

DATA REASONING IN THE EVALUATION OF DOMESTIC THERMAL ENERGY USE

Chris Fishlock, Keerthi Rajendran, Mohammad Royapoor and Anthony P. Roskilly
Sir Joseph Swan Centre For Energy Research, Newcastle University, UK

ABSTRACT

The decreasing cost of sensing equipment and developments in the field of data science are providing increased opportunities for the validation and enhancement of existing knowledge and solutions across many fields. With the primary aim of supporting the optimisation of domestic thermal energy use, this paper documents the early stages of the application of a data centric approach to extend the understanding of energy use at an individual property level. To facilitate this, a Semantic Web platform is designed, providing the foundation on which factors influencing thermal energy use are inferred using data reasoning techniques.

INTRODUCTION

Use of energy within a domestic context is a significant component of overall energy consumption, contributing 29% of total within the UK (Department of Energy & Climate Change, 2013). When considering the final consumption of energy within a domestic setting, 82% is used to meet space and water heating requirements.

With current energy use a primary contributor to greenhouse gas emissions, and in light of UK commitment to an 80% emissions reduction by 2050 (UK Government, 2008), optimisation of energy use within the home provides a significant opportunity. Optimisation is achievable through two overlapping approaches, the reduction of overall energy use, and the utilisation of low emission sources to meet demand.

In considering optimisation from the perspective of an individual property, a clear understanding is required as to what, when, why and how energy is used within the property. To address these aspects, this study proposes a data centric approach to the evaluation of thermal energy use, applying data analytics to gathered information to infer detail about any specific property of study. The selected approach will be evaluated against 4 main objectives.

Ability to facilitate decision support for optimisation. Through computational analysis of the data collected, the inferred understanding will be evaluated against the ability to provide information to residents and building managers with the aim of

reducing the environmental burden of energy use.

Data evaluation. Provide evaluation of the effectiveness of different collected data in analysis, considering both time series sensor data and information describing the building, its contents and use. Assessment of data will also consider the wider aspects of its collection, covering cost, complexity, reliability and the privacy of residents.

Reuse. Any analysis techniques developed should not be specific to a building under study and should be applicable to the evaluation of any domestic property.

Extensibility. Consideration given to the ability for further extension of the approach adopted to a wider selection of data sources, buildings, and for the use of the information beyond the core aims of the study.

APPROACH

The approach selected by the study results from initial experimentation on the capture of temperature data in multiple locations across a single property. From initial observational review of the data it was apparent that variations between rooms could be observed, with events such as heating and shower use also identifiable.

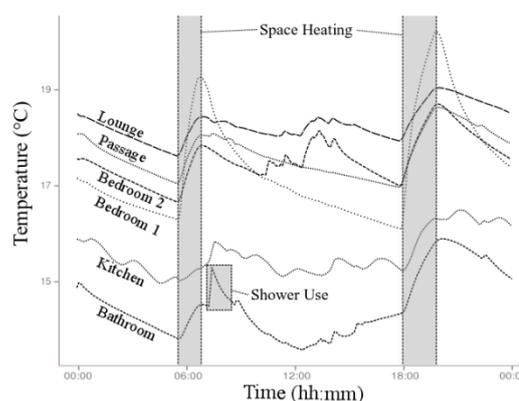


Figure 1 Initial Data Observations Over 24hr period

The initial human interpretation of the gathered information provided no quantitative assessment of the scenario, reliant on both the data captured and an understanding of both the building in which the sensors were located, and the building's use.

With the human undertaking of quantitative

assessment of the studied scenario significantly time consuming, especially when evaluating many identified features within the data, or multiple buildings, consideration turned to the opportunity for computational evaluation. In order for computational interpretation of the information to be conducted, an understanding of the processes used to infer an event was used. Using the example above:-

The heating event (central heating) was identified from the knowledge that the property has central heating serving all the rooms and that a rise in temperature occurred across the rooms, initiated concurrently.

The shower use event was identified by the knowledge that the room in which a temperature increase occurred, independently of the other rooms, was the bathroom, a room in which a shower is located.

Within the examples, although other causes of the observations could exist, reasonable confidence is possible from the wider understanding of the building and its use, assisted by the relative homogeneity of the domestic context (Holzmann et al., 2013). Depending upon the situation and identified event, additional information could increase confidence by ruling out other potential options. For example, in the case of shower use, this could be supplemented by humidity data for the room, and energy data for the property. The consideration of such additional information in providing confidence to inferencing techniques is central to meeting the data evaluation objective of the study.

With reference to the objectives of the study, although the identification of events provides an interesting computational challenge, this alone is not sufficient for the provision of feedback and decision support. The identification and quantitative description of features inferred from the data enables further evaluation, conducting comparative and trend analysis against time, between locations within an individual building, and between different buildings. Within the example in Figure 1, this could be demonstrated by the detailed profiling of temperature increases during heating, and the different rates of subsequent cooling, highlighting the relatively high temperature rise and rate of cooling within Bedroom 1.

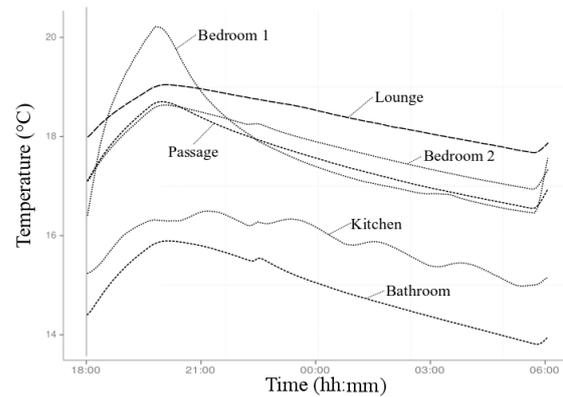


Figure 2 Heating and cooling across a property

Such an approach provides supplementary analysis to other approaches adopted such as surveys and building modelling. By directly observing the property and its use, both physical and occupancy characteristics can be evaluated. With low cost sensing technologies increasingly available, monitoring can be conducted at low overall cost, whilst providing the opportunity to capture variation over an extended period. Although some detail is required on the physical aspects of the building to provide context to sensor data, the use of inference against held information is evaluated against minimising this. The approach targets easy implementation by home owners, through simple sensor configuration and the minimising of questions to inform the model.

In facilitating the study, with reference to the primary objectives and computational requirements of the work, a Semantic Web data platform was selected, the justification of which is detailed in the following overview.

THE SEMANTIC WEB

Background

The concept of the Semantic Web builds on the key concepts forming the foundation of the World Wide Web, providing increased structure so that “information is given well-defined meaning, better enabling computers and people to work in cooperation” (Berners-Lee et al., 2001). In reducing the term Semantic Web, the two key components are highlighted. Here, 'Web' highlights the scalable interlinking of information, a foundation of the World Wide Web from inception, whilst 'Semantic' highlights the adding of meaning to stored data.

In assigning meaning to information it is clear that additional information will be required to define this structure, for example in describing a house :-

A house can have many rooms.

Rooms within a property can have specific usage, for example a kitchen for cooking, or a bathroom for washing.

The structure of this “data about data”, or metadata, within the context of the domain of study builds to form an ontology describing the possibilities of the information held. Such a template against which data is stored, provides a knowledge of the structure and relationships expected within the information, importantly encoded in a machine interpretable form.

Central to the Semantic Web approach is the standardisation of language or syntax on which a domain specific ontology is defined. A base for such languages is provided by the Resource Description Foundation (RDF), a standard governed by the World Wide Web Consortium (W3C).

Although the RDF format provides a foundation on which a Semantic Web ontology may be built, increasingly expressive languages have subsequently been defined, providing the ability to define information with more detail and clarity. Of these, the Web Ontology Language (OWL) and its derivatives, again governed by the W3C, are the most widely adopted (Grzybek et al., 2014).

A powerful aspect of the approach is that, although a developed ontology is specific to its field of use, ontology languages, such as RDF and OWL, on which the ontology is defined, are not. Tools have been developed to assist all elements of data lifecycle, covering the design, storage, querying and, specifically important to the study, inference and reasoning against the information held against an ontology. The ability to utilise predefined tools for significant elements of the work reduces the requirement for bespoke development, whilst the adoption of existing standards by the study could assist further future adoption and extension by other parties.

With machine interpretation of the compiled data enabled by the ontology language, reasoning tools provide the capability of evaluating data held against the ontology definition to which the data conforms. Such tools provide the ability to highlight discrepancies between data and the ontology, and to infer additional relationships defined within the ontology that may not have been specifically specified in the data. This ability is central to the interpretation of information gathered by the study, enabling a number of observations to be combined to identify cause from existing knowledge held in ontological form.

Applications in Building Research

Although far less adopted in building research when compared to fields such as biomedicine, evidence of the study of semantic web techniques has been undertaken within both domestic and commercial building research.

Within a domestic context, the study of domotics provides the demonstration of use of a Semantic Web platform in home automation. Here the DOGOnt ontology, initially defined by Bonino and Corno,

(2008), provides a single platform for the integration of multiple technologies, an approach analogous to the integration of multiple sensor data sources within this study. Subsequent studies have extended the ontology, energy efficiency measures through automated appliance control (Kofler et al., 2012) (Görner et al., 2015), utilising reasoning techniques to enable this.

Within a commercial building context, the ontological extension of Building Information Modelling (BIM) remains an on-going work where data reasoning has been used to conduct cost estimation (Lee et al., 2014), safety assessment (Zhang et al., 2015), and to identify the embodied energy of individual buildings (Hou et al., 2015). With the studies providing a justification for the increasing adoption of the Semantic Web in building projects, work undertaken in the ontological representation of BIM data must be noted, as evaluated by Pauwels and Terkaj, (2016). In response to the lack of a Semantic Web standard for the representation of building data, a working group has been founded facilitating the exchange of approaches and best practices within the field (“Linked Building Data Community Group,” 2016).

Although of limited detail, possibly due to the commercial nature of the study, the multi dwelling work of Chaussecourte et al., (2013), demonstrates decision support from energy meter readings utilising OWL based reasoning. An interesting aspect of the study is the encapsulation of decision support information directly into the ontology itself, an approach to be considered within this study once analysis techniques have been developed.

METHODOLOGY

Experimentation

To facilitate thorough evaluation of both the data and computational techniques, a multi-building study was defined. In addition to providing multiple scenarios against which to test the identification and classification of features within the collected data, the approach provides the opportunity to conduct comparative and trend analysis both within and between the properties of study.

The deployment of 21 Tinytag TK-4014 temperature sensors provides an initial foundation for the consideration of temperature data, with 7 sensors selected per house, 6 for internal use in the core areas of the building, and 1 mounted in a sheltered external location. In order to increase the range of properties covered, it was decided to implement week long sampling of a greater number of buildings, returning the sensors to each property every 3 months, thus providing quarterly data for each building. With the Tinytag sensors requiring the manual downloading of data, the adoption of week-long sampling provided regular access to the data, allowing for a timely review of the methods adopted.

The consideration of wider parameters is achieved through the implementation of Open Energy Monitor

and Tensor plc solutions in single properties for the duration of the study, as outlined in Table 1.

Table 1 Sensing Equipment Initially Deployed

Product	Tinytag	Open Energy Monitor	Tensor plc
Type	Data Logger	Wireless (433Mhz)	Wireless (Zigbee)
Deployment	1 week samples from 10 properties	Continuous sensing, initially in one property	Continuous sensing, initially in one property
Data	Temperature (Supplemented by manual gas and electricity meter readings)	Temperature R. Humidity Electricity (Supplemented by manual gas meter readings)	Temperature R. Humidity Electricity Gas CO2 PIR

Providing an energy context to the Tinytag instrumented properties, data is supplemented by manual gas and electricity readings taken at the start and end of each monitoring period, prior to the sensors being returned to the university for data retrieval. Manual daily gas meter readings are also being taken from the Open Energy Monitor instrumented property.

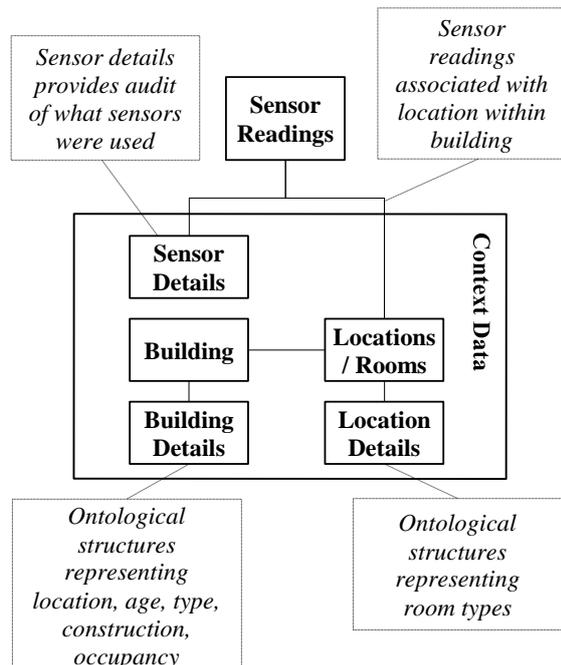


Figure 3 Building Ontology Components

Context Data

In initiating the experimentation, a questionnaire was completed by the occupants of each property in order to obtain a brief overview of the building, heating equipment and occupancy details. This data is combined with the captured sensor data to form the Building Ontology, providing a comprehensive

overview of an assessed property. In addition to enabling reasoning to be conducted within the study, this approach assists future researchers in explaining the context in which data was captured (Bechhofer et al., 2013).

With intrusion and detail of context data minimised at this stage, regular review of the information requirements is conducted, judging the data against the ability to undertake analysis of the scenario.

Analysis Design

Analysis of the captured data was divided into 3 separate interlinked stages, with the output of each stage forming an input to the next. This modular approach mirrors that of the platform design, allowing for differing approaches to each stage to be considered whilst minimising the impact on the wider analysis.

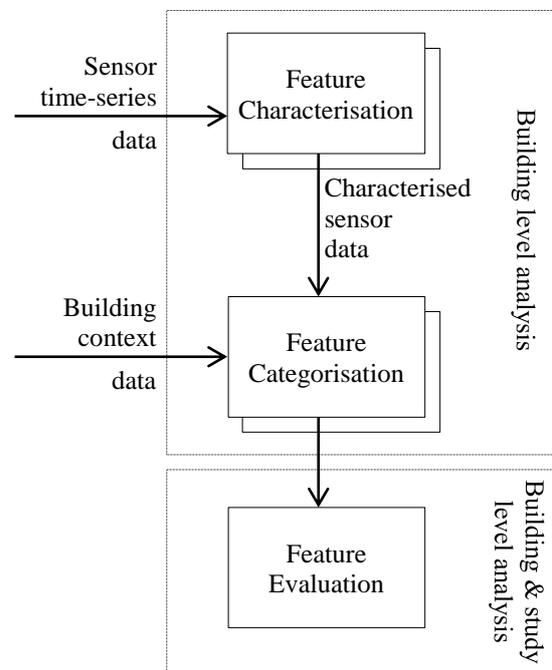


Figure 4 Analysis Components

1) Feature Characterisation. Through computation against the raw time-series data, the aim of this stage is to identify signatures within the raw time series data, representing for example temperature increases, periods of cooling and cyclical patterns within the data. In conducting this stage, the use of context data is minimised, maximising the reuse of the processing techniques across differing data sources.

Algorithm design for characterisation is conducted using the R statistical computing platform, a tool capable of directly interrogating data stored within a Semantic Web structure, using the SPARQL query language. Once algorithms are identified, the logic is implemented outside of the R platform where automation is used to test signature identification across all the data held.

This stage aims to produce semantically structured output representing forms identified within raw data, upon which inference is conducting in the following stage.

2) Feature Categorisation. With the characterised structures providing information regarding the form of the features or signatures identified in the raw sensor data, categorisation utilises reasoning to associate these features with a likely cause, though combining this information with the building context data. In order to conduct the reasoning, the stage develops an additional Analysis Ontology, on which knowledge of likely causes is defined. This stage requires minimal bespoke development of code, merely the design of data structures representing an understanding of the likely causes, against which generic OWL reasoning tools are deployed.

3) Feature Evaluation. Although the identification and characterisation of features within the sampled data provides an interesting computational study, this is insufficient to provide a foundation on which to conduct decision support. Through the automated application of the first two stages across all data held, analysis can be conducted across the features, comparing findings within properties, between properties, and against time.

Platform Design

The approach detailed in the analysis design has been selected to minimise the requirement for bespoke software development within this study. The core knowledge and intelligence of the platform is achieved through the deployment of generic reasoning against domain knowledge stored within a combination of the Building Ontology and Analysis Ontology.

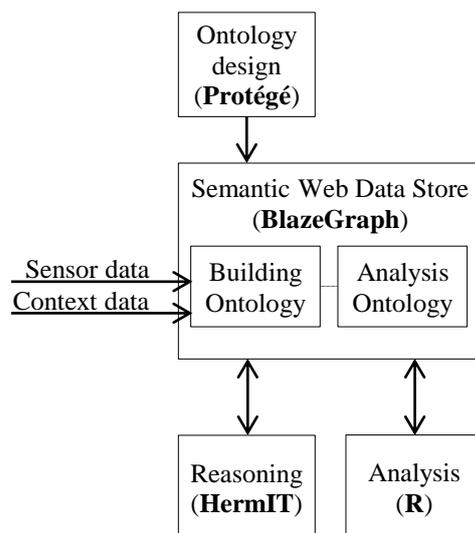


Figure 5 Platform Design

The decoupling of components within the solution allows for the future extension of the approach to incorporate a consumer facing view of the information, implementing decision support based on

the analysis undertaken.

DISCUSSION

The designed platform outlined earlier has been implemented and populated with data representing the properties being studied, covering detached, terraced and semi-detached dwellings, obtained from initial questionnaire. Time series data from the initial samples has also been incorporated into the Building Ontology, utilising software scripts to import data from file from each of the different sources. More advanced approaches are potentially possible providing near real-time access to the data obtained by the Tensor and Open Energy solutions, with data transmitted from the sensor platform to a remote data repository. Although not undertaken as part of the current platform, it may be possible to automate the extraction of data from the respective stores, depending upon the architecture of the two platforms.

The design of feature identification approaches is a current area of focus of the study, initially considering more isolated signatures, such as that observed during overnight cooling of the properties. Here external factors such as solar gain and occupant activity are removed or reduced, simplifying the detection of structure in the data.

In addition to the identification of anticipated features, characterisation of unwanted components has also been initiated. An example of this is the observation of a cyclical temperature signature from a kitchen sample, caused by the close proximity of the sensor to refrigeration equipment (observable within Figures 1 & 2). External readings obtained have also shown some exposure to direct sunlight, a factor potentially alleviated by supplementing gathered data with an external source of weather information for the location. Future automation of analysis could be used to highlight such erroneous data, enabling the timely correction of sensor configuration.

At this stage, no additional context data has been requested beyond that captured by questionnaire in initiating the study. In order to validate the understanding inferred during data reasoning and analysis, it is anticipated that some level of verification against resident activity may be required. Although as yet undefined, this may take the form of resident logging, or preferably as less intrusive, through direct regular reporting of analysis to the occupant for review.

CONCLUSION

From implementation and initial analysis, the data reasoning approach presented here provides a scalable, extensible platform against which an initial level of understanding can be inferred from a minimal range of captured data.

Such a platform provides a foundation on which increasingly intelligent reasoning can be conducted. Combining this with consumer feedback, the

gathering of additional information through targeted questioning could enable the extension of the knowledge base, forming a case based reasoning approach similar to that documented by Yang, (2013) in consideration of electricity use.

Beyond the direct use of the data within this study, data captured within the Building Ontology could provide benefit to other research, providing an opportunity to integrate with existing modelling and assessments such as SAP, the supplementing of which with empirical data is suggested by Laurent et al., (2013).

With analysis and knowledge also represented as data against the Analysis Ontology, the potential for use of this information beyond decision support is maximised. The greater adoption of the approach could provide a repository detailing real-world scenarios against which modelling and future building technologies could be evaluated.

REFERENCES

- Bechhofer, S., Buchan, I., De Roure, D., Missier, P., Ainsworth, J., Bhagat, J., Couch, P., Cruickshank, D., Delderfield, M., Dunlop, I., Gamble, M., Michaelides, D., Owen, S., Newman, D., Sufi, S., Goble, C., 2013. Why linked data is not enough for scientists. *Future Gener. Comput. Syst.* 29, 599–611. doi:10.1016/j.future.2011.08.004
- Berners-Lee, T., Hendler, J., Lassila, O., 2001. The semantic web. *Sci. Am.* 284, 28–37.
- Bonino, D., Corno, F., 2008. DogOnt - Ontology Modeling for Intelligent Domestic Environments, in: Sheth, A., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T., Thirunarayan, K. (Eds.), *The Semantic Web - ISWC 2008*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 790–803.
- Calderón, C., James, P., Urquizo, J., McLoughlin, A., 2015. A GIS domestic building framework to estimate energy end-use demand in UK sub-city areas. *Energy Build.* 96, 236–250. doi:10.1016/j.enbuild.2015.03.029
- Chaussecourte, P., Glimm, B., Horrocks, I., Motik, B., Pierre, L., 2013. The Energy Management Adviser at EDF, in: Alani, H., Kagal, L., Fokoue, A., Groth, P., Biemann, C., Parreira, J.X., Aroyo, L., Noy, N., Welty, C., Janowicz, K. (Eds.), *Proc. of the 12th Int. Semantic Web Conf. (ISWC 2013)*, In-Use Track, LNCS. Springer, Sydney, NSW, Australia, pp. 49–64.
- Department of Energy & Climate Change, 2013. *Energy Consumption In The UK*.
- Görner, M., Göschel, T., Kassel, S., Sander, S., Klein, T., 2015. An Ontology for Interoperability: Modeling of Composite Services in the Smart Home Environment, in: van Sinderen, M., Chapurlat, V. (Eds.), *Enterprise Interoperability, Lecture Notes in Business Information Processing*. Springer Berlin Heidelberg, pp. 30–38.
- Grzybek, H., Xu, S., Gulliver, S., Fillingham, V., 2014. Considering the Feasibility of Semantic Model Design in the Built-Environment. *Buildings* 4, 849–879. doi:10.3390/buildings4040849
- Holzmann, A., Adensam, H., Kratena, K., Schmid, E., 2013. Decomposing final energy use for heating in the residential sector in Austria. *Energy Policy* 62, 607–616. doi:10.1016/j.enpol.2013.07.027
- Hou, S., Li, H., Rezgui, Y., 2015. Ontology-based approach for structural design considering low embodied energy and carbon. *Energy Build.* 102, 75–90. doi:10.1016/j.enbuild.2015.04.051
- Kofler, M.J., Reinisch, C., Kastner, W., 2012. A semantic representation of energy-related information in future smart homes. *Energy Build.* 47, 169–179. doi:10.1016/j.enbuild.2011.11.044
- Laurent, M.-H., Allibe, B., Oreszczyn, T., Hamilton, I., Galvin, R., Tigchelaar, C., 2013. Back to reality: How domestic energy efficiency policies in four European countries can be improved by using empirical data instead of normative calculation.
- Lee, S.-K., Kim, K.-R., Yu, J.-H., 2014. BIM and ontology-based approach for building cost estimation. *Autom. Constr.* 41, 96–105. doi:10.1016/j.autcon.2013.10.020
- Linked Building Data Community Group [WWW Document], n.d. URL <https://www.w3.org/community/lbd/> (accessed 9.28.15).
- Pauwels, P., Terkaj, W., 2016. EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology. *Autom. Constr.* 63, 100–133. doi:10.1016/j.autcon.2015.12.003
- UK Government, 2008. *Climate Change Act 2008*.
- Yang, S.-Y., 2013. Developing an energy-saving and case-based reasoning information agent with Web service and ontology techniques. *Expert Syst. Appl.* 40, 3351–3369. doi:10.1016/j.eswa.2012.12.044
- Zhang, S., Boukamp, F., Teizer, J., 2015. Ontology-based semantic modeling of construction safety knowledge: Towards automated safety planning for job hazard analysis (JHA). *Autom. Constr.* 52, 29–41. doi:10.1016/j.autcon.2015.02.005