With the rapid growth of sensors and devices that communicate—that is, the Internet of Things (IoT)—smart devices have permeated every facet of modern life. These IoT devices are within our bodies, on our bodies, in the environment both inside and outside our homes, observing our behavior patterns on a day-to-day basis, and assisting in production systems and surveillance. Figure 1 highlights some of the more popular IoT applications in the world.

However, with these sensors’ ubiquity and pervasiveness comes vast amounts of data that need to be processed and analyzed to extract meaningful or actionable information from the data for recommending appropriate changes in the real world. This requires using not only semantic approaches, but also data streamlining to ensure that the decisions made are not erroneous. Moreover, due to the sheer volume of the data from these IoT devices, any errors from user entry, data corruption, data accumulation, data integration, or data processing can snowball, causing massive errors that can detrimentally affect the decision-making process. Consequently, there needs to be a clear understanding of the challenges associated with data quality and a way to evaluate and ensure that data quality is maintained for different applications.

Mapping the OSI Framework to IoT Quality

There are several implementation concepts behind determining the IoT architectures that are application-driven and face different challenges, whether at the hardware level, software level, or integration. Specifically from the Open Systems Interconnection (OSI) perspective, three quality areas need to be examined: the device level (data link layer), the network level (network layer), and the application level (presentation and application layers). The choice of IoT devices and design protocols in the IoT systems determine the choice of implementation design, as well as analysis algorithms to achieve the optimal quality of service. Specifically, the lower OSI layers have been extensively investigated in several studies to extract and transmit the raw sensor data from IoT devices through protocols such as MAC or IEEE 802.3. However, the higher layers often get overlooked, especially from the perspective of the target domain. Using two scenarios, we will show how data quality can be evaluated contextually and how the different OSI layers are affected by the specific user needs of the system.

Before we get into the use cases, we need to understand data quality, “the degree to which a set of inherent characteristics fulfills the requirements.” Within quality, we have two categories: specification and conformance quality. Specification quality refers to how well a device matches with other similar devices in that domain. Conformance quality looks at the “correctness” or the veracity of the readings from the device. We add one more quality control that needs to be examined in this setting: the device’s semantic quality. Clearly, the interpretation of the data from the sensor has a key role to play in an application targeted toward a specific healthcare requirement, as our two use cases show.

Use Case 1: Lung Function for the Elderly

With the rapid growth of wearable technologies in the mobile industry, the healthcare industry is pushing the boundaries of continuous activity monitoring using wearables. In this scenario, we consider the use of a popular wearable vest called Hexoskin for measuring physiological changes in older adults. Specifically, let’s look at one of the physiological sensors, the lung volume measures that compute the tidal volume of the lungs using the last inspiration (80 mL to 10 L) at a frequency of 1 Hz, and the frequency of the inspiration and expiration events.
Now let us look at some of the challenges in the OSI layers that need to be considered for this use case.

**Applying Data Semantics in the Presentation Layer**

In terms of preprocessing the sensor data on the Hexoskin, the vendor’s website (www.hexoskin.com) mentions baseline change and noise detection but does not provide details into what this entails. In particular, one aspect of these wearables that is overlooked when it comes to its use in the nonactive population is that the baseline or even the noise measurement can differ significantly in older populations. In fact, articles and studies describe the lung change in the older population as similar to emphysema. This can lead to a low correlation in lung function and activity, and must be considered when vital sign functions are interpreted. Furthermore, this can affect the analysis when using the vest on the elderly population. This constraint is not surprising given the fact that the sensor was primarily built for fitness measurement. Figure 2 shows how understanding the challenges at the upper layers of the OSI model can improve the performance of wearable systems in healthcare applications.

Consider an older man, Bob, who is wearing the Hexoskin vest for activity measurement. Suppose the lung function readings (data) for him are FEV1 (forced exhaled volume in 1 second) = 2.04 L and FVC (forced vital capacity) = 3 L. Although the values themselves have no significance for chronic obstructive pulmonary disease (COPD), the FEV1/FVC ratio (also called the Tiffeneau-Pinelli index) represents the information as the proportion of a person’s vital capacity that he or she can expire in the first second of forced expiration. We see from the FEV1/FVC ratio that this index is greater than 0.7. This is where the knowledge indicates that there may be concern for obstructive pulmonary disease such as emphysema. However, this could also be a manifestation of aging, so we need to dig one step further in the process and ask contextual questions, such as whether Bob has a history of smoking, or conduct additional tests, such as a flow volume loop, to diagnose whether Bob has a lung condition. If in fact the additional tests indicate that Bob has emphysema, the wisdom (relevant actionable medical science) comes in its treatment and handling the day-to-day variabilities in the symptoms that allow Bob an improved quality of life. In this scenario, the data and information part of the pyramid map to the presentation layer, which checks the context and veracity of the data readings, wherein the data are interpreted to enable semantic quality control of the IoT device readings for this application. Similarly, as Figure 2 shows, the knowledge and wisdom portions of the pyramid map to the application layer to incorporate additional data sources for more meaningful data analysis.

**Incorporating Knowledge at the Application Layer**

One way to overcome this manifestation of aging in the sensor readings is to use the raw sensor information and change the baselines for the older population or for populations with certain health conditions like emphysema. Here, the semantic mapping between the sensor values and knowledge of the population using the system needs to be taken into account to ensure that the results are accurate and useful. To potentially incorporate this information, we use ontologies or knowledge representations to annotate the data specifically for this application. An example of such an ontology is the semantic sensor network ontology. Creating such shared semantic definitions helps integrate new data into historical, temporal, and spatial contexts. Definitions of sensors and their capabilities are also useful for quality reasoning. For example, if the accuracy of a sensor depends on phenomena other than that which it measures, then a specification of this can be used as a guide to search for spatially and temporally related measurements of the phenomena on which the accuracy depends, which then defines the application’s quality metrics.
Specifically, in our example from Figure 2, for the older adult population, we can restrict the operating range (property `measurementRange`) to accommodate the differences in the breathing measurements for the target population, which can help improve the accuracy of the readings through identification of more errors or discrepancies in the data values. Many of the other parameters or entities described in the ontology are specifically left undefined to fit user requirements and enable reusability, which can be leveraged for specific target populations to improve user-centric applications, especially in healthcare.

**Use Case 2: Smart Home System**

In this use case, consider a smart home-monitoring system that uses several IoT devices to measure the physical and cognitive health conditions of older adults living independently in their homes. Such a system can be used to detect falls and to study the residents’ continuous behavior pattern and generate alerts when the residents deviate from their normal behavior. Apart from the wearable sensors discussed earlier that can measure physiological changes in the residents, environmental sensors are also placed in this smart environment. These include a depth-sensor-based system that is in place for continuous anonymized fall monitoring and activity analysis, including sedentary behavior.12,13 Moreover, wireless motion sensors are in place in the environment to study the interactions of the people in the environment, as well as to further examine their behavior patterns, such as measuring the time away from home14 and movement within the living spaces.13,14 Figure 3 shows such a system in the home of an older couple; the IoT devices are all routed through a common channel via the Internet and stored in a secure database. Corresponding behavior analysis is shared with the clinician as well as the couple’s family members. In this use case, we will discuss two challenges in terms of the network and application layers that relate to a more complex multimodal data fusion IoT-focused application.

**Incorporating Data Integration at the Network Layer**

The fall-detection system we described earlier requires a common framework that can integrate wearable sensors with environmental sensors such as the wireless body/personal area network (WBAN).15 WBAN allows long-term,
unobtrusive, ambulatory health monitoring with instantaneous feedback to the user about the current health status. The devices are connected wirelessly via low-powered networking protocols such as Zigbee (motion sensors), Zwave, or Bluetooth (activity trackers), as well as through high-powered wired connections (Kinect). Device interoperability is crucial to ensure that all the data are recorded simultaneously, continually, and accurately. However, systems such as the semantic gateway bypass the network interoperability that acts as a bridge between the IoT devices and the Internet to allow part of the data processing to occur in the gateway, enabling faster decision support.16

**Effect of Application-Driven Quality of Service at the OSI Application Layer**

A crucial and often overlooked challenge in terms of quality of service for the smart home system is that quality is dominated by its weakest link—that is, the lowest-quality sensor device. This could include a failed sensor, device-specific network connectivity issues, or even database malfunctions. As an example, consider the fall-detection system in the smart home setting. This system comprises heterogeneous sensors, such as depth and audio sensors, to detect falls inside the home and alert the clinicians. However, if the lowest frame rate among the sensing devices (say, the depth camera) is 2 frames per second, that will be the resolution of the overall system, conservatively speaking. Although this frame rate might be sufficient for activity-monitoring systems, the resolution could be too low if we want to detect falls occurring inside the home. Moreover, the depth sensor’s low frame rate can seriously affect the fall-recognition system if it is a combination of multimodal sensors such as depth and audio, which relies on the sensor fusion for detecting the fall occurrence. To address this, we can use two factors for activity recognition: the individual sensors’ data quality and knowledge of the activity itself.

An important point to note here is that both of the factors discussed in
Table 1. Factors affecting quality of service in smart home fall-detection systems.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Information (example)</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity understanding</td>
<td>A fall event corresponds to a “sudden change in height, with a sudden increased downward movement, as well as a corresponding trigger to the person being on the ground.”13</td>
<td>Incorporating the semantic data quality check to look for the temporal sequence information of the activities.</td>
</tr>
<tr>
<td>Disparity in sensor data quality</td>
<td>The depth sensor is more prone to generate false alarms in detecting falls as compared to the motion sensors in the living area.</td>
<td>Weighted aggregation of the sensor devices depending on the location of the fall event. Fuzzy aggregation methods such as Sugeno and Choquet integrals also allow uncertainty to be incorporated, which can take into account sensor data noise.</td>
</tr>
</tbody>
</table>

Table 1 require prior knowledge of the activities, as well as evaluation of the sensor data quality that can be leveraged to learn the aggregation measures for multimodal data fusion.

Internet of Everything or Indispensable Role of Humans in Quality Control

The two factors discussed in the quality of service of the smart home system in Table 1 are essential for the effectiveness of the smart home system. However, despite the incorporation of activity knowledge as well as the IoT device quality, the performance of the state-of-the-art fall-detection system is still low. One way to improve the system’s performance is to incorporate human knowledge into the IoT system architecture. In fact, the Internet of Everything (IoE) is a concept that extends the IoT framework on machine-to-machine (M2M) communications to encompass people and processes for a much larger scale of data analytics. For our smart home-monitoring system, by incorporating a human-in-the-loop, the fall-detection system can achieve a much lower false-alarm rate that will alert the clinician and family members of a fall only after a technical nursing staff dedicated for this purpose has confirmed its occurrence. This can not only reduce clinician fatigue but also prevent undue panic through a more mediated IoE approach. Moreover, through an active learning process, the existing fall-detection system can update the algorithm to reduce the number of instances where the human-in-the-loop is required using the IoE design.

In a report from Cisco on IoE innovations,18 a key insight was on the increased usage of mobile applications for interacting with IoE processes. For a fall-detection system, using a fall-detection mobile application will further allow clinicians and family members to access real-time feedback on their patient or loved one’s status, thereby transforming the current clinical decision support system.

Overall, we see the effect of the IoT, and even the IoE, on two use cases in daily life. Through improved IoT data quality, the IoT can have a staggering impact on different facets of our existence, from entertainment, surveillance, transportation, and daily activities to industry applications and healthcare.

Acknowledgments

We thank John Hughes from Wright State Internal Medicine and Mary Catherine Schafer from Jefferson City Medical Group for their valuable suggestions on the clinical aspects of the use cases described in this article.

References


Tanvi Banerjee is an assistant professor of computer science and engineering at Wright State University and Kno.e.sis. Her research interests include sensor validation and tying machine learning with sensor data for actionable information in healthcare applications, specifically in the domain of eldercare technologies. Contact her at tanvi@knoesis.org.

Amit Sheth is the LexisNexis Ohio Eminent Scholar, executive director of the Ohio Center of Excellence in Knowledge-Enabled Computing (Kno.e.sis) at Wright State University, and an IEEE Fellow. Contact him at amit@knoesis.org; http://knoesis.org.amit.