ECAI 2014
T. Schaub et al. (Eds.)

doi:10.3233/978-1-61499-419-0-1211

Extending Semantic Sensor Networks for Automatically Tackling Smart Building Problems

Joern Ploennigs and **Anika Schumann** and **Freddy Lecue**[1]

Abstract. Sensor systems are constantly growing in all application areas and become elements of our environment. Semantic Sensor Networks (SSN) support this development and provide standardized semantic access for reasoning on this information. Unfortunately they do not model internal system knowledge or simple correlations between sensors and hence they cannot be used to automatically perform analytics tasks based on sensor data only. We show how SSN ontology can be extended and demonstrate its benefits for the task of diagnosing smart building problems using real-world data.

1 Introduction

With the development of embedded cyber physical systems and large computation resources in the cloud, the availability of sensor information continuously increases. Semantic Sensor Networks (SSN) play an important role in this development as they provide a homogeneous semantic layer for sensor information and simplify the detection and retrieval of data [1]. This in turn is key for automating high level analytics tasks that are based on these data. However, many analytics tasks require system information apart from sensor data such as cause-effect relationships between sensors [2] for diagnosis. Adding these information manually can turn into a very tedious task and prevent the use of SSN for large scale analytic applications.

Building automation systems are one established example of large scale sensor and control networks that can contain thousands of devices. Analyzing the data allows to improve the building's energy consumption with large environmental impact as buildings consume about 40 % of the energy in industrialized countries [3]. This could be significantly reduced if malfunctioning equipment could be identified quickly. Approaches that tackle this problem [2, 4] need to be configured manually since they need to correlate sensor information. Thus, they require not just data as input but also information about the physical processes in buildings.

This paper shows that few extensions to the SSN ontology are sufficient for automatically deriving complex physical process models and thus for automating the diagnosis task in buildings. Our approach is solely based on semantics techniques such as SPARQL update. Starting from our extended SSN description we first derive the physical process model to capture the correlations among the sensors. This allows us to derive diagnosis rules, i. e. the cause-effect relationships of sensor observations, and finally to obtain the diagnosis results.

2 Extending the SSN Ontology

As basis for our ontology we use the Semantic Sensor Network (SSN) skeleton ontology defined by the W3C incubator group [1].

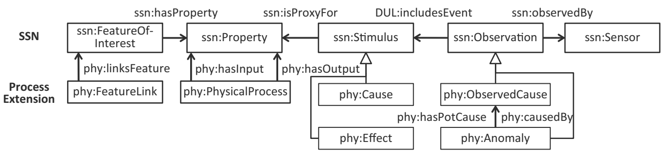

Figure 1. Extension to the SSN ontology

It provides basic concepts to model a *ssn:Sensor* that observes a *ssn:Observation* of a *ssn:Stimulus* occurring in the environment of a specific *ssn:Property* of a *ssn:FeatureOfInterest*. We extend the SSN ontology using the name space *phy* for generic concepts to model physical relationships for enabling diagnosis. Figure 1 illustrates the extensions, which are defined as follows:[2]

- $phy{:}PhysicalProcess$: The properties of many real world features are related by physical processes. A $phy{:}PhysicalProcess$ models this as directed relationship of a source property ($phy{:}hasInput$) that influences a target property ($phy{:}hasOutput$). We differentiate positive and negative correlation processes. The influence is a $phy{:}PosCorrProc$ if a factor increases with its influence and it is a $phy{:}NegCorrProc$ if it decreases with an increase of the influence. The temperature we feel, for example, is influenced by the inner energy in a room via a positive correlation process, as it feels warmer when we increase the energy. The cooling system uses a negative correlation process as it removes heat from this energy.
- $phy{:}FeatureLink$: Physical processes occur not only between properties within the same feature, but also between related features. A FeatureLink denotes such a relationship between features that are defined by $phy{:}linksFeature$ object properties. They are, for instance, used to model spatial relationships between two adjacent rooms connected by a wall.
- $phy{:}Cause$, $phy{:}Effect$: are subconcepts of $ssn{:}Stimulus$ and describe the not necessarily observed stimulus of a cause and the resulting effect.
- $phy{:}Anomaly$: is a subconcept of $ssn{:}Observation$ that is used to describe abnormal observations that should be diagnosed. An anomaly may be for example a high room temperature.
- $phy{:}ObservedCause$: is another subconcept of $ssn{:}Observation$ describing the observable discrete states of potential causes of an anomaly. A cause for a high temperature in a room might be an inactive cooling system.

3 Approach

We first create automatically a process model describing the physical relationships between sensors in a given SSN consisting of sensors,

[1] IBM Research, Ireland, email: {Forename.Surname}@ie.ibm.com

[2] Examples: *https://www.dropbox.com/s/z369tzmn00f1jv9/demopackage.zip*

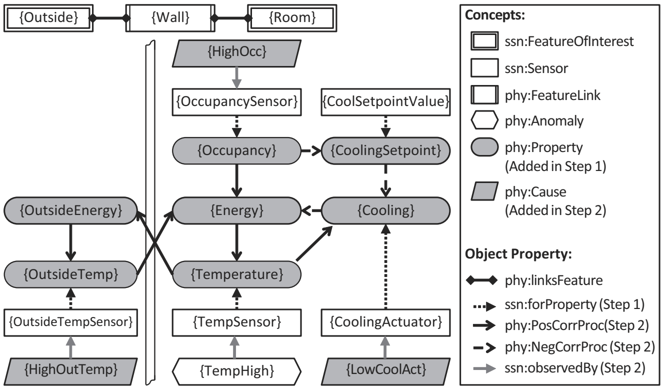

Figure 2. Process relationships for a example room and outdoor.

features, and feature links. Defining such relationships manually is not feasible for large ontologies with thousands of sensors. Therefore, we define such potential relationships once on conceptual level in a domain ontology using annotations. We then apply this knowledge using a series of SPARQL 1.1 update (SPARUL) queries to a specific building SSN. First, we create all missing property instances for the monitored feature instances. We differentiate: (i) mandatory properties that are characteristics for a feature and are always created; and (ii) optional properties that are only created if they are observed by a sensor. Second, we connect these property instances by physical processes within features and also between features connected by feature links.

We illustrate this for a single room in Fig. 2 connected to the outside. The white elements are existing instances in the SSN. The first step adds the properties in gray, such as the mandatory properties temperature and energy and the optional property cooling that is not added to the outside since there is no cooling actuator sensor. We then add the physical processes for the properties in the room and the linked outside. They are indicated by solid and dashed arrows representing positively and negatively correlated processes, respectively.

Diagnosis starts by discretizing online data streams of the sensors. If a sensor observes an anomaly it will create an observation instance for each time sample. Anomalies are detected using building operation rules that define a normal range. We discretize values as *High* if it is above this range, and *Low* if it is below it. Causes are also discretized in *High* and *Low* using a statistical classification model [5]. The benefit of our semantic approach and the known process relationships is, that we can utilize the semantic type of processes to narrow down the nature of the potential cause. For example, if the anomaly is characterized by a *High* state, then a cause that is connected by a positive correlation process is probably also *High*. If the cause is connected by a negative process it is probably *Low*. For example, a *TempHigh* observed by the temperature sensor in Fig. 2 may be caused by a high occupancy, low cooling actuator, or high outside temperature as they are all linked by physical processes.

4 Experiments

Our test system currently manages the IBM Technology Campus in Dublin. The campus contains six buildings with more than 3,500 sensors. We investigated in more detail office building B5, which has about 271 sensors [5]. The diagnostic approach creates 747 properties, 4,498 processes and 1,097 causes for this example.

The building has temperature sensors and a heating system in most rooms. We defined as abnormal if the temperature falls two degrees below the setpoint. The potential causes for the anomaly are a low outside temperature, neighboring rooms with a low temperature, an inactive heating system, and a low setpoint of the heating system. For our building 4 % of the room temperature samples were abnormal due to an inactive heating system and / or unusually low outside temperature.

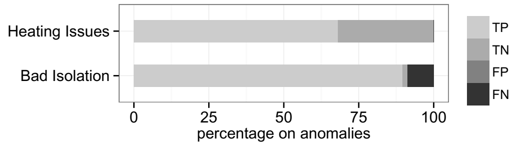

Figure 3. Results for a building: TP - true positives, TN - true negatives, FP - false positives, FN - false negatives in percentage of anomalies.

Figure 3 summarizes the results. Our approach retrieved that 67.95 % of the abnormal room temperature readings were related to an inactive heating system which could also be validated by the building operator. Furthermore our approach computed that 89.58 % of the abnormal cases were related to the outside air temperature which the operator largely confirmed to be the case.

Most importantly, our approach revealed that most abnormal temperature readings that were related to the outside temperature occurred in 11 rooms that had severe isolation problems. It turned out that they consumed an estimated 50 % of the buildings heating energy.

5 Conclusion

We have shown how to extend the SSN ontology for enabling the automated configuration and operation of analytics tasks such as diagnosis. The latter was achieved by extending the SSN by additional properties and then relating them by physical processes using SPARUL rules and generic knowledge of a domain ontology.

The approach was demonstrated for the smart building domain as it provides a good use case. Here the number of concepts for sensors, measureable properties, and processes in buildings is limited and strongly repetitive through the rooms. A specific SSN standard has not yet been established in the domain, but the required information of sensors and their locations is available for most systems and can be utilized to unlock the benefits of semantic models.

We have shown that we can indeed automatically localize the causes of anomalies for real buildings, i.e. our diagnosis result contained indeed the true cause of an anomaly. Thus the building operator only had to consider the possible causes of the diagnosis result for identifying the problem. Future work includes integrating test generation methods for assisting the operator in identifying which of the possible causes did indeed cause the anomaly.

REFERENCES

[1] Michael Compton, Payam Barnaghi, and et al., 'The SSN ontology of the W3C semantic sensor network incubator group', *Web Semantics*, **17**, 25–32, (2012).

[2] Jinsoo Han, Youn-Kwae Jeong, and Ilwoo Lee, 'Efficient building energy management system based on ontology, inference rules, and simulation', in *Int. Conf. on Int. Building and Mgmt.*, pp. 295–299, (2011).

[3] International Energy Agency. World energy outlook 2012, 2012.

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

[5] Joern Ploennigs, Bei Chen, Anika Schumann, and Niall Brady, 'Exploiting generalized additive models for diagnosing abnormal energy use in buildings', in *BuildSys - 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, pp. 1–8, (2013).