Framework for Heterogeneous Sensor Data Stream Management with BIM and WSN

Ozgun PINARER

Galatasaray University, Dept. Computer Engineering, Istanbul, Turkey

Highlights

• This paper focuses on major challenge posed by merging WSN technology with BIM.
• A novel approach is proposed to design and manage a smart building environment.
• Sensor data is processed to reflect the real time monitoring on the 3D model.

Abstract

By the 2010s, BIM is highly popular research area that gathers researchers from various domains. We tackle the major challenge of BIM and the main barrier for the application of BIM in smart environments: integration of BIM and wireless sensor networks. Existing studies in the literature have various dependencies and propose limited solutions. These solutions are dedicated to specific applications and domains. However, it is highly required to have a generic solution for real time indoor and outdoor environmental monitoring with full integration with the digital model of the building. To fulfill this need, in this study, we propose a novel approach to monitor and fulfill user requirements by a numerical model of the building, a novel framework and a BIM-WSN integrated architecture based on the BIM principles for smart building environments. All the experiments are performed on a real testbed, physical sensor devices deployed on the classrooms of the university building. Such numerical model presents the digital model of the environment by gathering acquired data. Besides, data is gathered on a big data platform where it may be served to various other applications such as machine learning studies.

1 INTRODUCTION

Use of smart technology has an important impact on enhancing the living/working environment. Designing a smart building requires an attentive interdisciplinary work. In a traditional building construction process, electrical, urban, environment, mechanical engineers, and an architect work in common to proceed the project. Besides, in the smart building construction process, computer engineers are also involved. With this interdisciplinary team, a new research domain has born: Building Information Modelling (BIM).

With the great improvement in the small size sensor technology and 3D modelling, BIM represents one of the most popular research areas from different domains. BIM provides a shared platform on which every participant may work on the 3D model at the same time without interrupting each other. Such working process provides a great working environment for agile teams. To achieve these benefits, IT-based knowledge and systems are highly necessary during the construction process of a smart building. Besides, reducing energy consumption of any environment requires a monitoring and management system which can be carried through by deploying wireless sensor devices and actuators. Deploying such equipment for managing the environment requires a highly developed framework built on suitable infrastructure. Managing wireless sensor devices and communication protocols are well-known topics in the literature. However, merging WSN technology with BIM, integration of such IoT based devices require a novel approach to merge various independent systems about designing and managing a smart building environment [1].

* e-mail: opinarer@gsu.edu.tr
The rest of the paper is organized as follows: Requirement and importance of such integration is presented in Section 2. Existing studies and more relevant research are given in Section 3. Section 4 presents the approach and system architecture. Besides, it gives details about the proposed approach, and defines all the components, tools and services integrated into the framework. Built environment and testbed are described in Section 5. Finally, conclusion and discussion are given in Section 6.

2 IMPORTANCE OF BUILDING INFORMATION MODELLING IN CONSTRUCTION

Intelligence and sustainability are two concepts that complement each other in this domain. Besides, high energy consumption of buildings is our major challenge to solve. It is proved that a significant part of the energy sources is consumed by the traditional buildings. For efficient use of resources on earth, many studies and research are performed on traditional buildings for energy harvesting and applying energy saving methods, and a new dimension has been brought to construction standards on behalf of sustainable: Building Management and Monitoring Systems via BIM.

Such smart building management systems aim to increase the living standards of the people living in the buildings and to increase the living standards, as well as to improve energy monitoring and energy consumption. Since earth resources will end one day, there is always a need for smart systems that care about energy resources and try to minimize consumption as much as possible in order to leave a livable world for future generations. Figure 1 presents the awareness and the usage of BIM in real-life. Due to the energy concern, in many developed countries, using such technologies in the building construction process becomes compulsory.

![Figure 1. Awareness and usage of BIM [2]](image)

3 LITERATURE REVIEW

As mentioned, BIM is an interdisciplinary research area in which various researchers from different domains work to get the benefits of BIM and to integrate BIM workflow to their internal work processes. Detailed literature review of BIM and its usage in different domains are deeply presented in [3]. As indicated there, BIM is used not only to manage smart building environments but also, it is commonly used in early stages of construction environments. Such environments can be very critical in terms of workers' health. BIM is highly important to prevent from industrial accidents.

Table 1 presents the enabled technologies and most popular platforms in this research domain. Current studies in the literature suggest local and limited solutions. Besides, most of the proposed approaches are prototypes and are not suitable for automation management systems. It is quite hard to adopt these approaches to real life scenarios. Moreover, in these studies, data management of multiple sensor nodes and
Actuators are not addressed, and these are limited to predefined building practices. As a concrete solution, in this study, we present a complete architecture and a well-developed framework for a smart building environment that handles the data from its acquisition till its usage in third party applications running on our framework. As indicated in several studies, existing BIM applications and implementation still require more experiences and concrete solutions about data sharing policies, insufficient collaboration, and integration.

These studies are specialized solutions and application-based approaches however a proposed solution should be generic and be able to fit to different environments. Moreover, the BIM concept, existing studies on IoT and BIM integration commonly intent energy saving and finding energy harvesting solutions [4]. Not only IoT or smart technologies but Industry 4.0 approaches are also involved in this domain. Concept of Industry 4.0 focuses on the design stage of construction in terms of using renewable and energy-efficient materials or technologies. Intention of Industry 4.0 in BIM is to increase and improve the productivity level, safety, and quality of the construction [5].

**Table 1. Enabled platforms for BIM**

<table>
<thead>
<tr>
<th>Platform</th>
<th>Research</th>
<th>Proposed Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wang et al. 2013 [8]</td>
<td>RBIVS - Visualisation of university campus</td>
</tr>
<tr>
<td></td>
<td>Woo et al. 2014 [9]</td>
<td>Analysis system to facilitate decision-making and economic modernization of aging commercial buildings</td>
</tr>
<tr>
<td></td>
<td>Hu et al. 2014 [10]</td>
<td>Optimise sensor device location</td>
</tr>
</tbody>
</table>

3.1. Usage of BIM and WSN in cultural heritage

BIM is also preferred in protecting the cultural heritage of every country. As there are very few researchers and research teams working in this field, it appears to be a highly open field for development. Within the scope of cultural heritage, historical buildings and interior decorations are evaluated within this scope. For a concrete living example, due to the fire in January 2013 in the historical palace building at the Galatasaray University Ortaköy campus, many historical paintings and wall decorations inside the building have disappeared. In addition, the 675-year-old Notre Dame Cathedral in Paris, France, was heavily damaged by fire in April 2019. Such structures are structures that should be passed down from generation to generation as the common heritage of humanity and should be preserved to the finest detail. These developments and research are referred to as Historical Building Information Modeling (HBIM).

In this sense, BIM offers a digital solution to prevent such structures from being lost. BIM, by definition, is the technology to digitize the 3D detailed model of the media. In this way, detailed information of each object in the environment is recorded. As BIM can provide detailed 3D models of buildings within the
scope of cultural heritage, it is a very important resource in repair/reconstruction etc. after accident cases such as fire.

Lee et al. proposes a new metadata to facilitate the digitization of cultural heritage structures in their work [14]. With the new data structure presented, on the virtual environment, it is thought that real time calculations can be made on the risk management of the building. In cultural heritage research domain, Logothetis et al. discusses the use and evolution of building information modeling [15].

4. PROPOSED APPROACH

In this paper, a novel framework that benefits from building model and wireless sensor nodes for a smart building environment is presented. As mentioned before, to manage such a pervasive environment like a smart building environment, various components should be merged to achieve a full automated management system.

4.1. System Design

Our proposed declarative architecture is composed of three principal layers as presented in Figure 2. Besides, on the same figure, sub layers of the developed framework are also introduced. This framework is developed not only for a smart building environment but can be easily adapted to any automation system for a pervasive environment willing to be monitored and managed.

![Figure 2. Generic architecture of the proposed approach](image_url)

**Application Layer:** This layer is responsible for the connection between the user and the proposed framework. Requests from the user are managed in this layer and transmitted by converting them to the lower layers. Likewise, information from the lower layers is extracted and served to the relevant users or parts. Visualization and the reporting issues are processed in this layer. This layer stands for the graphical user interface of the framework that interacts with the users.

**Automation Management Layer:** This layer is the core of our approach and the developed framework. Each critical action takes place on this layer. This layer has a precision role that connects the upper and lower layers of the declarative architecture. Since this layer manages a huge part of the job, it consists of several sub layers. Main objectives of these sub layers are the storage, data processing and the subscription. These layers are responsible for storing the data, streaming, processing and subscription model between the framework and the physical layer.

Besides, this layer also servesstreams the requested data to outer platforms or systems. This layer interacts with many other outer components such as 3D building and physical models. These outer components are explained in the detailed presentation of the framework.
**Physical Layer:** This layer is the lowest layer of the architecture and represents the physical environment. Wired/wireless physical sensor and actuator nodes deployed in the environment to be monitored/managed.

On this generic architecture, user requirements are processed as a top-down process. Besides, the data processing and serving to users by the integration with BIM principals follow a bottom-up process.

### 4.2. System Components

In this study, a novel framework is proposed to manage a smart building environment. Introduced framework serves not only building management but also provides a platform on which further machine learning or big data research can be performed. Figure 3 introduces the detailed architecture and the components of the system that compose the framework and other layers.

<table>
<thead>
<tr>
<th>Application 1</th>
<th>Application 2</th>
<th>...</th>
<th>Application N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subscription</td>
<td>...</td>
<td></td>
<td>Answer</td>
</tr>
<tr>
<td>Hive</td>
<td>Insert into tables</td>
<td></td>
<td>HDFS</td>
</tr>
<tr>
<td>Apache Kafka</td>
<td>CVS data files</td>
<td></td>
<td>MQTT - Mosquito</td>
</tr>
<tr>
<td>NodeMCU</td>
<td>Sensor Stream</td>
<td></td>
<td>Sensor Stream</td>
</tr>
<tr>
<td>Sensor Subscription</td>
<td>NodeMCU</td>
<td></td>
<td>Sensor Subscription</td>
</tr>
</tbody>
</table>

*Figure 3. Detailed architecture of the proposed approach*

### 4.3. Sensor Data Acquisition

Physical layer of our architecture consists of IoT driven wireless sensor devices. Several sensor nodes provide physical measures from the environment and send these sensor data to the lowest layer of the framework. Here, sensor data acquisition is achieved through MQTT protocol. In a basic network architecture, devices are communicating via HTTP. Here, instead of using request/response model, MQTT protocol uses a publish/subscribe architecture which is event-driven and enables messages to be pushed to clients. MQTT Broker is the central communication point. To receive the sensor data, a subscription should be done. Once a subscription to a sensor node is made, the sensor node publishes its sensor data. MQTT receives the sensor data and serves this data if there exists a client that is subscribed to that published data.

Each NodeMcu connects online based on the connection rules for establishing a connection to MQTT Broker. Due to the assignment of a static IP address to the microcontroller, the gateway, the subnet, the DNS, IP address and SSID with its password should be described in Arduino programming (on Arduino IDE).

On the MQTT Broker side, a configuration should be set. There exists a configuration block that is responsible for internet connection and start-up command to the sensors to make a robust and stable connection during the streaming.

### 4.4. Stream Management via Apache Kafka
A basic smart system is based on the acquired data and management of the stream data. Hence, such systems require a stream management tool to process the data flow and forward it to the right clients. In this study, Apache Kafka is preferred for the sensor data stream management. At this layer of the architecture, we focus on modeling of data streams. A tool is required to both transport these streams of data to all the systems (as a pipeline) and applications that need them to build real-time applications on top of them. Thus, Apache Kafka is a suitable solution to manage our sensor data streams.

Such stream management tools adopt the MQTT, subscribe/publish model. Here, sensor devices are the producers that produce and publish data, upper layers are our consumers that are subscribed to these streams. Besides, a topic is needed to be created. A topic can be considered as a category or feed name in which acquired data are published (inserted). The main purpose of existence of topics is that there always exist multiple subscribers for the same stream. A broker handles the consumers, producers, and the topics.

After receiving the sensor data on the MQTT Broker, a python script is executed to save this stream in a csv file. Still, sensor data should be transferred to Kafka. To get there, firstly Zookeeper is used. It is a centralized service for maintaining configuration information before distributing to other services. It provides highly reliable data registry, serialization for encoding data according to specific rules. Zookeeper ensures that the data streams run consistently. Once Zookeeper is launched, Kafka is started to run.

Kafka's main role is the real-time data feeds, collecting big data and performing real-time analysis on that data. It adopts a producer/consumer model. Hence, a creation of a topic is mandatory, and the producer should be defined. In our case, the producers are the sensor nodes. They provide a sensor data stream to our framework. While the bottom-up process, sensor source is hidden in the framework and from the application point of view, they are only service providers, and the applications benefit these services. On the other hand, parts that benefit these services are the consumers of the system.

To establish a proper producer and consumer connection, it is mandatory to create a topic. Once the topic is created, the pipeline is ready for use and Kafka sends the csv file through this pipeline to Apache Flume. Flume is a service for collecting, aggregating, and moving large amounts of data efficiently. The flume configuration file defines the rules and the paths for the messages coming through the pipelines. Via Flume, csv files are transmitted to HDFS for storage.

4.5. HDFS and Required Tools

Hadoop distributed file system (HDFS) is used to store unstructured and semi-structured data. HDFS has a master and slave models in terms of architecture: NameNode (master node)- DataNode (slave node). Data coming from the Apache Flume is unstructured. Data is kept unstructured on purpose, so that developed frameworks may handle heterogeneous data formats from various service providers (sensor nodes). To manage unstructured data in HDFS, several tools should be initialized. Once the csv file is reached to the HDFS, it should be put in a format suitable for analysis (map) and performing mathematical operations.

To be able to read the data in HDFS, another tool is also required: Hadoop Common. It provides the systems to read the data stored under the Hadoop file system. Besides, it contains libraries and utilities needed by other Hadoop modules as well.

Here, the data is successfully formatted and stored on HDFS and prepared to be read. However, data is the main resource of the system and to manage this resource, YARN is also required. YARN manages computing resources in clusters. It contains the parameters of Node Manager (framework agent, responsible for containers, monitoring their resource usage), Resource Manager, Application Manager.

4.6. HIVE

Data stored in HDFS is unstructured. So, to be able to analyze and structure it, the Hive is used. Data stored in HDFS is sent to Hive and via Hive and HiveQL, data can be processed, analyzed and can be used for...
other applications such as machine learning algorithms etc. Figure 4 shows the HiveQL "SELECT * FROM data.dissertation ORDER BY datetime DESC LIMIT 10" and the result of the given query.

![HiveQL Query Result](image)

**Figure 4. Hive and HiveQL**

### 4.7. BIM-WSN Integration via Revit Dynamo

The major novelty of the developed framework that differs from the existing studies in the literature is the BIM-WSN integration. Proposed framework provides not only sensor data management to process and to serve other applications, but also integration with the 3D model. The main benefit of such integration is to monitor the environment on the 3D model in real time.

For the BIM-WSN integration, a programming tool of Autodesk is used: Revit Dynamo. Revit Dynamo is an extension of Autodesk that uses Revit API for visualized programming. Simply it provides a building information workflow. The main advantage of using Dynamo and creating scripts on Dynamo is to automate the repetitive tasks such as fetching sensor data from MQTT Broker and refresh the 3D model in real time. Besides, it provides combining data from different categories (from different products from MQTT's point of view).

An illustration of the Revit Dynamo for our experiment environment is given in Figure 5. Revit Dynamo finds the NodeMCU ports and starts listening to them. As presented above, in our physical layer, three different sensor types are used. Based on the acquired data from these sensors, a real-time monitoring system on 3D is established and the data communication is achieved through the Revit Dynamo. Figure 6 presents the real-time temperature stream monitoring and the modification of the color of the parts according to the values read from the ports.
Figure 5. Connecting sensor data stream to Revit Dynamo

(a) Initial state
(b) after the refreshment
(c) temperature change of a room (room focused)

Figure 6. Detecting real time temperature change of a room through Revit Dynamo

5. ENVIRONMENT SETUP

For the environment, Galatasaray University building is chosen, and wireless sensor devices are deployed into offices.

5.1. Model Transformation - 2D to 3D

In traditional building construction, the model of the building is designed in two dimensions commonly via AutoCad®. To enable the developed framework presented in this study, 2D models should be transformed
into 3D models. For 3D design, in this study, Autodesk Revit is preferred. Autodesk Revit is Building Information Modeling (BIM) software offering a multi-disciplinary and collaborative approach to design and construction projects. Here, for testing our testbed built in our university building, 2D models of our engineering department building are obtained and as the first step of this study, it is transformed into 3D models. Figure 7 presents this transformation from 2D to 3D on the platform Autodesk Revit.

![2D Model](image1)

![3D Model](image2)

**Figure 7.** University building floor - Transformation process from 2D to 3D on Autodesk Revit platform

5.2. Physical devices

In this study, for the data acquisition, a sensor node is built using 3 sensors: DHT11 for temperature and humidity measurements, BH1750 for luminosity measurement and finally MQ-2 for various gas detection and measurement. All these sensors are connected to a NodeMcu ESP8266. A sample node with connected sensors is presented in Figure 8. As seen from the figure, these devices are small and light weight factors. Deployment such devices in suitable locations of the building provide a concrete environmental monitoring system. Specifications and the data acquisition parameters are indicated in Table 2.

![Sensor devices and NodeMcu with its motor shield](image3)

**Figure 8.** Sensor devices and NodeMcu with its motor shield

<table>
<thead>
<tr>
<th>Feature</th>
<th>DHT11</th>
<th>BH1750</th>
<th>MQ-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>22 x 20.5 x 1.6mm</td>
<td>21 x 16 x 3.3mm</td>
<td>35 x 22 x 23mm</td>
</tr>
<tr>
<td>Operating Voltage</td>
<td>3.5V to 5V</td>
<td>2.4V-3.6V</td>
<td>5V</td>
</tr>
<tr>
<td>Operating current</td>
<td>0.3mA</td>
<td>0.12mA</td>
<td>40mA</td>
</tr>
</tbody>
</table>

**Table 2.** Specifications and Data Acquisition Parameters
5.3. Physical representation of 3D model

As described in the architecture representation, acquired data is directed to three different parts of the system: upper layers in the same framework for the storage and to serve to applications, Revit Dynamo to monitor environment from the 3D model in real-time and finally a concrete model printed in 3D printer. This model is enriched with RGB leds for each room represented in 3D model. This lightning mechanism represents the environment status based on the acquired data. A single NodeMcu (without branching any sensor) fetches the sensor data from MQTT Broker and processes it. Based on the sensor data processing, relevant alarm or status indications are performed on the model.

5.4. Data Visualization

A graphical user interface is implemented for the user interaction. Figure 9 presents the GUI of the system. In Figure 9a, last acquired physical measures from the chosen location are presented at top. It is always possible to set the threshold values for each one. Based on the threshold values and the measured value, a warning or a critical alarm can be produced. A small sign on each measure indicates if the physical measure is between the suitable interval or a warning/alarm is produced. It is also possible to observe the previously acquired data as indicated in Figure 9b with the paging system.

![GUI of the system](image)

(a) Last values
6. CONCLUSION AND DISCUSSION

In the recent years, smart building environments and management systems became one of the trend topics in computer science. Various analysis realized by researchers from different domains prove that traditional buildings are primary consumers of a significant portion of energy resources.

In this study, we highlight the major challenge posed by merging WSN technology with BIM, integration of IoT based devices requires a novel approach to merge various independent systems about designing and managing a smart building environment. Here we focus on integrating into smart building environments and connected systems. Based on the declarative principles of PEMS, we present a sustainable declarative surveillance architecture and a novel framework adopted to that architecture for allowing the processing of massive raw data from sensors, assisted by a digital model of the building.

The most popular studies in this domain are listed with their usage and supported features in Table 3. These studies include several dependencies and propose limited solutions. They are dedicated to specific applications and domains such as tower crane operations, energy saving lighting systems etc. However, the automation management systems in practice need more generic and convenient solutions in real life scenarios. On the other hand, as indicated on the last line of the table, our proposed framework provides a real time indoor and outdoor environmental monitoring with full integration with the digital model of the building. Not only a visualization mechanism but also an alarm system is also covered by this solution.

Table 3. Data Visualization of selected location

<table>
<thead>
<tr>
<th>Study</th>
<th>Usage</th>
<th>Visualization</th>
<th>Alarm mechanism</th>
<th>Integration with BIM platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Riaz et al. 2014, 2017 [6, 16]</td>
<td>Confined space</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Kiani et al. [7]</td>
<td>Indoor</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lee et al. [11]</td>
<td>Outdoor</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Indoor/Outdoor</td>
<td>Yes</td>
<td>No</td>
<td>Semi</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------</td>
<td>-------</td>
<td>----</td>
<td>------</td>
</tr>
<tr>
<td>Liu et al. [12]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Park et al. [17-19]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hu et al. [10]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim et al. [20]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee et al. [14]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logothetis et al. [15]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our framework</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To manage a smart building environment, we propose a framework and BIM-WSN integration. In this study, for the tested environment university building is chosen and wireless sensor devices are deployed into offices. Firstly, 2D model of the relevant environment is transformed into 3D to obtain a numerical model of the environment. Besides, a highly complex framework is developed to manage the sensor data streams and makes the streams available for diverse applications. As a result, sensor data is acquired by the sensor device and is transferred to the framework. Here, sensor data is passed by several pipelines between tools and services and meanwhile it is processed to reflect the real-time monitoring on the 3D model (virtual and printed one). Concurrently, data is stored in a big data platform in order to serve to other applications such as machine learning applications, data science studies etc.

For the moment, sensor data is gathering successfully, the real dataset is being constructed and the framework executes and analyzes data in real time. In the future works, we are planning to integrate the machine learning approaches based on the acquired data. Various data analysis and machine learning methodologies can be applied to improve the current system.

**CONFLICTS OF INTEREST**

No conflict of interest was declared by the author.

**REFERENCES**


