Challenges in Developing a Generic Monitoring Framework for Pan European Energy Usage and Environmental Monitoring

David While, Jan Krasniewicz, Sharon Cox

School of Computing and Digital Technology
Birmingham City University
Birmingham, UK
Jan.Krasniewicz@bcu.ac.uk
Sharon.cox@bcu.ac.uk

Abstract— **The ability to easily monitor different aspects of the environment is essential to achieve the aspirations of smart homes, smart buildings and smart cities across Europe. A wide range of sensors are available for both the domestic and commercial markets to enable different aspects of the environment to be monitored. These sensors are disparate, requiring different interfaces and utilizing conflicting data formats. This paper reports on the development of a generic monitoring framework to capture and analyse data from ubiquitous sensing devices in smart cities. The framework has been tested by capturing energy usage data from both public and commercial buildings, and domestic homes in three cities across two European countries. Dashboards were created to enable facility managers and home owners to compare energy usage with similar buildings in different cities. The paper discusses the technical and data quality challenges encountered with capturing data from domestic and non-domestic buildings and highlights the need for a generic context framework to support monitoring and analysis of the pan-European data captured.[1]**

Keywords— *API, BEMS, energy monitoring, sensor networks, smart cities, smart homes, software architecture.*

I. Introduction

A third of the world's energy consumption is attributed to buildings; in the UK, buildings account for 39% of energy consumption; across Europe this increases to 42% [1]. Environmental monitoring systems and systems that capture data about energy usage in buildings are therefore important to support data driven approaches to improve energy efficiency [2]. The embedding of information communication technology (ICT) within an environment to transform an activity is one view of a smart city [3]. Smart cities are based on the premise that improving access to information improves decision making about consumption of resources, such as energy [4]. This requires data to be available to support both real-time and long-term data analysis [5] to provide timely and meaningful information that can be acted upon.

A range of sensing devices is readily available in both domestic and commercial markets to monitor all aspects of the environment such as temperature, air quality, lighting levels, and energy consumption, and to detect events such as movement. Sensor networks are inherently distributed [6] and heterogeneous [7] as sensor technologies have evolved independently, providing limited opportunity for integration. A city incorporates a range of domestic and non-domestic buildings, therefore in order for a city to be truly smart, data from a range of sources needs to be consolidated, integrated and organized to provide actionable information. A ubiquitous infrastructure is needed to support context-aware applications [8] that make sense of data received from diverse sensing devices. The infrastructure needs to have the ability to aggregate data from disparate sensors, be able to drill down through the data [9], and be able to interrogate the sensor. Action sequences can then be initiated to respond to and act upon the data received from the environment.

Previous research has mainly focused on domestic energy monitoring systems with limited studies focusing on non-domestic buildings [2]. This paper builds on and extends previous research by:

- Designing a generic monitoring framework to capture and analyse heterogeneous data.

- Capturing data from a wide range of domestic and non-domestic properties.

- Testing the framework by collecting data from three European cities.

The contributions of the paper include:

- The design and implementation of a flexible technical architecture that enables a range of unstructured data (including data from energy monitoring devices and

---

control sensors) to be remotely collected from both public, commercial and domestic buildings.

- The development approach from installation in one domestic property through to capturing data from 255 buildings (including car parks, schools and civic buildings) in three cities in two countries across Europe.

- A generic monitoring framework that integrates pan-European data to support intelligent analysis for smart cities, enabling data from different European cities to be compared.

- The identification of challenges to consolidate environmental data from domestic and non-domestic buildings across Europe.

This paper reports on the development of a generic monitoring framework for environmental monitoring that incorporates data from both domestic homes and a wide range of non-domestic buildings in two European countries. The need for the framework is first outlined and the design of the framework is then presented in two stages. Stage 1 explains how the framework was first tested in a domestic environment and the results are discussed. In stage 2, the framework was used to capture data from both public and private non-domestic buildings in a major UK city. The scope of the data captured and stored was then extended to incorporate domestic and non-domestic buildings in three cities across two European countries. The results of the testing are discussed, considering both the technological challenges, such as scalability and the analytical challenges of determining useful information from the environmental data captured. The paper concludes by highlighting the challenges that need to be addressed in the implementation of pan-European energy and environmental monitoring frameworks.

II. BACKGROUND

Smart cities provide a technical infrastructure enabling citizens to make informed decisions [4]. The citizens of a smart city include the owners and occupants of domestic and commercial properties, educational institutions, and public buildings. Improving energy usage across a city requires the participation and inclusion of both domestic and non-domestic properties. Cities also need to learn from one another, sharing knowledge and best practice to strive for sustainable improvement.

A. Domestic Buildings

Providing consumers with information about their energy usage can help them understand how to reduce their consumption [10]. For example, one study [11] reported that energy use in the home was reduced by 10-15% when residents were given information about their energy usage.

Although a number of systems have been developed to display energy usage, domestic energy users lack clear information about energy usage on which they can base decisions to reduce their consumption [12], and to establish sustainable routine practices. One of the limitations of existing systems is that they display energy usage data but do not provide sufficient information for home owners to interpret, evaluate,

and act on the data. Real-time information is becoming the expected level of data access and use demanded by citizens [13].

In addition to seeing their own real-time energy usage, householders have suggested that they would like to see which devices were consuming the energy so that they could alter their behaviour [14]. The ability to see the status of specific devices (on/off/standby) would enable residents to make decisions about specific actions to be taken [14] in response to real-time consumption data. However, research by [15] discovered that home owners left more items on stand-by as a result of seeing how little energy this actually consumed.

The ability of social norms to influence energy conservation behaviour has been widely discussed [for example, 4]. A study by [14] reported that occupants wanted to compare and benchmark their energy consumption with their neighbours and engage in competition to reduce energy consumption. This included information at country, regional and neighbourhood level, using similar demographics and household type [14]. In contrast, [10] reported that participants did not think it would be helpful to compare energy consumption data with other households due to differences in appliances and lifestyles. Sharing energy consumption data between households can lead to concerns about privacy as detailed information about daily routines could be inferred from shared data. Privacy concerns could be addressed by aggregating consumption data so that neighbours could only see total daily consumption data [14]. End-user programming is also needed to control devices and empower home owners, addressing concerns of devices controlling their lives [16] and giving occupants control of the type and granularity of data that can be shared.

According to [10] energy display monitors should be based on a three step model of behavioural change: raise awareness, inform complex changes and maintain sustainable routines. In the first step of raising awareness, detailed information on energy usage patterns should be provided to consumers together with a feedback system to reflect how changes in behaviour have impacted energy consumption. When consumers are familiar with their energy consumption, the second step involves setting goals to reduce consumption. Limited work has been done about how to set goals and provide feedback [10], which requires consolidation of information in order to inform goals. Finally, in order to maintain sustainable routines, systems need to evolve with the home environment and consumers' changing needs [10].

Consumers need to understand the baseline energy consumption in their home [10]. Research by [14] found that when people moved to a new home they had to learn the average consumption usage of the new home. The ability to establish baseline consumption levels and understand consumption patterns in a new home is more difficult and takes more time when occupants move to a new city or a new country. Energy performance is an important factor affecting property sales [17]. Information about energy usage from different geographical areas would therefore be useful to inform relocation decisions. This information would also help occupants to understand energy usage and set usage expectations in a new home more quickly.

In a smart city, occupants need the ability to access a range of information to help them make informed decisions about their energy usage. This includes data relating to:

- Their own energy consumption in order to understand their own energy usage patterns.

- The energy consumption of similar properties with similar household structures in the area or similar areas in order to benchmark their own usage.

- The energy consumption of properties in other areas and in different countries to inform relocation decisions.

B. Public and Commercial Buildings

The EU Directive requires that new buildings occupied and owned by public authorities should be nearly zero-energy buildings after 31 December 2018, and that by 31 December 2020 all new buildings are to be nearly zero-energy [18]. However, it is estimated that 60% of the buildings that have already been built will still be standing in 2050 [5]. Analysing energy consumption in existing buildings in order to inform behavioural change is therefore important. Furthermore, despite the pressure to build energy efficient buildings, there is a difference between predicted energy usage and actual energy usage in buildings [19], [20]. Two main factors contribute to the potential discrepancy between prediction and actual energy usage in a building. First, performance targets change during a building project as conflicts arise and compromises are agreed [21]. Second, energy consumption predictions are often based on incomplete data simulations. This has led, for example, to some green buildings consuming more energy than similar conventional buildings [22]. Energy certification occurs using regulated-demand simulations pre-occupancy; unregulated demand (that is, plug load from electrical equipment) is not considered and therefore simulations are not intended to predict total building consumption [23].

Electrical demand in buildings is expected to increase as the number of appliances continues to increase, reducing energy consumption in non-domestic buildings is therefore urgently needed [5]. A study by [20] identified differences in energy saving behaviour between occupants in green buildings and traditional buildings. For example, more occupants in green buildings reduced the number of multiple screen monitors in use and worked on laptops rather than desktop computers to reduce energy; however, more occupants in traditional buildings reduced energy consumption by shutting computers down or putting screens to sleep rather than using screen savers [20]. Post-occupancy evaluation is needed to feed-forward into building design, feedback to improve efficiency, and benchmark to measure progress towards sustainability targets [24].

Building information management is often poor, resulting in fragmented, redundant, and overlapping information [25]. Energy usage data in buildings is often held by the building occupants or their supply companies, which hinders the ability to analyse demand patterns and efficiently manage supply distribution [5]. Existing information systems lack detail about for example, the type of energy used and the unique nature of the building [26], which hinders the ability to support benchmarking between similar buildings, within industries, and across cities and countries. The potential impact of energy consumption data to initiate behavioural change is less understood in non-domestic buildings because of a lack of sub-metering [2]; sub-hourly analysis is needed [27].

Building energy management systems (BEMS) are centralized control systems that monitor building performance, integrating mechanical, electrical and plumbing (MEP) controls, in order to satisfy the comfort and safety needs of occupants [1]. The performance of BEMS is central to energy consumption in buildings [28], however, optimizing heating, ventilation and air conditioning systems (HVACS) to maximize the comfort for occupants can conflict with the requirements to minimize energy consumption [1].

Most energy management systems do not enable occupant participation [1] as they are based on "mechanical logic" [28]. Smart buildings remove decisions and remove control from occupants, making occupants passive and ill-informed about energy use, however, humans are a critical component in resource consumption decisions [29]. A graphical interface is therefore needed [28] to provide a socio-technical information feedback system to monitor and decrease building energy consumption [29].

Archived data from BEMS provide a useful resource to improve the energy efficiency of a building [21]; however, existing building management systems often lack the ability to integrate monitoring data and to generate actionable information [28]. BEMS can be used to support the automatic control of individual devices in the Internet of Things (IoT) to create smart buildings [5]. This requires sensors to be connected to BEMS to measure, for example, occupancy, temperature and lighting. Previously, cost has been a prohibiting factor, however, the availability of low-cost sensors and alternative ways to capture data (such as through mobile devices) has made the vision of a smart building more achievable [5].

A three phase philosophy is outlined by [28] to reduce energy use in buildings: first, install sensors to monitor and control the environment (such as energy consumption, temperature, motion and brightness sensors); second, monitor real-time energy data (requiring data collection, statistical analysis and data storage); third, implement optimization (using control scenarios and end-user software).

The ability to process, analyse, and manage data from sensors remains a significant challenge [5]. A large-scale platform independent, diverse-application IoT infrastructure is needed to process, manage, and analyse data in order to enhance decision making to form smart cities [30]. The following section outlines the development of a generic monitoring framework to process, analyse and compare data from domestic and non-domestic buildings across Europe. Improving building information management will facilitate improvements in energy performance by informing decisions about behavioural change of occupants, building maintenance, and future building design.

III. Generic Monitoring Framework

A proliferation of hardware systems is available in the UK domestic market to enable home owners to monitor aspects of their environment. Examples include: X10[2], Z-Wave[3] and KNX[4]. Data are supplied by disparate sensing equipment, which differ in terms of the manufacturer, the format used, and the values returned by the equipment. The problem is how to provide a unified interface through which the end user can access the data. A framework is needed to support interoperability between the different sensing and monitoring devices available from different manufacturers [2]. In addition, sensor network management systems require a low communication overhead, efficient search methods, and long-term storage [9]. A number of solutions require significant rewiring of premises and although this is feasible during construction, to retrofit can become costly. This research concentrates on using readily available, low-cost sensors, which do not require significant cabling in order to function effectively.
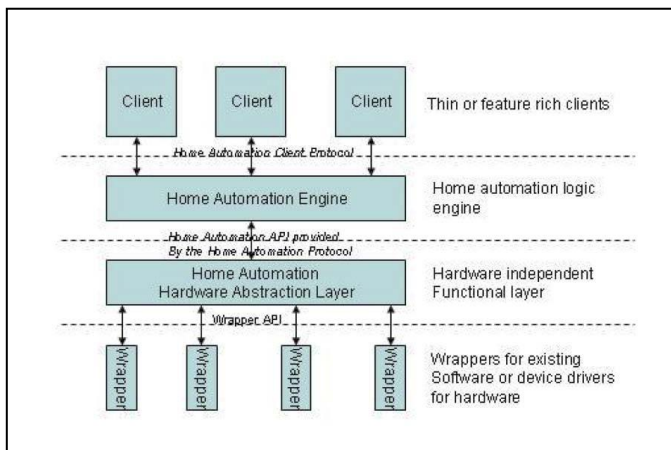

Fig. 1.   Generic monitoring framework [31].

A generalized framework for data collection is needed [30] [1]. In earlier work [31] proposed a layered approach to address the problem of integrating disparate systems and provide a common user interface. The layers, shown in Fig. 1, each perform an independent and well-defined function supporting flexibility. This layered approach separates the functionality into discrete layers.

Fig. 1 shows that the generic monitoring framework adopts an approach of creating a wrapper to manage the interface between the sensing device and the client. This creates a loosely coupled, language independent ubiquitous platform, using well established protocols for data transfer. A key feature of this approach is that the sensing devices are not restricted to being connected to a computer. The following sections discuss the implementation of this generic monitoring framework to capture energy usage data from a range of domestic and non-domestic properties across Europe.

IV. Stage 1: Domestic Environment

The generic monitoring framework (GMF) in Fig. 1 was implemented with a simple trial set up using readily available commercial, off-the-shelf sensors in a domestic environment. This provided the opportunity to test the feasibility of the concept. Initially a single house was equipped with a variety of sensors, which returned data to the GMF.

A. Sensors in Domestic Environment

A range of sensors were all connected to a single computer running the Linux operating system. This was chosen because of the flexibility it provides in accessing the hardware and also because of cost. The version of Linux chosen was Fedora, which is regularly updated allowing access to modern hardware and is also free. The sensors included an electricity monitor, temperature sensors, barometric pressure sensor, and X10 controlled devices. Sensors need to be uniquely identifiable with their location and functionality without affecting scalability, network performance or device functionality using IPv6 [30].

1) Electricity Monitor

The electricity monitor chosen was a commercial unit supplied by CurrentCost (www.currentcost.com). This device uses a current clamp on the main live feed into the premises enabling the electricity consumption of the whole house to be monitored. The clamp is connected to a sending unit, which uses wireless technology on 433MHz to transmit the data to the display unit.

The display unit is designed to show the current consumption along with other information such as cost and historical data. The advantage of this unit is that it has a serial connection allowing the display to be connected to a computer. The unit is supplied with a USB cable and software for Microsoft Windows Operating System™. This unit transmits its data through the serial port every 6 seconds as an Extensible Markup Language (XML) stream of data.

The display unit contains a temperature sensor so it was also possible to obtain the temperature of the room where the display unit was located. Software was written in Perl to listen to the serial data and identify the relevant items in the XML data stream. These data were then sent to the GMF. Perl was chosen as the data returned was XML as a text data stream. Perl was designed for text extraction and it seemed logical to use this language to extract the data from the text stream.

a) Temperature Sensors

Along with the CurrentCost device being able to supply the temperature of the location of the display unit, a number of other temperature sensors were also deployed. These were Arexx[5] sensors. They consist of a temperature sensor unit and a receiving base station. The sensor units relay information to the base unit using RF at 433MHz. The base unit connects to a computer using USB. The protocol used for communication between the computer and the base unit is proprietary. Open source software to support these sensors was obtained from the

Internet [32] and modified to allow the data obtained to be sent to the GMF. This software was written in C.

2) One Wire Devices

External temperature was obtained using a one wire sensor. These sensors are designed to be connected using one wire and are powered via that wire. Consequently the wiring for these devices can be achieved unobtrusively. The wire connects to the computer using the USB interface via a converter. There is a variety of software available to access the devices since the interface to the computer is via USB. Since the devices were connected to a computer running Linux, software was chosen that allows the one wire sensors to appear as part of the file system[6] and consequently the software needed to get data from these sensors simply requires to access a file. Since this now was effectively a text manipulation problem, the software was written in PERL to extract the data from the relevant files and send the data to the GMF.

3) X10 Devices

X10 devices allow control of individual electrical appliances. These devices differ from the sensors described above in that they are two way. They not only provide information but can also use information to control the state of the connected device. There are a number of different devices allowing control of a wide variety of appliances. The implementation here used X10 to control both internal and external lighting as well as a number of electrical items in the garden of the property.

There already exists a piece of software for X10 under Linux called Heyu[7], which allows control of X10 devices and the creation of macros to respond to various events happening. This software provides a command line interface. Although the protocol to interface to the X10 control unit is widely available, the protocol is complex and implementing new software could be prone to errors. It was decided to utilize this existing software and to create a wrapper around it to provide an interface to the GMF. This wrapper was written in PERL due to its text manipulation ability.

B. GMF Implementation in Domestic Environment

The GMF needed to be able to provide a number of functions both for the collection of data and also for the dissemination of that data. It was decided initially to create the GMF as a server application running under Linux. The initial implementation was done using PERL. The protocol for communication with the GMF was to be TCP sockets and port 2345 was arbitrarily chosen as the port for communication.

The server was created using a request mechanism where the client would send a request on the TCP port. For example to set a sensor value, the client would send S: <sensor id> <value> and to read a sensor value a client would send G: <sensor id>. The server would respond to this with <sensor value> <last update time>.

As the initial implementation was such that both clients and server were on the same machine, no consideration was given to allowing the client to pass the update time to the GMF. Instead the GMF would store the value of the sensor supplied along with the time from the local server. This time was stored as a Unix epoch time (i.e. number of seconds since midnight 01-01-1970).

A number of experiments were conducted looking at whether clients should send values when they have them or only when the value changed. Trying to implement a client to show meaningful data quickly, revealed that the best solution was to store all the data and then filter it rather than making the sensor clients perform the filtering. This is in line with the layered approach adopted where the processing is done in the relevant layer.

The underlying database used for the GMF was MySQL[8]. This was chosen due to its reliability, performance, and being an established industry standard. The GMF consequently took the data from the requests coming in and used SQL queries to store the data in the database.

C. Presentation Clients

Once data is in the GMF a means to retrieve the data is needed in order to be able to provide useful information. Again the presentation clients would interact with the GMF using TCP sockets on port 2345. A number of different clients were developed to demonstrate the flexibility and interoperability of the GMF. In order to show the data on a web site, a simple web page was created using PHP to display the last recorded values of a number of the sensors. These can be seen at http://host.while.org.uk/stats. A simple Windows-based client was produced in C# which displays a list of available sensors and the user can then select a sensor to see the latest values or a graph of the recent data (see Fig. 2).
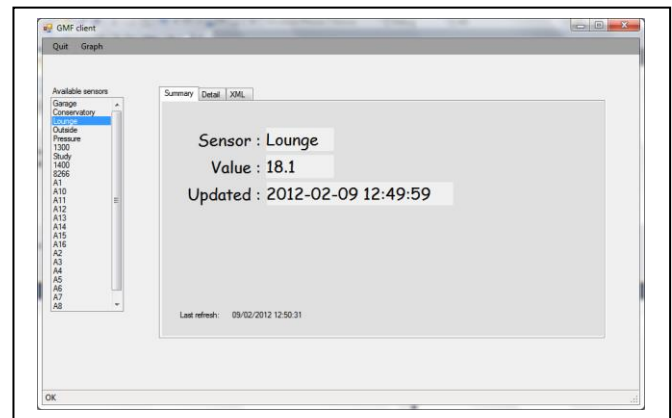

Fig. 2.   Windows client

---

6   http://owfs.org
7   http://heyu.org
8   http://www.mysql.com

Data visualization is important to help users interpret sensory data [30]. A graph of the data was prepared to help show the data in a meaningful way. A webpage already existed showing data on the current operation of the server, e.g. load, network traffic, etc. It was decided that this webpage would be a suitable host to also display sensor data. This page used rrdtool[9] an open source tool for holding time related data and producing graphs from this. A number of simple clients were created in PERL, which queried the GMF and presented the data in a form that rrdtool could use.
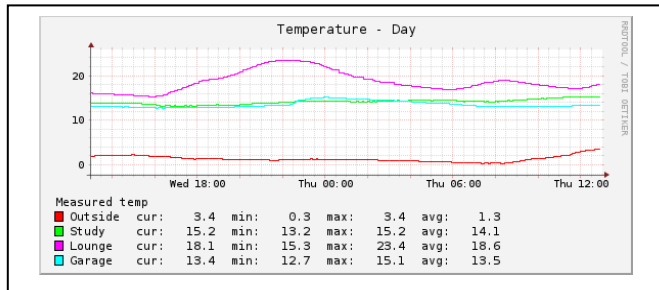

Fig. 3. Temperature data.

Fig. 3 demonstrates this using a graph showing 24 hours of data from four temperature sensors. The outside temperature is monitored using a one wire device, the study is the temperature reading from the electricity monitor display, and the other two are Arexx temperature sensors.

D. Challenges Identified from Stage 1 Pilot of GMF

A number of problems were identified during the pilot relating to data context, data volumes, and timestamps.

1) Data Context

Data from a sensor typically returns a value, such as a temperature, and the time when the data was captured. Additional data needs to be captured about the context in which the data were captured in order to make sense of the data and create meaningful information. In the pilot, sensors were given a unique identifier and a look-up table was used to identify the location of the sensing device (e.g. room 1).

Fig. 3 shows temperature data captured from rooms over a period of time. The graph shows a clear rise in the temperature, however, further data need to be captured from other sensing devices in order to understand what factors caused the sudden significant change in temperature.

The pilot demonstrated that a wide range of contextual factors needs to be captured, analysed, and interpreted to fully understand the context of the data reported within the GMF. In addition, data from a variety of sensors captured at the same time period need to be interpreted collectively to gain meaning from the data collected. This requires preprocessing of the data due to multivariate time series data [30].

2) Data Volumes

The pilot demonstrated that the data volumes required to monitor an environment over a period of time quickly become extensive. For the limited number of sensors in the pilot, the

database was around 20MBytes per day. This has implications for the capture, storage and analysis of the data collected. The pilot showed that the throughput of the data was highly dependent upon the capabilities of the hardware and communication network. Further consideration therefore needs to be given as to where data should be aggregated. The high data volumes also demand that attention is given to the expiry of data [30]. Historical data have an important role in analysis, prediction, simulation, and decision making, therefore data retention periods need careful consideration [33].

3) Timestamp Results

In the pilot, the sensors used the time from the server as the timestamp for the data values captured. If data are to be later compared with data captured in different time zones around the world then further consideration needs to be given to whether universal time should be used or location information needs to be obtained.

V. STAGE 2: EXTENDING THE GMF IMPLEMENTATION

This section discusses how the problems identified from the initial pilot were addressed and explains the changes made to the GMF to capture data from multiple buildings. Fig. 4 shows the process of data collection in the GMF.
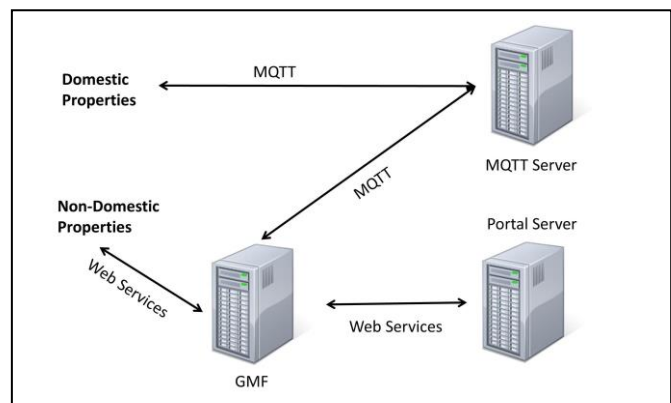

Fig. 4. Data collection.

A. Scaling the Technical Infrastructure

The initial implementation adopted a server-based approach to the design and implementation of the GMF. This led to three main problems. First, a new protocol had to be developed to allow the clients to talk to the server to provide and request information. This meant that any client wishing to make use of the GMF needed to be written specifically due to the proprietary protocol. Second, making the GMF a server required that this software was running all the time. For new software this can lead to problems of reliability due to problems such as memory leaks. Third, using TCP sockets on port 2345 meant that clients requesting to talk to the server had to be able to communicate on port 2345. As port 2345 was unused by well-known applications, it was unlikely that firewalls would be configured to allow traffic on port 2345 through.

Taking the above problems into account discussion led to the question as to whether web services could be utilized to

---

[9]   http://oss.oetiker.ch/rrdtool

overcome some or all of these problems. Clearly the protocol for web services is well defined [34] and the problem of using a separate port can be overcome by using web services on port 80 (the standard http port).

In order to ascertain whether development of the GMF using web services should be pursued, a simple web service was developed on the existing GMF. This web service returned the last sensor value and update time for a supplied sensor id. Given that the existing GMF was running under Linux it was decided to utilize the Apache web server as that was already configured and running on the server. The web service was developed in C# using Visual Studio and this led to a problem of deployment under Linux. Fortunately there is an open source solution in the form of mono[10]. It is possible to configure Apache to use mod_mono, which allows asp.Net applications to be run under Linux. This was configured and the web service deployed. The web service directly accesses the MySQL database using SQL to retrieve the data and supply the data to the client.

Initial results would seem to indicate that this method overcomes the problems identified above with no discernable performance issues.

B. Implementation of sensors

As the implementation of version 1 of the GMF concentrated on a single property all sensors were connected to the server and relevant software written to obtain the data. Clearly this was not feasible for the next implementation where multiple properties were being used.

The server was developed using web services to provide relevant functions that could be used by clients.

It was considered that the most appropriate devices for collecting energy data was the CurrentCost device used in the previous trial. Therefore some mechanism to retrieve the data from this device and call the appropriate web service was required.

It was not possible to rely on the owner of the property having access to a computer so it was decided to provide a complete solution for the user. This required a cheap computer based system.

A TP-Link TP-300 router was selected as it was possible to reprogram this using an open source Linux based system allowing for development of software to provide the relevant functionality.

The user was provided with a CurrentCost device, A TP-Link router suitably programmed and a set of instructions which simply required them to connect the energy monitor and plug the TP-Link into their broadband router.

The transmission of the data used the MQTT[11] protocol (Fig 4) as this was designed for transmission of small amounts of data. The software on the TP-Link would get the data from the CurrentCost device and post it on an MQTT channel having connected to an MQTT server. Within the server hosting the GMF there was a client that listened on the appropriate MQTT

channel for messages and took the data and called the appropriate web services to add the data to the GMF.

C. Scaling the Data Capture from Domestic Properties

Extending the use of the GMF to capture data from multiple domestic properties in different countries required more detailed consideration to be given to the contextual information that needed to be captured with the sensor data in order to provide actionable information. Residential occupants required the ability to:

- Securely log-into the GMF dashboard.
- Select known devices from a database.
- Add various types of devices to the GMF.
- Configure devices for a required data set.
- Set the access levels to the data and the granularity of the data.
- View their own data.
- Compare their data with data from similar properties.

Data therefore needed to be captured about property size, property type, occupancy level, country code, post code (though only partial post codes were recorded to address privacy concerns), year of build, number of rooms, and energy type.

Energy data was collected from 77 domestic buildings including: one from Belgium, two from Romania and 74 from the West Midlands, Cheshire and Greater Manchester areas of the UK. The range of properties included: terraced houses, detached houses, semi-detached houses, houses converted into flats, apartments, and a converted barn. The properties differed in the number of rooms, occupancy level, and energy source; for example, some homes only used one type of energy whilst other used multiple sources.

As the number of properties introduced into the GMF increased, the complexity of the data context also increased. For example:

- The GMF stored the year a property was built but some occupants were unsure of the year a property was built. Six categories were therefore used for build year (pre 1900, 1900-1939, 1940-1954, 1955-1972, 1973-1999, 2000 onwards). One property was an 18th Century barn converted in 1980 and renovated 2011, which challenged the significance of build year.
- Energy type was ambiguous because many properties incorporate multiple energy sources, therefore a data field of primary heating type was used with the values of gas, electricity, oil, solid fuel, or other fuel.
- The registration of properties included an optional free-text description field. The data entered into this field highlighted contextual factors that were important to residents, such as "House built to the Lichfield design.

Fully double glazed with gas central heating." The free-text descriptions identify data attributes that could be incorporated in future developments of the GMF to improve the ability to compare energy consumption in like-for-like buildings.

Fig. 5 shows a screenshot from the domestic dashboard to show how data from two buildings can be compared. On the right hand side, users can enter the selection criteria to identify properties of similar type and heating source. On the left side, the graph shows two lines: the energy consumed at the property and the average consumption of similar properties during the same timeframe. The data can be accessed at different levels of granularity.
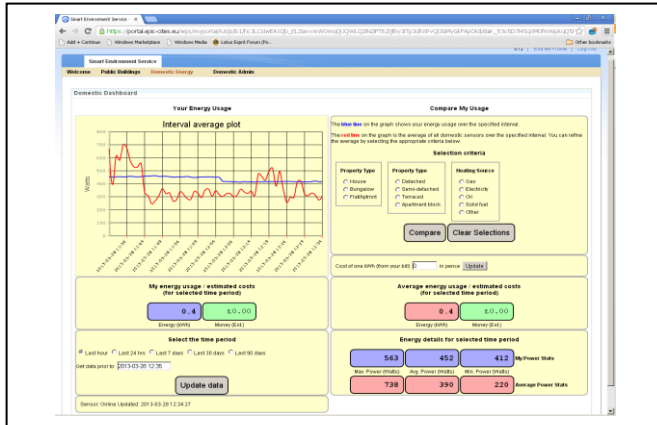

Fig. 5. Comparing data consumption.

The following section outlines the extension of the GMF to include energy consumption data from building energy management systems in non-domestic buildings.

D. Scaling the Data Capture to Include Public and Commercial Buildings

Building management systems are diverse, complex, and heterogeneous [1]. This project used data captured from existing fiscal smart meters and building management systems, rather than introducing any additional sensors into the non-domestic buildings. The GMF was used to aggregate the data and provide energy dashboards for commercial and public premises using an open source platform using commercial grade applications such as mySQL and industry standards. Web services in Java were used, as the framework was deployed on IBM websphere. Problems identified include the need to understand the context of the data, managing the large data volumes, and how to timestamp at the server or at the client.

The context of the data introduced further challenges in relation to non-domestic properties. Many factors affect energy use in buildings [20] including physical structural conditions, socio-demographic characteristics, supporting infrastructure, and cultural and economic factors [4]. Environmental monitoring of public buildings differs to that of domestic properties in that a public building often comprises a complex of buildings, which each may be subdivided into zones. It was a requirement of the project that facility managers must be able to view and analyse energy consumption data at the level of an individual building. This is because different spaces have

different uses such as office space, meeting space, and public space.

Data was captured from 175 public and commercial properties in Birmingham (UK), Manchester (UK) and Tirgu Mures (Romania). This included a wide range of properties such as: civic buildings, education buildings, leisure centre facilities, markets, multi-storey car parks, museums, office accommodation over multiple floors, public roads (including underpasses and tunnels), and waste management sites.

The introduction of each building type introduced further data attributes relevant to the specific types of spaces within each building. For example:

- Civic buildings included office accommodation and extensive event and conferencing facilities.

- Education buildings included primary, secondary and higher education institutions comprising office space, general teaching space, specialized teaching space (for example, laboratories), social spaces, and student accommodation.

- Leisure centres included swimming pools and external floodlighting.

- Markets included indoor and outdoor, public and wholesale markets.

- Multi-storey car parks included a number of levels and were measured by the number and type of parking spaces.

This highlighted the complexity and granularity of data that need to be collected in order to analyse energy usage data to make informed decisions. Energy use in buildings needs to benchmarked internally and externally [2], 25 buildings in one organization complex were therefore included in the sample to enable the organization to compare energy consumption across a building complex.

E. Implementation of non-domestic sensors

Most of the buildings used had intelligent building management systems that provided data periodically for fiscal purposes. It was possible to get this information via email and then use software to automatically process this calling the appropriate web services to add the data to the GMF.

Problems occurred in the Romanian non-domestic buildings as they did not have building management systems so another solution had to be developed.

The CurrentCost device is able to monitor three phase supplies so it was decided to modify the domestic sensor setup to monitor a three phase supply. The software had to be modified to aggregate the supply data into a single figure.

F. Challenges Identified from Stage 2 Scaling the GMF

During the project problems with the reliability of the Internet connection reduced the ability to gather large volumes of data from Romania. Other problems were identified relating to the data context and software architecture.

1) Data Quality

The metadata for each domestic building was entered by the user. The size of domestic properties in the UK is normally considered by number of bedrooms, but elsewhere in Europe properties are advertised by floor space measured in square metres. This highlighted problems with data quality as some data were missing or incomplete, limiting the ability to identify similar properties for comparison.

Data about public buildings, such as size, was not consistently entered, because building size can be measured in different ways. Further attributes therefore need to be defined for specific building types. Building information such as size, age, history, and use could be attained by accessing existing systems but the information is likely to be held in a number of different systems in varied formats. This relates to the work of [35] on primary and secondary context. The GMF captures data about the primary context of buildings, which can be used to access related data in different systems, providing the secondary context. The project has identified the need for the existing information systems in the city, such as building planning systems, to be integrated with the GMF in order to provide contextual information in which to interpret energy consumption data.

The date and time of readings were captured from the sensors. This was based on the data from the building management systems that were sent in batches. A problem arose with sensor data captured from domestic properties. The date and time had to be manually set by the occupant when the sensor was activated. This relied on occupants setting the date and time correctly. The data collected includes sensor readings that include, for example, the wrong year. This therefore adversely affects the quality of the data and limits the ability to use the data for both benchmarking between buildings or informing decisions about energy consumption in the same building. The purpose of this project was to test the feasibility of the software architecture therefore data quality did not affect the results attained, but would need to be addressed before analysing the data captured using the GMF.

2) Software Architecture

Apart from the problems identified above relating to data quality and context, there are also some software architecture problems to be resolved. If, as it seems likely, future development of the GMF will be around web services, then there will need to be some investigation into whether the web services should directly access the database using SQL or whether there should be an intermediate database access layer to be consistent with the layered approach, a database access layer should be used. This would clearly provide more flexibility allowing for a range of underlying databases to be utilized without having to modify the web services. Further consideration also needs to be given to the aggregation of data over a certain age, in order to improve performance of queries as the volume of data in the GMF increases.

3) Hardware Architecture

The implementation of the GMF and associated clients was undertaken before the widespread availability of small low cost computer systems such as the Raspberry Pi. It is likely that future work on the clients would make use of these devices as the development of the client software would be made easier.

VI. CONCLUSION

A technical architecture, referred to as the generic monitoring framework (GMF) has been developed to remotely capture and integrate energy consumption data from domestic properties, public and commercial buildings across three European cities. This has involved addressing challenges relating to the capture of conflicting data formats from heterogeneous devices, large volumes of data transmissions and accommodating existing facility management systems. Information about energy consumption has an important role in reducing energy consumption, enabling residents to make decisions about their use of energy, however, further research is needed to understand how people use energy. The system enables homeowners and facility managers to interrogate their own data usage and compare their energy consumption with comparable properties in Europe. This investigation into the feasibility of the suggested framework in [31] shows that technically, the proposed software architecture for a generic framework for environmental monitoring of smart cities is possible. However there are a number of technical and data quality challenges that require further investigation.

The GMF provides a web-based infrastructure supporting a data-driven approach to enable cities to benchmark energy consumption and establish best practice to strive for sustainable improvement in energy consumption. It has been demonstrated that the proposed architecture allows a variety of disparate sensors with different protocols to provide data to clients so that a unified view of the data can be seen. It has also demonstrated that the architecture is independent of programming language with clients being developed in PERL, PHP, C#, and C under both Linux and Windows. The GMF has captured data from 255 buildings in two countries and has identified challenges that need to be addressed to improve the quality and reliability of data captured from domestic and non-domestic properties. Further analysis of the data is being undertaken to explore how the data can be used to inform actions and change energy consumption behaviour. This requires further work in contextual inquiry to identify the additional data needed to interpret energy usage data. Systems within a smart city then need to be integrated to enable secondary context data to be accessed to provide actionable information to inform decision-making in a smart city.

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," *Renew. Sust. Energ. Rev.*, vol. 34, pp. 409–429, June 2014.

[2] M. W. Ahmad, M. Mourshed, D. Mundow, M. Sisinni, and Y. Rezgui, "Building energy metering and environmental monitoring – A state-of-the-art review and directions for future research," *Energy and Buildings*, vol. 120, pp. 85–102, May 2016.

[3] N. Kominos, *Intelligent Cities and Globalisation of Innovation Networks*. Abingdon, UK: Routledge, 2008.

[4] M. S. Khan, M. S. Islam, and H. Deng, "Design of a reconfigurable RFID sensing tag as a generic sensing platform toward the future internet of things," *IEEE Internet Things J.*, vol. 1, no. 4, pp. 300–310, Aug. 2014.

[5] P. Kumar, C. Martani, L. Morawska, L. Norford, R. Choudhary, M. Bell, and M. Leach, "Indoor air quality and energy management through real-time sensing in commercial buildings," *Energy and Buildings*, vol. 111, pp. 145–153, Jan. 2016.

[6] R. Meier, A. Harrington, K. Bechmann, and V. Cahill, "A framework for incremental construction of real global smart space applications," *Pervasive Mob. Comput.,* vol. 5, no. 4, pp. 350-368, 2009.

[7] U. Ramachandran, R. Kumar, M. Wolenetz, B. Cooper, B. Agarwalla, J. Shin, P. Hutto, and A. Paul, "Dynamic data fusion for future sensor networks," *ACM Trans. Sens. Netw.,* vol. 2, no. 3, pp. 404-443, 2006.

[8] D. Sathan, A. Meetoo, and R. K. Subramaniam, "Context aware lightweight energy efficient framework," *Int. J. Human Soc. Sci.,* vol. 5, no. 4, pp. 24-255, 2010.

[9] D. Ganesan, B. Greenstein, D. Estrin, J. Heidemann, and R. Govindan, "Multiresolution storage and search in sensor networks," *ACM T. Storage,* vol. 1, no. 3, pp. 277-315, 2005.

[10] Y. Riche, J. Dodge, and R. A. Metoyer, "Studying always-on electricity feedback in the home," in *Proc. SIGCHI Conf. Hum. Factor Comput. Syst.*, 2010, pp. 1995–1998.

[11] J. Froehlich, "Promoting energy efficient behaviors in the home through feedback: The role of human-computer interaction," in *Proc. HCIC Workshop*, 2009, pp. 1–11.

[12] J. Suppers, and M. Apperley, "Developing useful visualizations of domestic energy usage," in *Proc. 7th Int. Sym. Visual Inf. Comm. Inter.,* pp. 139–148.

[13] F. Scorza, "Smart Monitoring System for Energy Performance in Public Building," in *ICCA 2015*, *Lect. Notes Comput. Sc.,* vol. 9156, pp. 767–774.

[14] M. Chetty, D. Tran, and E. R. Grinter, "Getting to green: understanding resource consumption in the home," in *Proc. 10th Int. Conf. Ubi. Comp.,* 2008, pp. 242-251.

[15] L. Liikkanen, "Extreme-user approach and the design of energy feedback systems," in *Int. Conf. Energ. Effic. Dom. Appli. Light.,* 2009, pp. 16-18.

[16] S. Davidoff, M. K. Lee, C. Yiu, J. Zimmerman, and A. K. Dey, "Principles of smart home control," in: *UbiComp 2006, Lect. Notes Comput. Sc.,* vol. 4206, 2006, pp. 19-34.

[17] O. M. Jensen, A. R. Hansen, and J. Kragh, "Market response to the public display of energy performance rating at property sales," *Energy Policy*, vol. 93, pp. 229–235, Jun. 2016.

[18] European Union, "Directive 2010/31/EU of the European Parliament and of the Council on the Energy Performance of Buildings," Brussels, Belgium, 2010.

[19] J. Pan, R. Jain, S. Paul, T. Vu, A. Saifullah, and M. Sha, "An Internet of Things framework for smart energy in buildings: designs, prototype, and experiments", *IEEE Internet Things J.*, vol. 2, no. 6, pp. 527–537, 2015.

[20] N. S. Mokhtar Azizi, S. Wilkinson, and E. Fassman, "Strategies for improving energy saving behaviour in commercial buildings in Malaysia," *Engineering, Construction and Architectural Management,* vol. 22, no. 1, pp. 73–90, Jan. 2015.

[21] D. T. J. O'Sullivan, M. M. Keane, D. Kelliher, and R. J. Hitchcock, "Improving building operation by tracking performance metrics throughout the building lifecycle (BLC)," *Energy and Buildings*, vol. 36, no. 11, pp. 1075–1090, Nov. 2004.

[22] G. R. Newsham, S. Mancini, and B. J. Birt, '"Do LEED-certified buildings save energy? Yes, but…," *Energy and Buildings*, vol. 41, no. 8, pp. 897–905, Aug. 2009.

[23] J. Gabe, "Market implications of operational performance variability in certified green buildings," presented at the 17th Annual Pacific Rim Real Estate Society Conference, Gold Cost, 16-19 Jan., 2011.

[24] I. Cooper, "Post-occupancy evaluation - where are you?," *Building Research & Information*, vol. 29, no. 2, pp. 158–163, Mar. 2001.

[25] I. G. Dino, and R. Stouffs, "Evaluation of reference modeling for building performance assessment," *Automat. Constr.*, vol. 40, pp. 44–59, Apr. 2014.

[26] J. O'Donnell, E. Corry, S. Hasan, M. Keane, and E. Curry, "Building performance optimization using cross-domain scenario modeling, linked data, and complex event processing," *Build. Environ.*, vol. 62, pp. 102–111, Apr. 2013.

[27] J. Salom, J. Widén, J. Candanedo, and K.B. Lindberg, "Analysis of grid interaction indicators in net zero-energy buildings with sub-hourly collected data," *Adv. Build. Energy Res.* vol. 9, no. 1, pp. 89-106, 2015.

[28] V. Marinakis, C. Karakosta, H. Doukas, S. Androulaki, and J. Psarras, "A building automation and control tool for remote and real time monitoring of energy consumption," *Sustainable Cities and Society*, vol. 6, pp. 11–15, Feb. 2013.

[29] E. J. Petersen, V. Shunturov, and K. Janda, "Dormitory residents reduce electricity consumption when exposed to real-time visual feedback and incentives," *Int. J. Sustainability in HE,* vol. 8, no. 1, pp. 16–33, 2007.

[30] J. Jin, J. Gubbi, S. Marusic, and M. Palaniswami, "An information framework for creating a smart city through Internet of Things," *IEEE Internet Things J.*, vol. 1, no. 2, pp. 112–121, 2014.

[31] D. While, J. Krasniewicz, and S. Cox, "The development of a unified API for smart home systems," in *The Internet of Things and Services, 1st International Research Workshop co-located with EuroTRUSTAmi 2008 Conf.,* 18-19 September 2008, Sophia-Antipolis, French Riviera, 2008.

[32] Algorithm-Forge.com, "Linux USB driver for the Arexx TL-500 –part II," [Online]. http://algorithm-forge.com/techblog/2010/01/linux-usb-driver-for-the-arexx-tl-500-part-ii/, Accessed on: Jul. 1, 2016.

[33] S. Cox, *Managing Information in Organizations: A Practical Guide to Implementing an Information Management Strategy.* UK: Palgrave, 2004, pp. 125-126.

[34] T. Berners-Lee, "Web Services," [Online]. http://www.w3.org/DesignIssues/WebServices.html, Accessed on: Jul. 1, 2016.

[35] *A. K.* Dey, and G. D. Abowd, "Towards a better understanding of context and context-Awareness," in *Proc. CHI 2000 Workshop on The What, Who, Where, When, and How of Context-Awareness*, The Hague, Netherlands, Apr. 2000.