A PID Inspired Feature Extraction for HVAC Terminal Units

Maitreyee Dey*, Manik Gupta*, Soumya Prakash Rana*, Mikdam Turkey*, Sandra Dudley*
*School of Engineering,
London South Bank University, London, United Kingdom

***Abstract*—Retrofitting older buildings and embedding new building stock with Energy Management Systems (BEMS) is paving the way for smarter energy use and increased well-being awareness and initiatives for occupants. BEMS can discover problems related to energy wastage, user comfort and building maintenance. Remote analysis and categorization of the different Heating, Ventilation and Air-Conditioning (HVAC) Terminal Unit (TU) behaviours based on a unique set of features using BEMS data is the main aim of the proposed work. Hence, a novel feature extraction method inspired by the Proportional Integral Derivative (PID) controller response curve to define events from TU data is proposed and applied to multidimensional, real-time data streams remotely retrieved from a building based in the city of London. The feature extraction method executing across different TUs and the feature sets obtained, have been used to identify different TU behaviour patterns. Subsequently, unsupervised machine learning has been employed to investigate automated TU fault detection and diagnosis.**

***Index Terms*—Heating, Ventilation and Air-Conditioning (HVAC), Building Energy Management System (BEMS), Terminal Unit (TU), Feature Extraction, Proportional Integral Derivative (PID), Fault Detection and Diagnosis (FDD)**

I. Introduction

HVAC systems are the key enabling technology deployed in residential, commercial, and industrial buildings to provide thermal comfort and indoor air quality. The focus of current research is on Terminal Units, a specific subcomponent of HVAC systems, which is responsible for the final delivery of comfort inside built environments. A TU is a common and simple device consisting of a heating and/or cooling heat exchanger or 'coil' and fan used to control the temperature of a single room. It is generally ceiling-mounted and usually controllable by local thermostats in order to control the throughput of water to the heat exchanger using a control valve and/or fan speed. They may either primarily recirculate internal air, or can introduce proportion of 'fresh' air with the re-circulated air. Usually inside buildings, there is a central chiller plant that distributes cold water to all the cooling coils, and a central boiler that pumps out hot water to all the heating coils. The fan is operational at all the times. When the environment becomes too warm, the thermostat senses and signals the chilled water valve, and cold water is subsequently passed through the coil, thus extracting the heat from the air being blown by the fan. If it gets too cold depending on the local set point, the heating coil begins working in the same way. Fig. 1 shows the schematic of a typical TU used in HVAC systems. Poorly controlled or faulty terminal units like fan coil units can be responsible for significant energy wastage and user discomfort in buildings [1]. For example, a faulty fan coil unit can signal a false heating demand to the boiler, causing the boiler and ancillary equipment to activate and begin distributing hot water, unnecessary overheating other room spaces.

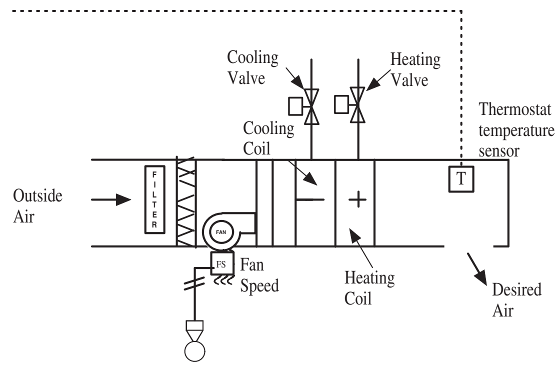


Fig. 1. TU schematic diagram

In order to improve energy efficiency, buildings are now being retrofitted with Building Energy Management Systems (BEMS) that can be used to extract valuable building data. BEMS data can be further analysed to discover faulty terminal units (TUs) and identify their associated problems. There has been tremendous research in FDD for different HVAC subsystems like chillers, air handling units etc. in terms of model based, rule based and data-driven based approaches [2]–[4]. The primary goal of this research is to develop novel data driven approaches harnessing the hidden information buried in the vast amount of historical building data in order to understand different TU behaviours.

Remote observation and fault identification can be an excellent contribution to BEMS in terms of creating future augmentable systems for remote problem solving. This approach effectively leads to empowerment of the building managers who are not always experienced in the systems within the buildings they manage. Remote problem identification and tracking can not only lead to energy wastage reduction, but also provides further benefits like increased operational cost savings and a greater appreciation and understanding of human-building interaction for future smart city applications.

Section 2 gives more details about the proposed feature extraction. Section 3 provides explanations about the different

TU behaviours and clustering results based on the TU data. Section 4 draws conclusions on the paper while stating the future research directions.

II. Proposed Feature Extraction Method

Some of the potential problems that can result in faulty TU behaviours can be poor control due to narrow dead bands settings and / or over aggressive proportional, integral and differential (PID) control, poor sensor location, varying setpoints, out of hours operation, incorrect TU sizing for actual demand, TUs unable to receive adequate flow or upstream temperatures, stuck-open valve, competition from nearby TUs leading to simultaneous space heating and cooling, localised effects due to high solar gains or a TU being placed very close to high energy consuming equipment, and unachievable set point (e.g. 17 degree Celsius ambient temperature requests in a room that has continuous sunshine exposure).

Thus, it's obvious that there can still be a multitude of issues leading to faulty TU behaviour that require expert building engineer knowledge to identify each one of these issues as they occur or identify that malfunctions observed are the summation of individual faults. The manual TU data investigation can be extremely tedious and impossible with the ever-increasing amount of building data exhibiting big data characteristics. Adding human factors into this places even more pressure on managers to identify faults or mishandling of items by occupants. Hence to be more effective intelligent and automated approaches employing data mining and machine learning techniques are required to identify possible TU issues.

This work proposes a novel, data-driven feature extraction method and subsequent unsupervised clustering to identify different TU behaviours to assist both expert and non-expert building managers. The event discovery is inspired by the Proportional-Integral-Derivative (PID) controller response curve [5]. PID controller provides a continuous variation of output within a control loop feedback mechanism to accurately control the process, remove oscillations and increase efficiency. Fig. 2 shows a typical step response curve after a controller responds to a set point change. In Fig. 2, settling

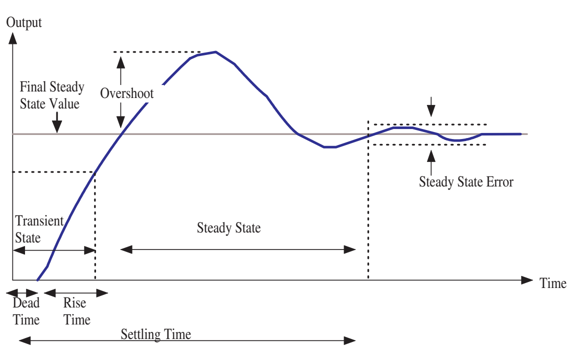
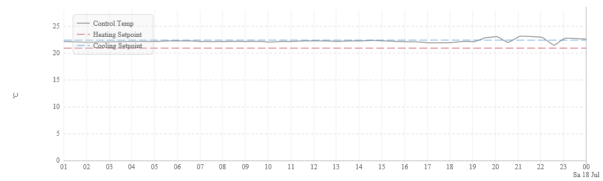
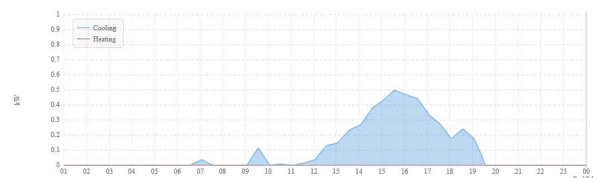


Fig. 2. PID controller response curve

time is the time required for the process variable to settle to within a certain percentage (commonly 5%) of the final steady state value. Steady State Error is the final difference between the process variable and set point. Percent Overshoot is the amount that the process variable overshoots the final value, expressed as a percentage of the final value. While, the curve rises from 10% to 90% of final steady state value within a period known as the rise time, dead time is a delay between when a process variable changes, and when that change can be observed. For instance, if a temperature sensor is placed far away from a cold-water fluid inlet valve, it will not measure a change in temperature immediately if the valve is opened or closed. A system or an output actuator that is slow to respond to the control command, for instance, a valve that is slow to open or close, can cause dead time. A common source of dead time in HVAC systems is the delay caused by the flow of fluids through pipes.



(a) Control temperature and set points



(b) Cooling effort

Fig. 3. Various data streams obtained from a TU

Based on the response curve shown in Fig. 2, the pre-processed data streams retrieved from an individual TU is divided into different events in an event discovery stage. The various data streams obtained for a terminal unit as shown in Fig. 3 are as follows:

1) Control temperature [$°C$] (reported space temperature)
2) Set point temperature and dead band [$°C$]
3) Effort exerted by heating or cooling valve (Average power is calculated from valve demand [%]. A nominal rated power of 1 kW has been assumed for all TUs and provides an estimate of the energy consumption of each TU)
4) Enabled signal to indicate the operational hours

Fig 3(b) shows the heating effort exerted by a TU enabled within the operational hours from 6:00am to 18:00pm to maintain the control temperature in Fig. 3(a) within the desired cooling and heating set point settings.

On a given day, when the heating and cooling units inside a building are enabled during the operational hours, the temperature begins to change depending upon the environmental demand. Now depending on the temperature variations, the data stream is sliced into different time periods and used to designate different events. Whenever the temperature values change with respect to the set point value, an event is considered to happen. The four different types of events are identified in the event discovery stage for both heating and cooling activity as shown in Fig. 4.

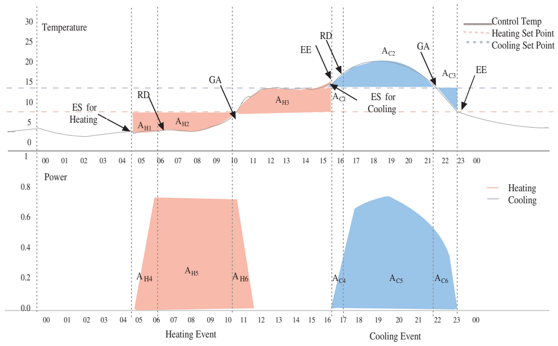


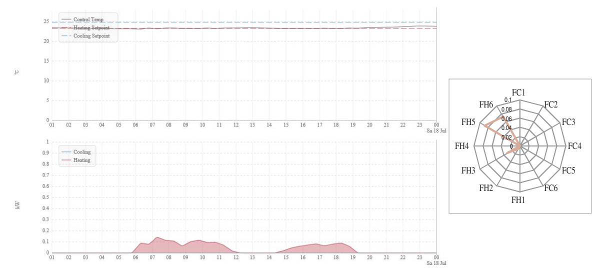
Fig. 4. Event discovery process for daily TU data

1) Event Start (ES): An event start is assumed to happen when the BMS is first switched on a particular day (enabled signal gets switched on) and the time instant when the temperature starts to change.
2) Response Delay (RD): It has been mentioned previously that due to the process variable delay during the dead time, the temperature starts to respond only after a certain delay from the previous point when the BMS is switched on and this event is termed as response delay. This is essentially the time spent by the TU during dead time as shown in Fig. 2.
3) Goal Achieved (GA): A goal achieved event is assumed to happen when the control temperature reaches the desired set point. GA can be considered as the time instant when the process variable reaches the steady state, or final value. This is essentially the time spent by the TU during rise time as shown in Fig. 2.
4) Event End (EE): Once the control temperature reaches the set point, it may either continue to be within the dead band till it exceeds the dead band, and an event end is supposed to happen at that time instant. This is essentially the time spent by the TU in the steady state and there could be a percentage overshoot above the final set point value as shown in Fig. 2.

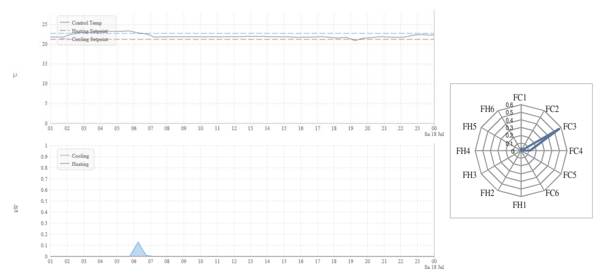
Following the event discovery stage, the event area calculations for both the temperature and power curves are executed with respect to the marked events for all the heating and cooling events happening during the day. Six areas are calculated for a heating event ($A_{H_1} - A_{H_3}$ for temperature curve, $A_{H_4} - A_{H_6}$ for power curve) and similarly, six areas are calculated for a cooling event ($A_{C_1} - A_{C_3}, A_{C_4} - A_{C_6}$). These areas are further normalized and aggregated to provide the input features for unsupervised machine learning using X-means clustering [6] technique.

III. Results and Experimental Analysis

The present case study is based on a building located in the city of London. The building has 17 floors and 731 terminal units spread across the different floors. The experimental analysis has been performed using a Java implementation on an Apache Spark cluster with the BEMS data being retrieved from a cloud based Cassandra database. Based on the data retrieved from these BEMS TUs and the application of the proposed feature extraction method, radars graphs as shown in Fig. 5 and Fig. 6 are obtained. Each radar graph represents an individual TU behaviour and each of its axes represents a separate feature. The six cooling features ($F_{C_1} - F_{C_6}$) are represented using axes 1 - 6 and six heating features ($F_{H_1} - F_{H_6}$) are represented using axes 7 - 12 respectively. Fig. 5 show good TU behaviours where the control temperature achieves the desired set point and little effort is required to maintain the temperature within the dead band.



(a) Temperature within dead band and low heating power

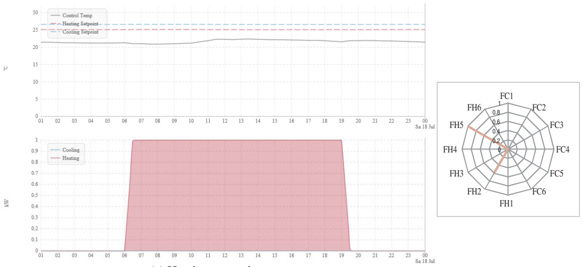


(b) Temperature within dead band with low heating power

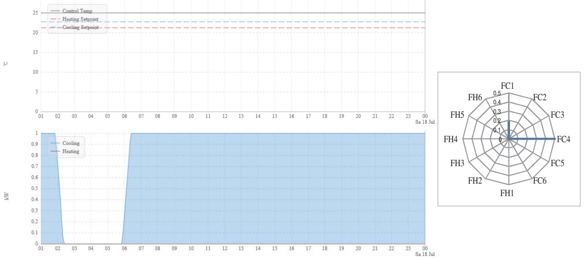
Fig. 5. "Good" TU behaviours and their corresponding feature radar graphs

Fig. 6 shows different bad TU behaviours with saturation, hunting and high temperature error patterns. Fig. 6(a) and (b) show saturation behaviours indicating high proportions of time over a full day that the valve or damper is open at maximum. Fig. 6 (c) shows hunting behaviour and shows how much the control temperature fluctuates over a day. Further, it can be seen from Fig. 6(b) and (c) that the TU are operating out of hours. This behaviour indicates a high degree of on-ness that is the proportion of time that a TU had any heating or cooling demand over a 24 hour period. Moreover, all of these TU with high average power also have high temperature errors that is control temperature deviates highly from the set point, indicating that they need to prioritised for further investigation and they are not only poorly controlling the temperature, but also consuming a relatively large amount of energy too.

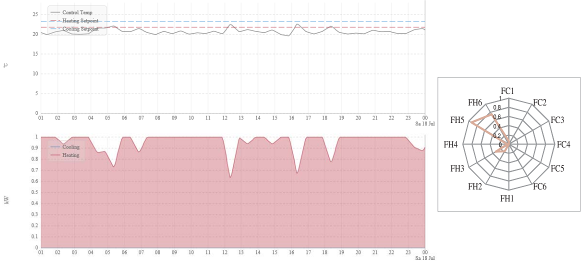
The aim of the clustering algorithm is to identify these different TU behaviours and Fig. 7 shows the clustering outcomes. The number of clusters has been chosen based on the BIC criterion [6] and further validated using the Silhouette



(a) Heating saturation



(b) Cooling saturation and on-ness



(c) Heating hunting and on-ness

Fig. 6. "Bad" TU behaviours and their corresponding feature radar graphs

indexing [7]. The distinct TU behaviours that are obtained as a result of clustering are described as follows:

1) Cluster C0: Most of the TUs achieve their goal. There is larger area under the curve between the RD to GA events for the temperature curve with minimal area under the curve for the ES to RD events for the power curve showing that these TUs exert little power to reach their goal.
2) Cluster C1: This cluster captures the TU behaviour where the areas are mostly in the GA to EE events and the TU applies medium to high power to reach the set point, therefore having area both under ES to RD events and RD to GA events for the power curve. The TUs in cluster C1 use more power than the TUs in cluster C0.
3) Cluster C2: These TU's have more area under GA to EE events as well as more area under the ES to RD events for the power curve which implies that the TU exerts initial medium power to achieve the goal.
4) Cluster C3: These TU's have more area under RD to GA events for both the temperature and power curves which indicates higher heating power is required, and the goal necessarily might not have been achieved.
5) Cluster C4: These TU's spend more time from the ES to RD events for both the temperature and power curves indicating that the TUs take longer time to reach the set point and hence struggle to achieve the goal.
6) Cluster C5: This cluster is similar to cluster C3 in terms of more areas under the RD to GA events for both the temperature and power curves, but captures the cooling TU behaviour and also the associated power levels are higher.

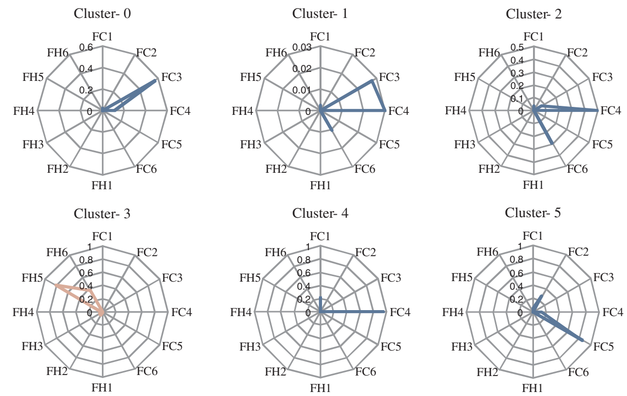


Fig. 7. Cluster wise feature distribution for daily TU analysis

IV. Conclusions and Future Work

Experimental results supports that the proposed new feature extraction technique is a very good approximation for HVAC TU data. Furthermore with the application of a clustering technique, a number of different cluster patterns have been identified that help in the identification of different TU behaviours. Currently, the work is implemented to a particular type of TU's (Fan coil units); but will be extended to different types of terminal units such as variable air volume (VAVs) and chiller elements etc. Additionally, based on the obtained clustering behaviours, categorical label assignment will be carried out to create an automated rule based TU classification system.

V. Acknowledgements

This work has been carried out as a part of the Innovate UK funded research project "Energy Management & Analysis Exploiting Existing BMS Infrastructure & Data" and the authors are grateful to the Demand Logic team for providing us with the BEMS data and invaluable information and feedback on building services and related performance issues.

References

[1] P. H. Shaikh, N. B. M. Nor, P. Nallagownden, I. Elamvazuthi, and T. Ibrahim, *"A review on optimized control systems for building energy and comfort management of smart sustainable buildings"*, Renewable and Sustainable Energy Reviews, 34, 409-429, 2014.

[2] Z. Gao, C. Carlo, and S. X. Ding, *"A survey of fault diagnosis and fault-tolerant techniques-Part I: Fault diagnosis with model-based and signal-based approaches"*, IEEE Transactions on Industrial Electronics 62(6):3757-3767, 2015.

[3] Y. Yu, Yuebin, D. Woradechjumroen, and D. Yu, *"A review of fault detection and diagnosis methodologies on air-handling units"*, Energy and Buildings 82, 550-562, 2014.

[4] K. Bruton, P. Raftery, B. Kennedy, M. M. Keane, and D. T. J. Osullivan, *"Review of automated fault detection and diagnostic tools in air handling units"*, Energy Efficiency, 7(2), 335-351, 2014.

[5] K. J. Astrom and T. Hagglund, *"Advanced PID control"*, ISA-The Instrumentation, Systems and Automation Society, 2006.

[6] Z. Huang, *"Extensions to the k-means algorithm for clustering large data sets with categorical values"*, Data mining and knowledge discovery, 2(3), 283-304, 1998.

[7] P. J. Rousseeuw, *"Silhouettes: a graphical aid to the interpretation and validation of cluster analysis"*, Journal of computational and applied mathematics, 20, 53-65, 1987.