Information Transfer Analysis for Ongoing Commissioning of Buildings

X. Rosalind Wang
CSIRO ICT Centre
PO Box 76, Epping
NSW 1710, Australia
Rosalind.Wang@csiro.au

Joseph T. Lizier
Max Planck Institute for
Mathematics in the Sciences
Inselstraße 22
04103 Leipzig, Germany
lizier@mis.mpg.de

Joshua Wall
CSIRO Energy Technology
PO Box 330, Newcastle
NSW 2300, Australia
Josh.Wall@csiro.au

ABSTRACT

Commissioning ensures buildings deliver the performance and energy savings promised by their design. Ongoing commissioning aims to fully automate the commissioning process, providing periodic or even continuous monitoring, identification and possibly remedial actions ensuring persistence in intended operation and performance . One of the major research areas inongoing commissioning at the moment is that of fault detection and diagnosis (FDD) of heating, ventilating, and air conditioning (HVAC) systems. First generation technology mainly focuses on detection of faults in sensors and actuators with manual diagnosis and remedial actions performed by system experts. In this paper, we propose an information theory method to analyse the information transfer between time series of the sensors and actuators (or the agents) of a HVAC system. The information transfer can help in the diagnosis of faults in the system by automatically identifying the related agents that caused the particular faults in another agent. We present some preliminary results from the proposed method using real data gathered from the a commercial office building in Newcastle, Australia.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

General Terms

Algorithms, Experimentation

Keywords

information dynamics, transfer entropy, energy efficiency, ongoing commissioning

1. INTRODUCTION

Literature on related building case studies suggest that virtually all buildings have some sort of *heating, ventilation, and air conditioning* (HVAC) operational problems, and the vast majority of buildings are not carefully commissioned [2, 11]. With considerable potential to achieve persistent energy savings at 'negative' cost, effective commissioning, post-occupancy evaluation, and *fault detection and diagnosis* (FDD) procedures could provide the key to immediate and significant emission cuts from the built environment.

We are interested in the problem of automating the FDD process for office buildings. Fault detection and diagnostics tools allows us to address many of the barriers to good commissioning and maintenance practice and consequential energy savings. Techniques from scientific disciplines such as artificial intelligence and machine learning [5, 6] are emerging as viable approaches where rules-based techniques may be less suitable.

Many current approaches focus on individual components, while this gives good fault detection algorithms, it does not allow for automatic fault diagnosis. In this paper, we examine the information transfer [8] between sensor and actuators (agents) within a building HVAC system. Information transfer is a recently published framework that characterises the information dynamics of distributed computation in terms of the elements of Turing universal computation. While information transfer is not a direct measure of causal effect [1, 7], it can be used for simple inference about causal relationships between two sets of measured data within the HVAC system. By knowing these relationships, we can then better diagnose faults by looking at those in related systems.

We begin the paper with an overview of building commissioning, HVAC systems and the continuous challenges in the field in Section 2. Section 3 introduces the framework for the information dynamics of distributed computation that was used. The experimental setup and data is shown in Section 4. We then present and discuss the results of information dynamics within the building energy sensors in Section 5. Finally, we discuss some future direction of this work and a conclusion.

2. ONGOING COMMISSIONING

The aim of commissioning new buildings is to ensure that they deliver, and sometimes exceed, the performance and energy savings promised by their design. When applied to existing buildings, commissioning identifies the almost inevitable "drift" from where things should be and puts the building back on course. Commissioning is a 'stealth' strategy for saving energy and money and reducing greenhouse gas emissions. Commissioning is also a risk-management strategy that ensures the building owners get what they paid for. [10]

Since each building is unique, the commissioning process would also be different for each new building. Commissioning is performed at widely different levels of effort and applied to buildings as a whole or a specific sub-system/end-use. Commissioning complements the energy efficiency practices by ensuring and maintaining building energy performance (and other benefits, such as indoor environmental quality). On the same token, it can simply focus on saving energy by improving conventional building systems, irrespective of whether or not the building is equipped to be particularly energy efficient.

In this paper, we focus on automated ongoing commissioning of the HVAC systems. Ongoing commissioning is a commissioning process that takes place in the building periodically. This is necessary since commissioning and retro-commissioning fixes do not persist over time [4].

The process of ongoing commissioning has two major steps: detection-diagnosis and repair. An automatic FDD system for HVAC will identify uncalibrated or malfunctioning sensors, valves, dampers and actuators; etc [4]. The repair step can be automated for tasks such as recalibration, new control strategies, and new set-points; or manual for tasks such as repair and re-installation of equipments. We are interested in the problem of automating the FDD process.

Fault detection and diagnosis are two separate processes. Detection aims to find data that deviates from the 'normal' behaviour of a system, while diagnosis is required to deduce the cause of the deviation. Automated FDD has become increasingly important in science and technology, however, it is a relatively recent discipline in building operation. Only within the last decade or so has the field been researched. A good review of the subject has been published by Katipamula and Brambley [5, 6].

FDD in building can be applied at various levels: from individual components such as heating coils, to specific systems such as the heating systems with all the components, to to the entire building, taking into account all systems such as heating, cooling, lighting, etc. Ideally, applications would complement each other, however, the current situation is far from achieving this goal.

Most FDD methods available now are designed and developed for individual components whose design and thermodynamic behaviour are already know. Special FDD methods are developed on the basis of individual faults. This detailed knowledge of the system is typical of FDD at component level and distinguishes it in principle from FDD operating at system or building level [3].

One reason for most of the research up until now to focus on individual components is that at the macro level of the entire building, most of the parameters are often not known. An example of these parameters is the effects of the occupants' behaviour on energy consumption in a building. At the moment, an accurate representation of the system or of the individual faults is only possible to a limited extent. Therefore, in more complex systems, fault analysis is often carried out manually by an expert. In this paper, we aim to use information theory to discover the relationships between various sensors and actuators within the HVAC system, thus allowing the FDD to operate at a system level, and ultimately, the building level. Knowing the relationship between various sensor and actuator's time series data will also allow the FDD to diagnose faults quickly and autonomously.

3. INFORMATION TRANSFER

Information theory [9] is an increasingly popular framework for the study of complex systems. In part, this is because complex systems can be viewed as distributed computing systems, and information theory is a natural way to study computation, e.g. [8]. Information theory is applicable to any system, provided that one can define probability distribution functions for its states. This is a particularly important characteristic since it means that information-theoretic insights can be directly compared across different system types.

The fundamental quantity for information theory is the (Shannon) *entropy*, which represents the uncertainty in a sample x of a random variable X. Shannon Information [13] was originally developed for reliable transmission of information from a source X to a receiver Y over noisy communication channels. When dealing with outcomes of imperfect probabilistic processes, it is useful to define the information content of an outcome x, which has the probability $P(x)$, as $-\log_2 P(x)$. Crucially, improbable outcomes convey more information than probable outcomes. Given a probability distribution P over the outcomes $x \in \mathcal{X}$, the average Shannon information content of an outcome is determined by

$$H(X) = -\sum_{x \in \mathcal{X}} P(x) \log_2 P(x), \tag{1}$$

Intuitively, Shannon information measures the amount of freedom of choice (or the degree of randomness) contained in the process — a process with many possible outcomes has high entropy.

Transfer entropy (TE) [12], a measure to formulate the information transfer based on information theory, can be useful in ongoing commissioning work. The information transfer between a source and a destination agent is defined as the information provided by the source about the destination's next state that was not contained in the past of the destination. While information transfer is not a direct measure of causal effect [1, 7], it can be used for simple inference about causal relationships between two sets of measured data within the HVAC system. For example, if the temperature within a zone is directly dependent on the speed of the fan, then there should be a large amount of information transferred from the fan speed to the temperature data.

Transfer entropy does not require a model of the interaction and is inherently non-linear. Transfer entropy addresses concerns that the mutual information, $I(X;Y) = H(X) - H(X|Y) = H(X) + H(Y) - H(X,Y)$, as a de facto measure of information transfer was a symmetric measure of statically shared information. The transfer entropy from a source agent Y to a destination agent X is the mutual information between the previous state of the source y_n and the next state of the destination x_{n+1}, *conditioned* on the semi-infinite past of the destination $x_n^{(k)}$ (as $k \to \infty$ [8]):

$$T_{Y \to X} = \lim_{k \to \infty} \sum_{\mathbf{u}_n} p(\mathbf{u}_n) \log_2 \frac{p(x_{n+1}|x_n^{(k)}, y_n)}{p(x_{n+1}|x_n^{(k)})}, \tag{2}$$

where $\mathbf{u}_n$ is the state transition tuple $(x_{n+1}, x^{(k)}, y_n)$, $T_{Y \to X}(n,k)$ represents finite-k approximation.

4. EXPERIMENTAL SETUP

The data used for this paper was gathered at the Australian Commonwealth Scientific and Research Organisation (CSIRO)'s Energy Centre in Newcastle, New South Wales. The building was completed in 2003, with a total air-conditioned area of 3320 m^2 divided into 15+ zones on 4 levels. A HVAC zone is a group of adjacent offices and/or

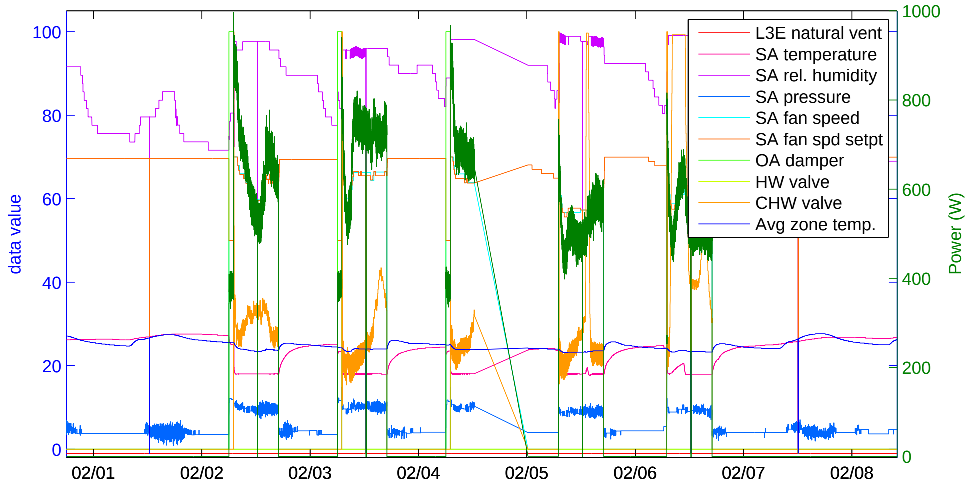

Figure 1: A example of the data used for the experiment, the days shown here represents a week from Sunday 1st to Sunday 8th.

spaces serviced by a common air-handling unit (AHU) or air-terminal device. Air is delivered to each zone via underfloor supply air plenums and floor air terminals. Return air is circulated back to the plant rooms via high level grilles. Space temperature control is provided through modulation of supply air temperature and and fan speed. The HVAC system accounts for approximately 57% of the total building electrical load.

A number of measurements have been conducted to gather historical data on Level 3 (East) office wing, in which three AHUs (AHU 08, 09 and 10) service three adjacent HVAC zones consisting of 18 offices and associated corridors. For this paper, we used the data gathered for AHU-10 (which services 8 adjacent offices) along with one set of data for level 3. The sensors and actuators used were:

1. Level 3 East natural ventilation on/off,
2. AHU-10 fan power (W),
3. AHU-10 supply air temperature (°C),
4. AHU-10 supply air relative humidity (%),
5. AHU-10 supply air pressure (Pa),
6. AHU-10 supply air fan speed (%),
7. AHU-10 supply air fan speed set point (%),
8. AHU-10 outside air damper (%),
9. AHU-10 hot water valve (%),
10. AHU-10 chilled water valve (%),
11. AHU-10 average zone temperature (°C).

We took the data collected between 0900hr 8th Jan and 1150hr 9th Feb 2009, the data was sampled at 5 seconds interval. It should be pointed out that this takes place in the middle of Australian summer, thus the HVAC system is mainly engaged to cool the building. Figure 1 shows a portion of the data used, the dates here represents a week in February.

5. RESULTS AND DISCUSSION

To calculate the transfer entropy between pairs of data sets, we set the history size of $k = 2$. There are a total of approximated 336000 data points in each set of data, therefore we used a kernel width of 1 for kernel estimation with a normalised kernel.

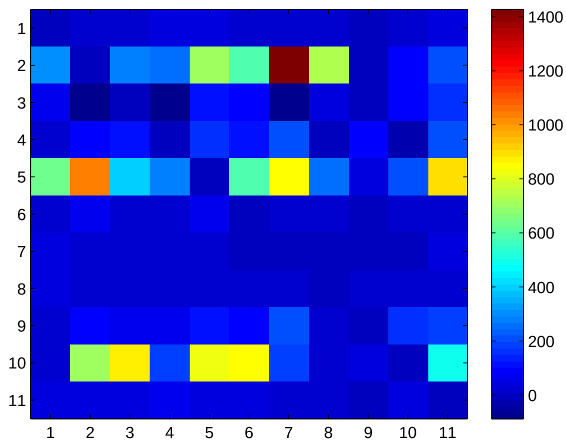

Figure 2: Transfer entropy calculated for the data sets described in Section 4. The rows represent the *destination* agent, and the columns represent the *source*.

Figure 2 shows the resulting total transfer entropy between pairs of the data sets as described in Section 4. Each row represents the destination agent of the transfer and each column represents the source agent. For example, cell (2, 7) represents the transfer entropy from data set 7 (which is AHU-10 supply air (SA) fan speed set point) to data set 2 (which is AHU-10 fan power). The units of shown in the image is bits.

One of the first feature to notice on Figure 2 is that it is not symmetrical. This is to be expected, since the causal relationship should be directional. So, while we expect the supply air's fan speed to affect the supply air temperature, we don't expect the reverse. This was the reason that mutual information isn't used in determining the causal relationship as it is unidirectional.

Another expected result that we see is that the diagonal is zero, since we wouldn't expect any transfer within one data source. This does not mean that an agent's past state doesn't affect it's future, rather transfer entropy calculates the information provided by a source about a destination's next state that was not contained in the past of the destination. Therefore, if a source is also the destination, then the terms $p(x_{n+1}|x_n^{(k)}, y_n)$ and $p(x_{n+1}|x_n^{(k)})$ in Equation 2 will be equal, thus TE is zero.

The highest transfer entropy shown in Figure 2 is that in cell (2, 7), indicating a positive transfer of information from the SA fan speed set point to the fan power. This might seem surprising at first since set points are usually constant for a long time and do not change often. However, looking closely at the data we have, we found that the set point in our case does get altered regularly, presumably to cool the building as outside temperature rose.

Of the other systems' relationship with fan power, we see that the water valves have very little transfer entropy, which is reasonable. The average temperature of the zone also doesn't have much effect, due to the fact we are only measuring one past step of the source agent, as it will take a lot more than 5 seconds (which is the time interval used in the data) for the average zone temperature to affect the fan power.

The other time series that was affected by many others was the SA pressure (row 5 in Figure 2. We can see from the results that it was most notably affected by (in order of transfer entropy value) fan power, average temperature of the zone, SA fan speed set point, and whether the natural ventilation of the floor was on or off which means whether the air conditioning system was on or off.

There are some relationships in Figure 2 that were expected to appear but didn't. For example we expect that the position of the chilled water valve will affect some of the time series such as the average zone temperature. This is due to the fact that the data was gathered at 5 second interval and we only used one step of the source agent's data, which would not be enough time for effects to be noticeable. Another reason for these none-appearance is that the data was not pre-processed before the TE calculations were made. This is especially a problem for systems that are turned off overnight, for example the fan power, as we see in Figure 1. Therefore, when we calculated the total TE using all the data, a lot of noise was introduced to the results, thus covering up some of the features that could showed up.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we measured the information transfer between time series data of various sensors and actuators in an HVAC system. The amount of information transfer can be used for simple inference about the causal relationship between two sets of data. This relationship can then be used to allow us better diagnose faults detected within the system.

We found with our preliminary investigation that for most pairs of time series, the measured information transfer provides a good indication of the causal relationship between the two. However, we also found with other sets, the expected causal relationship were not present in the results. There were two main reasons for those: firstly the data was recorded at a 5 second interval, which is too little time for expected effect to appear; secondly, as the data was not preprocessed, some time series still has a lot of noise in the data, which obscured the real features in the data. In future works, we would like to look into having a longer history length for the source agent's time series; and pre-process the data to eliminate the noise before calculating the transfer entropy. We also aim to apply this method to similar data from other buildings.

7. ACKNOWLEDGMENTS

We like to thank the High Performance Computing and Communications Centre (http://www.hpccc.gov.au/) for the use of their supercomputer clusters in performing the experiments for this paper.

8. REFERENCES

[1] N. Ay and D. Polani. Information flows in causal networks. *Adv. Complex Syst.*, 11(1):17–41, 2008.

[2] D. E. Claridge, J. Haberl, M. Liu, J. Houcek, and A. Athar. Can you achieve 150% of predicted retrofit savings? is it time for recommissioning? In *Proc. ACEEE 1994 Summer Study on Energy Efficiency in Buildings*, volume 5, Aug. 1994.

[3] L. Jagemar and D. Olsson. The EPBD and continuous commisioning. Technical report, CIT Energy Mangement AB, Göteborg, Sweden, Oct. 2007.

[4] S. Katipamula. Automated continuous commissioning of AHUs. In *ASHRAE Winter Annual Meeting*, Chicago, USA, Jan. 2003.

[5] S. Katipamula and M. R. Brambley. Methods for fault detection, diagnostics, and prognostics for building systems - a review, part i. *HVAC&R Research*, 11(1):3–25, Jan. 2005.

[6] S. Katipamula and M. R. Brambley. Methods for fault detection, diagnostics, and prognostics for building systems - a review, part ii. *HVAC&R Research*, 11(2):169–187, Apr. 2005.

[7] J. T. Lizier and M. Prokopenko. Differentiating information transfer and causal effect. *EPJ B*, 73(4):605–615, 2010.

[8] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya. Local information transfer as a spatiotemporal filter for complex systems. *Phys. Rev. E*, 77(2):026110, 2008.

[9] D. MacKay. *Information Theory, Inference, and Learning Algorithms.* Cambridge University Press, 2003.

[10] E. Mills. Building commisioning: A golden opportunity for reducing energy costs and greenhouse gas emissions. Technical report, Lawrence Berkeley National Laboratory, July 2009.

[11] M. A. Piette and B. Nordman. Costs and benefits from utility-funded commissioning of energy-efficiency measures in 16 buildings. *ASHRAE Transactions*, 102(1):LBL–36448, 1996.

[12] T. Schreiber. Measuring information transfer. *Phys. Rev. Lett.*, 85(2):461–464, 2000.

[13] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656, July, October, 1948.