplace environments has to be treated differently, using different kind of dataset unlike aforementioned ones. This paper proposes and presents one such dataset called *Office Plug Load Dataset (OPLD)*. Unlike homes, offices poses surprisingly different set of challenges such as – physically large space, diverse appliance types, several identical instances, overlapping operational timing and occupancy non-intrusiveness while conducting audits. The incentives for developing OPLD are manifold [7]. It offers scope to develop detailed device-type and instance-type characteristics, identify individual appliance type & their operating state from single-point measurement [10]. This will help to develop appropriate energy curtailment measures to the level of individual appliance states, automate energy savings measures, and improve occupancy engagement through eco-feedback [9].

The paper is organized as follows: Section 2, reviews four commonly used public non-intrusive load monitoring (NILM) datasets and discusses their applicability to office MELs audits. Section 3, introduces OPLD – discuss factors that influenced the experimental design for data collection, actual experimental design and data-preprocessing. It also presents robust approach to calibrating the hardware to maintain data integrity. Section 4, discusses results of individual device characteristics obtained from both aggregate signature and ground truth data for several appliance-type and instance from OPLD. Section 5 presents the direction of this research in improving office plug load audit technique.

II. Review of public energy datasets

Non-intrusive load monitoring is a term used to address methods and techniques to decipher individual appliance energy consumption and its state from a single, aggregate

978-1-5090-0160-6/15/$31.00 ©2015 IEEE

measurement point in a circuit. This section covers the overview of common public energy datasets used in NILM community.

A. REDD

The dataset consists of three levels of data – whole-home/aggregate, circuit and device level/ground truth measured from 10 homes. Each of them is measured at different sampling rates such as 15 kHz, 3 Hz and 1 Hz respectively. The raw data prepared consists of both high-frequency (order of kHz) and low-frequency (order of few Hz) voltage and current samples from which active and reactive power metrics are computed per second. A sample description on range of appliances measured in a house includes kitchen appliances, washers, lighting, refrigeration, electronics etc. While the approach to REDD data collection covers all possible levels of data for applying either supervised or un-supervised NILM techniques, the dataset has no clear description of electronic appliances measured neither in terms of appliance-level nor their states. Besides, when combined with multiple classes of appliances as in this case, the energy signature of office MELs will be masked by other high-power appliances in the home. Therefore, relying on such datasets particularly for the case of office plug load NILM research will be both challenging & misleading.

B. BLUED

Unlike REDD which is non-event based this fully labeled dataset can be particularly useful in application of event-based methods to energy disaggregation in homes. Similar to REDD, the data collection method in BLUED also follows multiple levels. At the top, entire house energy is collected at 12 kHz, while the ground-truth level data is collected by combining individual appliance meters and ambient environmental sensors such as light, sound, motion, vibration, humidity and pressure. However, the data from intermediate level using circuit-panel meters are collected in both datasets. The fundamental difference between the two lies in the fact that data for each appliance and their binary state-transitions are labeled. The list of appliance events (either *On* and *Off*) were created by data fusion from ground truth sensors, visual inspection of energy signature and hand-labeling transitions in appliance activity whenever power levels changed by 30Watts and last for 5 seconds. From the summary of appliances and its average power consumption along with their event count certain key observations can be drawn.

Firstly, there are quite a number of office plug load appliances whose appliance classes and steady-state characteristics are defined clearly.

Secondly, average power consumption of the appliances is computed based on a 5 second '*On'* event window. While this approach might appeal acceptable for binary appliance states but doesn't scale well especially when characterizing different appliance types across multiple states. This is because observations in OPLD indicate that certain ICT appliance states/events might actually last for more than 5 seconds and their variations over the time window are also considerably higher. This suggests that using a fixed time-window based average power as metric to characterize different office ICT appliances is ambiguous.

Finally, characterizing appliance deeply to the level of multiple states (e.g. computer monitor in different brightness levels, computer is sleep or hibernate and a multi-function printer in scan or print or copy) seems not quite possible with BLUED. Appliance characterization to such multiple-levels of operation can offer deeper insights to domain experts, facility managers and thus create informed choices about office appliance usage to building occupants.

C. AMPds

This dataset is one of such kind which combines data from three different energy resources such as water, natural gas and electricity. The notion of this integration is to aid disaggregate appliances that work on multiple fuel types. It consists of electricity readings from single house measuring 21 loads round the year at one sample per minute rate. A quick look at the range of appliances measured indicates no real presence of office equipment. However, a general sensible finding from this study suggest that current (in Amps) is a better metric in appliance recognition than real power (in Watts) based on variations in data points observed proved by disaggregation results. Therefore this dataset is also inapplicable for office NILM scenarios.

D. ECO

This dataset consists of electrical energy consumption of various home appliances along with occupancy information. Similar to other datasets it consists of data measured considerably for long duration (8 months) in 6 different houses. The real and apparent power for both aggregate and ground truth individual appliance data is sampled at 1Hz using smart meters and smart plugs respectively. Every house in the dataset contains about 7 to 8 appliances out of which either 1 or 2 are typically office electronic MELs.

A simple previous application of Weiss's algorithm on ECO which classifies individual appliances from aggregate signature using events (i.e. changes in real and apparent power) based on clustering technique indicate poor results in the case of laptops. This is because of the nature of measurement of loads in the ECO which poses the possibility of confusing switching events of laptops with variations in other appliances.

The following summarizes some learning from all the above public MELs datasets.

- The inter-event time in BLUED suggested that for higher (individual) appliance event detection rates from aggregate signature, datasets must be prepared at less than 1-minute resolution.
- The frequent overlaps in appliance events observed in both BLUED & ECO datasets degrades the individual appliance event detection rates. This in conjunction with higher event count observed in the case of office electronic loads makes disaggregation in offices challenging.

- The choice of current (Irms) as metric in appliance recognition is suggested through AMPds.

A common observation from all the above is that there is insignificant range of office MELs, lack of aggregate office plug load circuit measurements and their distinct power characteristics to enable disaggregation of plug loads in offices which is the focus of OPLD.

III. OPLD

Office Plug Load Dataset abbreviated as *OPLD* as the name suggests is

- A repository for energy signatures of several common office workstation plug loads. Currently, it encompasses four appliance types such as desktop & laptop PC, monitor and imaging devices.
- Characterized by low-frequency measurements of both individual appliance and aggregate circuit using metrics such as real power (P), apparent power (S) and current (I).
- A laboratory based experimental measurements and approximates real usage pattern of appliances in offices.
- First of this kind to aid office appliance characterization to several power states thus enabling deeper non-intrusive disaggregation.
- Statistically robust dataset unlike others with repeated measurements to aid better characterizing appliances and build models.

The experimental design for building OPLD is performed considering potential gaps and lessons learnt from previously available datasets for office plug load audits.

- *Data resolution:* The data collection both individual appliance and aggregate in OPLD is carried at 1Hz. Such low frequencies are characteristic of commercially available smart meters at floor/building level.
- *Power metrics:* OPLD offers real/active power (in Watts), apparent power (in Volt-Ampere) and current (in Amps) measured simultaneously both at individual appliance level and aggregate circuit level aiding both unsupervised and supervised NILM.
- *Overlaps in aggregate signature* due to multiple appliance events are handled in OPLD by two means. First, no two or more similar appliances are present in single (aggregate) circuit to aid characterizing, recognizing and disaggregating individual loads in aggregate signatures truthfully. Second, data collection is performed by controlled experimental setup which ensures no two or more dissimilar[1] appliance events occur concurrently (details of which is elaborated in next section). It is arguable that such an approach is not scalable to entire office plug load audit. But our experiences suggest that the proposed approach is promising in scaling the current state-of-art practices for office plug load audits atleast to next level. This would be better understood through illustration from section III-A.

[1] Dissimilarity in MELs context is treated based on functional usage. For example, a MFD which does print, scan, copy job is very different in usage to a PC.

A. Office NILM

The idea of OPLD can be better understood through an illustration of non-intrusive load monitoring in offices. A typical enterprise office workplace often has several workstations whose electrical network is designed with a separate plug load circuit. Such electrical sub-circuit typical branches from breaker panel either at floor-level or individual office rooms as shown by red line in *Fig. 1*. In such a context, auditing individual appliance on every workstation non-intrusively becomes expensive and time-consuming. One of the most common problems for applying NILM to offices is the presence of several identical appliance types in one (plug-load) circuit indicated by pink line in *Fig. 1*. To alleviate the complexity in such context, the proposed heuristic approach approximates every office workstation into a (dis)aggregating point (labeled A in *Fig. 1*) with three or four common office appliances. Thus the problem reduces the number of metering points, say for example 75 to 25 sockets.

The illustrated experimental design for data collection of OPLD thus becomes meaningful and approximates reality. It is carefully designed considering several commonly used office appliance types and scenarios, making individual appliance recognition/disaggregation from aggregate signature in offices possible.

Currently OPLD is designed to address common office plug load use case configurations with four appliance types - desktop PC, laptop, multi-functional printing device (MFD), and monitor. These appliances permit five different combinations of their aggregate usages in a typical office workstation.

1) 2 appliances in a circuit
    a) desktop PC + monitor
    b) laptop + auxiliary monitor
    c) laptop + MFD

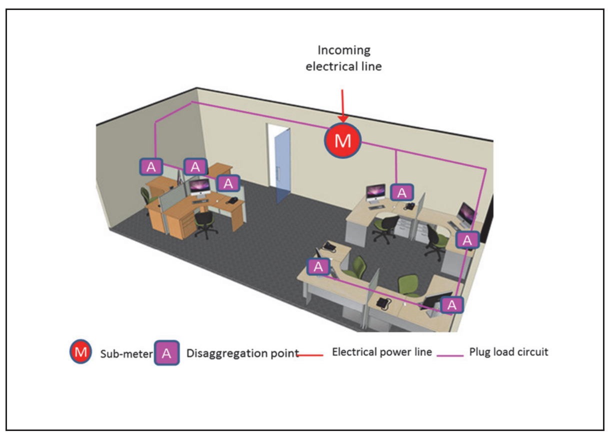

Fig. 1. Conceptual schematic of proposed approach to office plug load NILM

2) 3 appliances in a circuit

a) desktop PC + monitor + MFD

b) laptop PC + auxiliary monitor + MFD

Table I shows operating modes/states for various office MELs within OPLD. For instance, when a desktop PC, monitor and MFD are operating in tandem in a circuit as shown in *fig. 2*, there are 19 distinct use-case scenarios possible which can be referred to as *cases*. The table II shows the list of such *cases* in a circuit with three such appliances. The discussion about its characteristics in presented in section IV. Although there are 100 levels of brightness possible in a monitor, only 3 discrete levels are considered. Similarly a desktop or laptop can have multiple operating modes depending on the CPU workload, memory, IO disk access, network load etc. however in the context of offices it could be subjected to one of the three most common modes. To simulate the workload of (both desktop & & laptop) PCs in offices some common office applications are made to run in *On/Idle* mode (e.g. Microsoft Office Word, Excel, web-browsing, Adobe PDF reader and a calculator). However, in case of an MFD there are five distinct operating modes which is often used in offices. Our previous study [7, 8] has indicated potential savings in identifying them and switching between them either when not in use or during times to alternatively suggest users with other alternative eco-feedback actions.

All the measurements in OPLD both aggregate circuit and individual appliances are made at 1Hz. All possible cases are measured for duration of approximately 10 minutes each. This means for any given case (for example, desktop in *Idle* + monitor in *100% brightness* + MFD in *copy)* approximately 1500 to 1800 legitimate steady-state data points are possible to quantify appliance states in aggregate signature.

Measurements to this level of detail becomes especially useful in characterizing individual appliances, their steady & non-steady state operation and disaggregating them without having to rely on existing sparse datasets for office NILM. It is important to note that OPLD is built using measurements

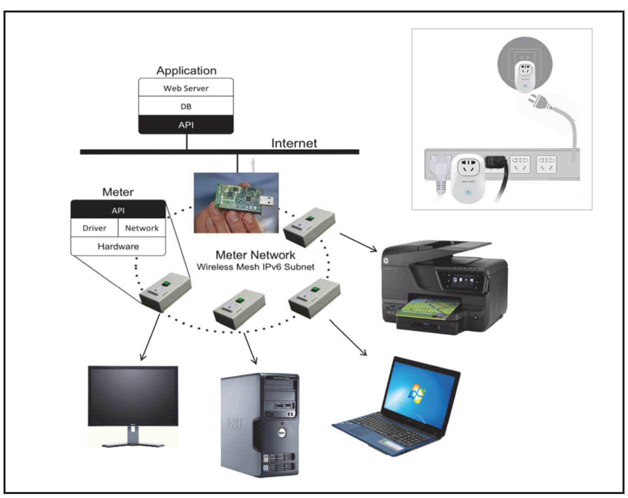

Fig. 2. Illustrative use case of OPLD in office NILM

TABLE I. DEVICE TYPE & THEIR POWER STATES CONSIDERED IN OPLD

Appliance type	Operating states
Desktop (or) laptop PC	*On/Idle* *Sleep* *Off*
Monitor	*100% brightness* *50% brightness* *Off*
MFD	*Idle* *Copy* *Scan* *Print* *Off*

-from three to five distinct models for each device type considered. This ensures a robust data set. Besides all possible permutations of device combinations are also considered in this measurement. Each participating device in every experiment is subjected to all valid appliances states manually. This accounts for over 3 million data points for aggregate and ground-truth appliance signature within the data set. Such an exhaustive dataset facilitates the notion of deep disaggregation which allows appliances to be treated as multi-state devices (e.g. 50% brightness, PC-sleep, MFD-scan) during NILM. This is in contrast to several existing public datasets that limit appliance states to binary on/off.

B. *Data sanctity measures*

Good data is key to characterizing MELs, especially the low-powered office appliances. The calibration as a data sanctity measure was considered apriori to OPLD data collection due to two reasons. The ACme smart plugs being used were "research grade" devices and not commercial grade and therefore not factory calibrated unlike several other metering devices used in building other datasets. The data output from smart plug is dimensionless and requires means to convert to meaningful quantities such as Watts, Volt-amperes, amperes etc. prior to actual data collection. A simple yet rigorous calibration procedure was thus carried out. In order to avoid any possible overlaps in appliance energy signatures at low power range due to measurement & instrumental errors. It was observed that calibration yielded a perfect straight line fit (R^2~0.99) and every smart plug mote had its own regression coefficients for real power (W), apparent power (VA) and current (A).

TABLE II. POSSIBLE APPLIANCE STATE COMBINATION WITH 3 APPLIANCES IN A CIRCUIT (WHERE D - DESKTOP PC, M – MONITOR, P – MULTI-FUNCTIONAL DEVICE)

D-Idle, M-100, P-Idle	*D-Idle, M-Off, P-Copy*
D-Idle, M-100, P-Copy	*D-Sleep, M-Off, P-Idle*
D-Idle, M-100, P-Scan	*D-Sleep, M-Off, P-Copy*
D-Idle, M-100, P-Print	*D-Idle, M-100, P-Off*
D-Idle, M-50, P-Idle	*D-Idle, M-50, P-Off*
D-Idle, M-50, P-Copy	*D-Sleep, M-Off, P-Off*
D-Idle, M-50, P-Scan	*D-Idle, M-Off, P-Off*
D-Idle, M-50, P-Print	*D-Off, M-Off, P-Idle*
D-Idle, M-Off, P-Idle	*D-Off, M-Off, P-Copy*
	D-Off, M-Off, P-Off

C. Data preparation and pre-processing

The raw energy signatures from both aggregate circuit and individual appliances at the end of every measurement cycle are exported from an application developed for this purpose and stored into local memory in the form of CSV files. This raw data has to be cleaned first, to account for duplicate & missing packets. Other pre-processing steps involve filtering the windows of time corresponding to different cases and combining the likely events (e.g. all monitor 100%, MFD Idle etc.) and extract meaningful steady-state statistics from the dataset. More details of analysis and findings from OPLD are discussed in section IV. However, preparing datasets like OPLD requires manually labeling the events and filtering time-slices systematically which is time-consuming and challenging. This process is semi-automated in this case using python scripts developed for this purpose. Since there is no real novelty employed in preprocessing and preparation steps, no further details are included in the scope of this paper.

IV. RESULTS AND DISCUSSION

While the data collection of OPLD that we are generating is an ongoing process, some rather interesting preliminary observations have been made from aggregate and individual appliance data. For the purposes of this discussion on some steady-state characteristics of appliances, a case of 3 appliance combination in a circuit (e.g. desktop, monitor and MFD) from OPLD is considered.

Both table III & IV summarize characteristics of 3 different instances of desktop PC (D), monitor (M) and MFD (P) from OPLD with suffices, column labeled *#data points* indicate the no. of distinct measurements used from OPLD in drawing statistical inferences of respective appliance steady states and finally notation P_{avg}, P_{SD}, S_{avg}, S_{SD}, I_{avg}, I_{SD} are used to represent statistical mean and standard deviation of active, apparent power and rms-current respectively based on the number of data points.

Table III presents the empirical characteristics of 3 different MFD instances in two steady-states from its ground-truth data. The following observations can be drawn.

- Steady state power consumption of three different MFD types (P1, P2 & P3) show difference in their active power, apparent power & current.
- Within the printer instance type both the *Idle* & *Off* state power consumption metrics remain consistent (i.e. row 1 to 6 & row 7 to 12) no matter whatever be the aggregate devices in the circuit, indicated by low standard deviation.
- While printer 1 draws substantial power approx. 8W in *Off* state, other two printer instances do not exhibit the same trend.

This indicates that there is variability within device class and they can't be treated alike. Table III presents only two steady-states of MFD leaving other states such as copy, print and scan which show non-steady characteristics. Hence these appliance states are to be treated differently.

TABLE III. STEADY-STATE CHARACTERISTICS OF 3 PRINTER INSTANCES FROM GROUND-TRUTH DATA

Aggregate appliances	Characterizing appliance/ steady-state	# data points	Power quality metrics					
			Active Power (in W)		Apparent Power (in VA)		Current (in A)	
			P_{avg}	P_{SD}	S_{avg}	S_{SD}	I_{avg}	I_{SD}
D1 + **M1** + P1	P1/Idle	8742	13.466	0.518	33.975	1.086	0.140	0.004
	P1/Off	8159	8.708	0.365	24.705	0.447	0.102	0.002
D1 + **M2** + P1	P1/Idle	8769	13.458	0.522	34.540	0.956	0.142	0.004
	P1/Off	8869	8.744	0.382	25.276	0.404	0.104	0.001
D1 + **M3** + P1	P1/Idle	8710	13.665	0.636	34.930	1.138	0.144	0.005
	P1/Off	8843	8.852	0.315	25.430	0.543	0.105	0.002
D1 + M2 + P2	P2/Idle	8256	3.135	0.567	15.062	0.884	0.062	0.004
	P2/Off	8026	0.016	0.318	9.675	0.653	0.039	0.003
D2 + M2 + P2	P2/Idle	6980	3.506	0.633	15.515	1.642	0.064	0.007
	P2/Off	7281	0.000	0.001	9.117	0.181	0.038	0.001
D3 + M2 + P2	P2/Idle	5388	3.447	0.564	15.468	1.255	0.064	0.005
	P2/Off	8008	0.000	0.001	9.335	0.171	0.038	0.001
D3 + M3 + P3	P3/Idle	8450	0.122	0.297	9.204	0.589	0.038	0.002
	P3/Off	8631	0.000	0.003	8.828	0.241	0.036	0.001

Similarly following observations can be drawn from table IV for monitors.

- Similar to MFD (presented in table III) steady state power consumption of different monitor types (M1, M2 & M3) also show difference in their active power, apparent power & current
- While monitor 2 & 3 exhibit approximate 50% reduction in power consumption moving between 100% to 50% or 50% to 0% as expected, but monitor 1 doesn't reflect similar behavior. This was found to be rather interesting.
- Within the device instance type all monitors' power metrics remain quite consistent indicated by low standard deviation.

Table V presents the summary of energy metric (e.g. active power alone) of how the individual appliances in the circuit influence the mean power in the aggregate signature through an example case of D1+M1+P1 across all legitimate *cases*. The following observations can be drawn from table V.

- It can be seen that both apparent power and current doesn't add up to form aggregate in contrast to observation made in AMPds. Hence it has been dropped for further analysis.

TABLE IV. STEADY-STATE CHARACTERISTICS OF 3 MONITOR INSTANCES FROM GROUND-TRUTH DATA

Aggregate appliances	Characterizing appliance/ steady-state	# data points	Power quality metrics					
			Active Power (in W)		Apparent Power (in VA)		Current (in A)	
			P_{avg}	P_{SD}	S_{avg}	S_{SD}	I_{avg}	I_{SD}
D1 + M1 + **P1**	M1/100%	8764	19.740	0.374	47.167	1.064	0.192	0.003
	M1/50%	8758	14.299	0.202	36.219	0.678	0.148	0.002
	M1/Off	15176	0.000	0.000	11.505	0.250	0.047	0.001
D1 + M1 + **P2**	M1/100%	8813	20.999	0.468	47.872	1.339	0.196	0.005
	M1/50%	8837	15.086	0.354	36.918	1.023	0.151	0.004
	M1/Off	15638	0.001	0.010	11.637	0.270	0.048	0.001
D2 + M2 + P2	M2/100%	7954	14.428	0.560	37.324	1.122	0.152	0.004
	M2/50%	8511	7.676	0.456	23.151	0.917	0.095	0.003
	M2/Off	12135	0.041	0.224	9.744	0.221	0.040	0.001
D3 + M2 + P2	M2/100%	8135	14.253	0.283	36.050	0.673	0.147	0.002
	M2/50%	8123	7.488	0.275	22.411	0.629	0.092	0.002
	M2/Off	12328	0.000	0.003	9.549	0.212	0.039	0.001
D3 + M3 + P3	M3/100%	8608	15.003	0.277	40.321	1.413	0.164	0.005
	M3/50%	8662	8.760	0.262	27.279	0.837	0.111	0.003
	M3/Off	15460	0.000	0.000	11.475	0.272	0.047	0.001

TABLE V. AN EXAMPLE OF STEADY-STATE STATISTICS OF BOTH AGGREGATE AND INDIVIDUAL APPLIANCES USING D1+M1+P1

Case #	Appliance state description	Active Power, P (in W)							
		Aggregate		Desktop		Monitor		MFD	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	*D-Idle + M-100 + P-Idle*	103.191	13.086	68.585	13.522	19.972	0.299	12.776	0.217
2	*D-Idle + M-100 + P-Copy*	109.214	11.509	67.248	10.744	19.660	0.321	20.423	4.959
3	*D-Idle + M-100 + P-Scan*	107.668	9.206	70.997	9.584	19.536	0.285	15.351	0.710
4	*D-Idle + M-100 + P-Print*	97.646	7.433	58.986	6.289	19.481	0.266	16.886	4.084
5	*D-Idle + M-50 + P-Idle*	94.861	7.767	65.215	8.072	14.332	0.197	13.376	0.152
6	*D-Idle + M-50 + P-Copy*	96.632	6.824	60.329	5.536	14.291	0.204	19.915	4.892
7	*D-Idle + M-50 + P-Scan*	94.735	8.859	63.113	9.173	14.262	0.228	15.351	0.680
8	*D-Idle + M-50 + P-Print*	95.982	8.451	62.864	7.576	14.283	0.186	16.813	4.031
9	*D-Idle + M-Off + P-Idle*	74.446	5.722	60.444	6.074	0.000	0.001	13.467	0.143
10	*D-Idle + M-Off + P-Copy*	81.655	7.400	61.101	6.161	0.000	0.000	20.148	4.698
11	*D-Sleep + M-Off + P-Idle*	17.983	0.276	1.352	0.141	0.000	0.000	14.045	0.156
12	*D-Sleep + M-Off + P-Copy*	24.489	4.583	1.341	0.139	0.000	0.000	20.815	4.803
13	*D-Idle + M-100 + P-Off*	93.758	7.896	63.359	8.025	20.058	0.290	8.347	0.145
14	*D-Idle + M-50 + P-Off*	85.408	6.100	60.648	6.450	14.328	0.183	8.437	0.149
15	*D-Sleep + M-Off + P-Off*	13.267	0.198	1.326	0.138	0.000	0.000	9.159	0.120
16	*D-Idle + M-Off + P-Off*	68.193	6.983	58.851	7.389	0.000	0.000	8.610	0.159
17	*D-Off + M-Off + P-Idle*	17.199	0.650	0.933	0.141	0.000	0.000	13.673	0.608
18	*D-Off + M-Off + P-Copy*	24.367	4.524	0.946	0.148	0.000	0.000	21.153	4.716
19	*D-Off + M-Off + P-Off*	12.818	0.186	0.921	0.144	0.000	0.000	9.133	0.461

- In most cases it can be seen that large variation in mean aggregate power is attributed to the corresponding large variability in desktop power consumption alone.
- Aggregate power measured in the circuit is always offset from aggregate of individual appliance load. This is because of the inherent load offered by the strip.
- As hinted before significant variations in mean power consumption of MFD can be observed in *cases* 2,4,6,8,10,12 & 18 which are characteristics of MFD in one of the non-steady states i.e. copy or print. This confirms that such appliance state combinations when modeled using these statistics could be inappropriate.

V. SUMMARY AND FUTURE DIRECTION

Our intent of this paper was to introduce the concept of an OPLD and what may constitute such a dataset. We have examined the possibility using existing public MELs datasets for office NILM, but found that the datasets have some lacunae especially when applied to office appliance. Hence the motivation for the development of an OPLD. The results presented in this paper give an indication to develop deeper understanding into office plug load appliances through simple characterization and aids clarifying few general misconceptions about them. It also opens-up a gamut of opportunities of applying such a dataset for detailed office MELs auditing.

The immediate next step is to understand the characteristics of other non-steady states and deal them differently in aggregate signature. A possible direction is to consider aggregate appliance signature as time-series subsequence, mine motifs in the periodic samples of data from OPLD and match the motifs with ground-truth for appliance state identification through pattern recognition truthfully. Till data very few studies have applied several variants of time-series subsequence based approach to mining aggregate signature for appliance state recognition which have found remarkable success in signature recognition in other time-series data mining problems.

To our best knowledge OPLD is the only dataset till date that has been developed with a specific aim to study office appliances for plug load disaggregation. Currently, our OPLD dataset is not publicly available (unlike REDD, AMPds) but our intent is to publish this dataset after we complete further analysis. When this is done it can spur greater interest among the NILM community to apply existing disaggregation algorithms and techniques which was previously found effective only on residential datasets. Thus it can steer future algorithm development to reflect on contrasting application scenarios based on availability of data.

ACKNOWLEDGMENTS

We would like to thank Jyothsna Devi in volunteering to help conduct lab measurements in building OPLD. Thanks to Terrence Tan & Prof. Shen Lijun in volunteering with their MELs. And our final thanks to the funding agency Singapore's Ministry of Education under AcRF grant R-296-000-145-112.

REFERENCES

[1] J. Z. Kolter and M. J. Johnson, "REDD: A public data set for energy disaggregation research," in Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA, 2011, pp. 59-62.

[2] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe, and M. Berges, "BLUED: A fully labeled public dataset for event-based non-intrusive load monitoring research," in Proceedings of the 2nd KDD workshop on data mining applications in sustainability (SustKDD), 2012, pp. 1-5.

[3] S. Makonin, F. Popowich, L. Bartram, B. Gill, and I. V. Bajic, "AMPds: A public dataset for load disaggregation and eco-feedback research," in Electrical Power & Energy Conference (EPEC), 2013 IEEE, 2013, pp. 1-6.

[4] C. Beckel, W. Kleiminger, R. Cicchetti, T. Staake, and S. Santini, "The eco data set and the performance of non-intrusive load monitoring algorithms," in Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings, 2014, pp. 80-89.

[5] G. Ghatikar, "Miscellaneous and Electronic Loads Energy Efficiency Opportunities for Commercial Buildings: A Collaborative Study by the United States and India," ed, 2014.

[6] A. E. Outlook, "Energy Information Administration," 2015.

[7] A. Kamilaris, D. T. H. Ngan, A. Pantazaras, B. Kalluri, S. Kondepudi, and T. K. Wai, "Good practices in the use of ICT equipment for electricity savings at a university campus," in Green Computing Conference (IGCC), 2014 International, 2014, pp. 1-11.

[8] A. Kamilaris, J. Neovino, S. Kondepudi, and B. Kalluri, "A case study on the individual energy use of personal computers in an office setting and assessment of various feedback types towards energy savings," Energy and Buildings, 2015.

[9] A. Kamilaris, B. Kalluri, S. Kondepudi, and T. K. Wai, "A literature survey on measuring energy usage for miscellaneous electric loads in offices and commercial buildings," Renewable and Sustainable Energy Reviews, vol. 34, pp. 536-550, 2014.

[10] C. A. Webber, J. A. Roberson, R. E. Brown, C. T. Payne, B. Nordman, and J. G. Koomey, "Field surveys of office equipment operating patterns," *Lawrence Berkeley National Laboratory,* 2001.