Methodology for Rapid Prototyping of IoT Solutions

Isabel Campos de Carvalho Portugal
isabel.portugal@tecnico.ulisboa.pt
Electrical and Computer Engineering
Instituto Superior Técnico, Lisboa, Portugal

Abstract

The Lean Diagnosis is a comprehensive process analysis aimed at identifying potential inefficiencies and facilitating the implementation of optimization strategies. The primary objective of this thesis is to create a suite of IoT tools that are non-intrusive and easy to install, specifically designed for gathering data in industrial environments to support Lean Diagnosis. To validate the effectiveness of these tools, we have chosen a particular factory as a case study, but these can apply to other areas with similar processes. The IoT system comprises three essential components: one for tracking product assembly on the production line, another for detecting employee proximity to locations, and a third for measuring performance of workers engaged in short, repetitive tasks. Each of these components comes with a straightforward implementation methodology. The findings from these tools align with the factory's overarching goal of improving production efficiency. The first node measures dwell times and production times, revealing a correlation between shorter average production times and higher product output when the sample is representative. The proximity system innovates in the weighting of baseline values and maximum RSSI (Received Signal Strength Indicator). Employee allocation varies, with dynamic employees consistently linked to high-output production lines. Tasks are often resumed on the following day and comparing durations with the reference unveils misclassification and non-standard work patterns. Moreover, employee presence state adds reliability and the possibility to classify scheduled and unscheduled breaks. System accuracy is evaluated using the testing system or node usage.
Keywords: Lean Diagnosis, Internet of Things, Quick Prototyping, Product Tracking, Employee Proximity System, Activity Monitoring

I. Introduction

Competitiveness is crucial for a company's success [1], and many have turned to Lean Manufacturing to enhance efficiency and productivity. This philosophy focuses on minimizing waste, reducing costs, improving quality, and shortening lead times [2].

Lean Diagnosis is a method from this philosophy to evaluate and analyze a production process to identify inefficiencies and areas of waste. Traditionally, this method relies on collecting data manually, and a subsequent summarization of this information to convert it into useful results [1]. According to [3], the whole process is very time-consuming for professionals. The emergence of Industry 4.0 (I4.0) technologies, such as the Internet of Things (IoT), offers opportunities for synergy with Lean practices. IoT allows real-time tracking and monitoring, transforming Lean Diagnosis [1].

Specifically, rapid prototyping, which consists of non-intrusive and easy-to-install IoT devices, can engage employees in the Continuous Improvement philosophy without disturbing their usual workflow [4]. This study explores integrating IoT-enabled rapid prototyping in manufacturing environments to enhance Lean Diagnosis. It aims to assess the effectiveness and value of the proposed methodology in assisting Lean professionals during their diagnostic processes.

The methodology involves several key steps: identifying the unique challenges and requirements of the industrial environment, designing and implementing a sensor network to monitor workstations, and establishing the necessary data storage and visualization. The ultimate aim is to provide insights into employee activities, their proximity to workstations, and the production line's status, all achieved through the use of a recently developed set of sensors.

The industrial setting chosen for this study possesses distinctive characteristics that differentiate it from cases previously examined by other researchers, comprising two production lines and workstations where manual tasks are performed by employees. The study leveraged data from an existing functional testing system at the end of the production line, which provided information about the number of products manufactured. While this study's context is centered on a specific factory with its intricate challenges, the foundational techniques can be adapted to a wide range of manufacturing environments.

In summary, this research strives to contribute to the field of Lean Manufacturing by seamlessly integrating IoT technologies into industrial settings, offering valuable insights and guidance to Lean professionals.

II. State of The Art

This chapter provides an overview of the main topics in this thesis: Lean Manufacturing and the Internet of Things. The application areas and enabling technologies of the Industrial Internet of Things and a comparison between them are mentioned. Finally, there is an overview of the key features that rapid prototyping IIoT solutions require and a benchmark of already existing systems.

A. *Lean Manufacturing*

Lean Manufacturing is an approach that emerged after World War II, implemented by Toyota, aiming to reduce waste and enhance productivity [5]. It centers on eliminating inefficiencies, streamlining processes, and empowering employees [6].

To implement Lean, professionals use various tools and techniques, including Value Stream Mapping, Kaizen, Just-in-Time, Preventive maintenance, and Standardized Work Methods. These tools help organizations improve efficiency, quality, cost-effectiveness, customer satisfaction, employee engagement, and competitiveness [7] [8].

In the context of Lean Diagnosis, there is a need to fully understand a process' steps and identify non-value-adding activities. This process involves collecting information about operators' locations, time spent on each step, machine states, and their durations [9].

Traditionally, this information is obtained through manual methods such as stopwatches and pen & paper. However, in this way, this process is time-consuming and inefficient. The Industrial Internet of Things (IIoT) addresses this challenge by significantly reducing the time and cost associated with Lean Diagnosis [10].

B. Industrial Internet of Things (IIoT)

IIoT involves implementing Internet of Things (IoT) technologies within manufacturing environments [11]. It encompasses several application areas that are crucial for enhancing industrial operations, including [12]:

- Connected Worker: This application provides a comprehensive view of daily operations by monitoring employee movements, task attendance, and working hours.
- Security Monitoring: Ensuring safety and security within industrial settings is paramount. These systems may prevent accidents, safeguard equipment usage, and monitor employees' health conditions like heart rate and body temperature.
- Asset Tracking: Asset tracking facilitates real-time monitoring of an asset's location and conditions.
- Predictive Maintenance: Through sensor-based monitoring, predictive maintenance tracks equipment conditions, identifying potential malfunctions. This allows for timely maintenance scheduling, reducing downtime and associated costs [13].
- Machine Monitoring/Performance: This application provides real-time data about machine performance, detecting abnormal patterns and equipment defects. It helps in resource management, scheduling, and usage forecasting [13].

The proposed solution primarily focuses on tracking and monitoring assets, operators, and machine states. To select the most appropriate communication protocol, it is important to consider the specific needs and characteristics of each technology.

C. Communication Protocols

Wi-Fi, a widely used network in IoT solutions, offers several advantages, including a long range. However, it tends to have higher power consumption compared to Bluetooth Low Energy (BLE) and LoRaWAN [14]. Therefore, it is best suited for devices connected to a power source.

BLE is designed for portable devices with a strong emphasis on low power consumption [15]. Given the solution's focus on localizing assets, BLE is well-suited for communication between beacons and a BLE receiver.

LoRaWAN (Low Range Wide Area Network) stands out for its extended range and low power consumption, making it a practical choice compared to BLE and Wi-Fi [14]. Moreover, it can be deployed with minimal infrastructure, making it a cost-effective option.

While Sigfox offers advantages like an extended range and low power usage, it operates with a lower data rate, meaning data transmission is slower. This might not be suitable for IoT solutions that require fast data transfer, particularly for assets in motion. Additionally, Sigfox has seen limited deployment and research, making it less suitable for the proposed solution's needs [14], [16].

As a result, Wi-Fi, BLE, and LoRaWAN are the more suitable options for the proposed solution, with each chosen based on specific use cases and requirements.

D. Tracking Technologies

Although effective for outdoor, GPS technology is ineffective for indoor localization such as in industrial settings, necessitating alternative methods. For this application, precision in exact location is not crucial; instead, proximity to key areas like workstations matters. Various proximity-sensing devices are available, presented next.

Time of Flight (ToF) technology calculates distance by measuring the time it takes for a signal to travel to a surface and return. It can be used to determine if an employee is in proximity to a specific predefined location, such as a workstation or machine [1].

Radio Frequency Identification (RFID) technology utilizes tags and readers to wirelessly identify and track assets. Tags store information that can include serial numbers, product details, or other data. RFID works well for asset tracking but can be influenced by materials like metal in industrial environments [1], [3].

Beacons, using BLE technology, are small, battery-powered devices that communicate with nearby devices. They transmit a Received Signal Strength Indicator (RSSI) to indicate proximity. Beacons offer benefits such as low power consumption, ease of deployment, and user accessibility [15].

Camera systems use images and videos to track assets and employees in a workstation. This method is efficient in building up a map of the tracked area and does not require prior knowledge of the location. However, it can be time-consuming for human analysis of video footage, poses privacy and legal challenges, and can be expensive to install [4], [17].

QR codes offer an efficient way to track and identify assets. They are easily recognized and decoded by cameras, eliminating the need for manual data entry. QR codes are versatile and suitable for capturing critical data related to assets. However, they require proper positioning of the camera and QR code for reliable data retrieval [17].

In scenarios where redundancy and noise reduction are critical considerations, low-cost options such as ToF sensors and beacons with the corresponding receiver offer an attractive solution. These options allow for the deployment of multiple sensors without incurring substantial costs. Additionally, cameras and RFID systems can be considered as viable, due to its low-cost.

E. Rapid Prototyping

Rapid prototyping is based on the deployment of non-intrusive and easy-to-install IoT sensors in manufacturing environments, with special attention to the involvement of employees in the Continuous Improvement philosophy [4].

Testing a solution in these environments may be very challenging since there are a lot of components. Thus, according to Martinho et al. [4] the solution must be robust, accurate, flexible, non-intrusive, easy to install, and cost-effective.

There is a need to know the location of employees, assets and their correlation with machines' states in factories. The objective is not to track every movement of workers but to gain awareness of their proximity to a certain workstation. Furthermore, it allows the identification of unnecessary motions from workers and also recognizes the need for a certain activity that is not immediately apparent.

Rapid prototyping, in this context, is an important development tool whose ultimate goal is to address a specific problem, or in other words, eliminate a waste, that may or may not be identified. It can hold that only after the analysis of the data, the inefficiencies and their cause can be understood.

There may be two outcomes in the end. In some instances, the company may see the solution as a helpful tool to implement permanently to apply continuously a Lean strategy to the factory, associated with the need to improve the system's accuracy. In the other scenario, it might just be a means to an end, the retrieval of data during a specific period to address one or more issues. Rapid prototyping turns out to be very beneficial in both situations.

F. Benchmark

Companies such as Cisco, Bosch, and Amazon have recognized the potential of IIoT technologies and have made forays into this field. A comprehensive analysis of existing market solutions conducted by Martinho et al. [4] indicates that these solutions predominantly integrate sensors for measuring parameters such as temperature, humidity, power, energy, pressure, acceleration, magnetic fields, and light.

However, they fail to meet the specifications to develop a solution for Lean Diagnosis that correlates operators and their relation with workstations. Furthermore, they also do not cover the identification of each worker and the measure of the distance to a certain location or asset by proximity sensors [4]. This study highlights the necessity for the proposed solution, a methodology that comes as a tool to ease the work done by Lean professionals when doing a Lean diagnosis.

In summary, this discussion delves into the convergence of Lean Manufacturing and the Industrial Internet of Things (IIoT) to streamline industrial processes, reduce waste, and enhance productivity. This solution must be designed according to the rapid prototyping features, highlighting its non-intrusiveness and easiness-to-install.

III. Use Case

In the course of our project, we have developed a comprehensive set of tools and methodologies. These tools are the result of extensive research, innovation, and a deep understanding of industry challenges. Now, we are eager to test these tools in a real-world setting.

This section focuses on the factory selection process and an extensive analysis of its operations to uncover areas in need of enhancement. Additionally, it involves a comparative evaluation to identify the most suitable sensors for data collection in specific areas. Finally, we outline the architecture of the selected nodes, covering the sensing, network, and application layers.

A. Analysis of Factory Processes

The practical application of IoT-enabled Lean Diagnosis requires the selection of a factory that exhibits the need to detect inefficiencies and optimize operational processes, while also demonstrating the availability to implement IoT technologies within its workplace. Here, we introduce Arquiled as our case study, where careful on-site observations provide insights into the LED street light production processes, presented in Figure 1.

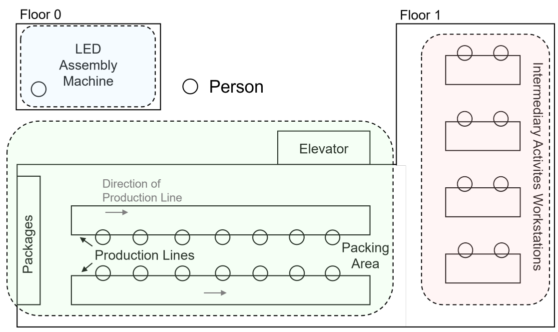

Fig. 1: Factory plant.

The LED assembly involves the automated placement of LEDs onto a plate. The machine is equipped with a semaphore light for issue alerts and is constantly monitored by an employee. Thus, No additional data retrieval is needed for this area.

There are two production lines, consisting of seven stages, with each lamp traveling through these stages within specially designed trays called *gabarits*. Notably, the third stage tends to have longer processing times, leading to an accumulation of *gabarits* in that area. One of the employees is based in the last stage tests the lamp's functionality, which provides details about the number of lamps produced. However, discrepancies in the interval between two tests raise questions about potential bottlenecks. Additionally, this worker is also responsible for moving to two other locations, one to handle materials and the other to deliver fully packed lamps to an elevator. Sensorization is proposed to monitor the lamp's progression through the production line, with a focus on the third and last stages, and to detect the relative location of the moving worker.

The factory contains a type of workstation where highly repetitive manual tasks are performed. These tasks include joining cables to terminal blocks ("Tekox - red and black cables" and "Tekox - blue and brown cables"), assembling the power supply ("MFALG assembly"), and placing screws on the board ("Screws in Board"). Work at these stations is flexible, responding to demand, and lacks a predefined sequence, resulting in varying task durations, quantities, and material requirements. The absence of a fixed program or any real-time data retrieval introduces an element of unpredictability in task execution and resource management. To enhance efficiency and employee performance in this environment, data collection is crucial. The proposed metrics to collect are the task type and duration, cumulative task duration, task quantity, employee identification, and presence data. These might enable the identification of top performers and target support for employees, fostering a culture of accountability and motivation for optimal performance.

B. Sensor Selection

The architecture of the network and sensing layer of every node is presented in Figure 2. The nodes are powered to a power source and after collecting the data, they send the information packages through Wi-Fi or Message Queuing Telemetry Transport (MQTT). Further, the data is then transmitted to ThigsBoard, a platform that stores and presents our data.

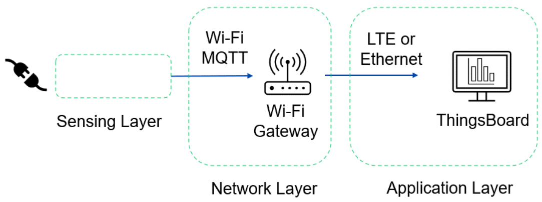

Fig. 2: Network and application layers of nodes.

To identify the *gabarits* along the production line, we initially considered the use of QR codes and cameras, due to easy deployment and scalability. However, this approach yielded inaccurate results due to challenges related to lighting and QR code readability. Additionally, the continuous movement of *gabarits* caused wear and tear on the QR codes.

The chosen alternative was the use of RFID tags attached to *gabarits*, and RFID readers placed at the desired monitoring stages. This solution is also scalable and easy to install. Nevertheless, there are challenges related to tag placement, potential metal interference, and the need for proximity between the tag and the reader. The schematic of this node is presented in Figure 3.

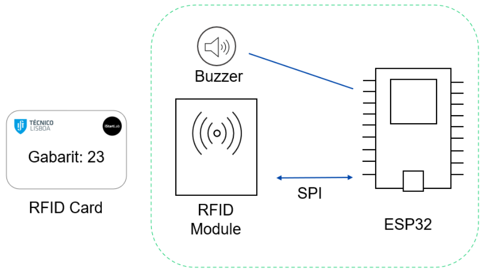

Fig. 3: Schematic of sensing layer of *gabarit* identifying node.

In the context of detecting employee proximity to certain locations, two potential solutions were explored. The first involved distance sensors and RFID cards for identification but faced limitations related to accuracy and workflow disruptions. The second, more promising solution, utilized beacons and receivers to measure proximity. This solution had minimal impact on employees' workflow and, with clear communication, we addressed concerns about being monitored. Additionally, the effectiveness of beacons may be influenced by environmental factors such as signal interference or obstructions, whether by the human body or other materials, especially metal surfaces, thus requiring proper placement and calibration. In Figure 4 we present the architecture of this node.

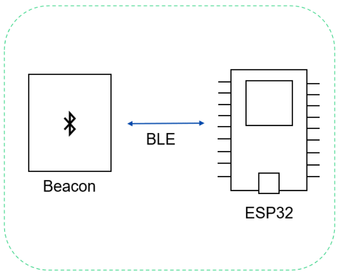

Fig. 4: Schematic of employee's relative location node.

The interaction with the task monitoring node allows the collection of data from repetitive tasks in a manufacturing environment and provides real-time feedback to employees. This node is equipped with several sensors and components, including buttons, RFID sensors, and a distance sensor.

For evaluation of individual performance, an RFID system functions with the actual identification cards used in the factory. To identify the tasks in operation, workers use tailored cards associated with a particular task in another RFID system. Due to the difference in frequencies, it is not possible to merge the two systems. Employees use a green button to indicate task completion, and a yellow button to signal breaks. Additionally, the node is equipped with a distance sensor that measures the distance of an object interrupting its signal. This helps determine when an employee is present at the workstation. The schematic of this node is present in Figure 5

C. Deployment

Recalling Figure 1 where the factory plant is presented, we now present the location of each node in the plant in Figure 6.

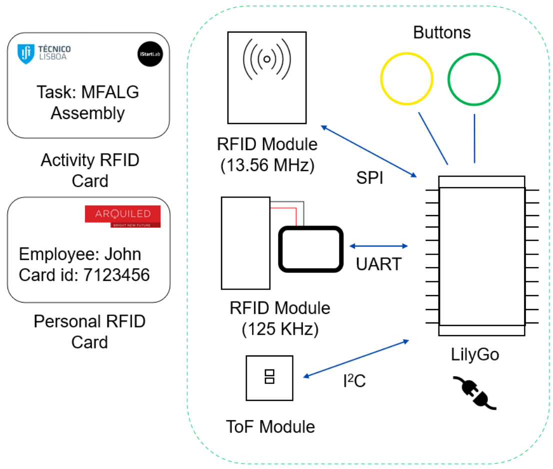

Fig. 5: Schematic of task monitoring node.

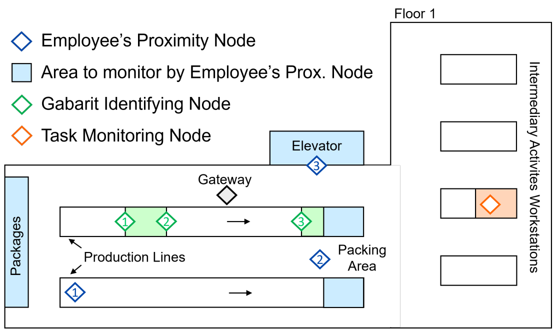

Fig. 6: Factory plant with installed nodes.

Here, the green shapes illustrate the *gabarit* identifying nodes. Nodes 1 and 2 scan the entry and exit of the *gabarits* at the third stage of the production line, known for taking a considerably higher processing time than the remaining. Node 3 monitors the entrance in the final stage, where the final functional tests are made. The blue shapes are the employee's proximity nodes, where node 2 is located at the end of the production line, the base station of the moving employee. Further, node 1 is near the packages area, where employees move to handle materials, and node 3 is placed in a high structure near the elevator, where the employee delivers complete products. Lastly, the orange node represents the task monitoring node and is attached to a table where intermediary activities take place.

D. Installation Methodologies

As the implementation methodologies for each node consist of a great part of the value of this work, we present them in APPENDIX. These instructions offer practical guidance for deploying each node. Furthermore, the appendix provides troubleshooting information for common issues encountered with the task monitoring node.

E. Installation Challenges

Throughout the installation phase, we faced various challenges. In this subsection, we provide a detailed account of the strategies implemented to overcome these challenges.

The product tracking nodes are implemented in one of the production lines. The system consisting of both the node and the card must be positioned near, to ensure scanning of the card.

The choice of where to place cards on the *gabarits* involves several considerations. Placing the card on top ensures smooth *gabarit* flow but increases the distance to the tracking node. Conversely, positioning the card below the base ensures better communication but introduces potential wear and tear due to friction. An alternative option is embedding the card within the *gabarit*, but this necessitates significant modifications and is currently unfeasible for easy temporary installation. After evaluation, the decision favors placing the card on top of the *gabarit* base, prioritizing uninterrupted flow, minimizing wear and tear, and avoiding significant modifications to the *gabarits*.

Initially, a structure was considered to place the node, but it was replaced with a simpler solution involving the removal of one conveyor roll and the placement of a support structure under the *gabarit*. The removal of the roll did not disrupt the usual workflow and this adjustment allowed for proximity between the two parts.

Another installation challenge to consider was related to the width of the conveyor, which exceeds that of the *gabarits*. This dimension disparity could potentially lead to errors in the reading of the tags attached to the *gabarits*. Having understood that employees typically work with the *gabarit* closer to them, the card is attached with the *gabarit* fully leaning toward that boundary of the conveyor. Though, this configuration does not ensure absolute accuracy in scanning every time, as it will be seen later.

After the implementation, we began receiving data that raised questions about the accuracy of the implementation. Time differences between the first and second nodes were often just a couple of seconds, insufficient to capture the entire process at that stage. To address this, future installations should consider greater node-to-node distance to account for process variability and possibly detect bottlenecks.

Each node has the goal to capture the BLE signals in the vicinity area where it is attached. The nodes were installed as mentioned in the area where packages are (station 1), in the last stage of the production line (the main station 2), and above the elevator (station 3).

It is relevant to highlight that the receiver positioned near the elevator is situated at an elevated position, which results in a greater distance between the receiver and employees, compared to the other nodes. The higher vertical distance affected the reliability of the detection of an employee at this station.

Employees assigned to the final stage of each production line (A and B) were provided with a beacon (associated with A and B). They placed the beacon in their uniform pocket and feedback revealed that this setup did not disrupt their standard workflow.

Given the characteristics of the proximity sensor, the task monitoring node needs to be positioned in the direction of the employee operating at the workstation. Furthermore, it is placed in proximity, so that it facilitates easy access for card replacements and button presses. The cards are conveniently stored in a nearby container, ensuring easy access to them. Further, the node was securely attached to a table using Velcro, providing flexibility for adjustments when necessary.

Typically, an employee is categorized as "Near" when the distance measures less than 80 cm. However, in specific cases where the distance reading remains nearly constant, it is deduced that an object was mistakenly left in front of the node. In such cases, regardless of the distance, the state is labeled as "Away".

IV. Results

To aggregate the data received, an API is utilized to extract it from the ThingsBoard platform. Whenever possible, we cross-reference this data with the information supplied by Arquiled's testing system, called as node 4 for simplicity. Finally, we present visual representations of this data to enhance understanding.

A. Product Tracking Node

In this subsection, our goal is to gather insights into the movement of products along the production line. Our analysis begins with an examination of the total time spent at each station, followed by an analysis of the system's detection rate and confrontation with factory data.

The total time spent in each stage accounts for dwell times (periods when the *gabarits* remains stationary near a node) and the time difference between nodes. Summing up these durations gives us the total production time for the monitored part of the production line, as shown in Table Table I.

TABLE I: Average duration of each stage and total monitored production time for models X and Y.

	Average time in (mm:ss)			
Model	**Stage 1**	**Stage 2**	**Stage 3**	**All Stages**
X	00:23	04:05	01:33	06:01
Y	00:48	03:46	01:56	06:30

Certain data points significantly exceeded this average production time, suggesting potential issues in the production line. To identify the responsible station or stations for such deviations, we can calculate the data points with the longest production times, as shown in Table II.

The cells highlighted in gray represent the process durations that are more than twice the average. We can conclude that products that took significantly longer to manufacture, approximately 23 minutes and 17 minutes and a half, appear to be affected by prolonged durations in two processes. Moreover, we note that stage 1 shows the most consistent delays concerning average.

The capacity to obtain accurate insights from the data collected is deeply influenced by the detection rate of the system, thus its study is crucial. We take the number of products tested by Arquiled as the reference for the number of products manufactured and provide the average detection rate for each node as well as for all nodes in Table III. Additionally, the average number of *gabarits* detected by all nodes is discriminated in the last two columns by its functioning status (OK or KO).

TABLE II: Highest peaks in the total production time of dataset. Cells highlighted in gray indicate process durations that are more than twice the average.

		Time in (mm:ss)			
Date	**Model**	**Stage 1**	**Stage 2**	**Stage 3**	**All stages**
25-09-2023	X	01:21	05:05	03:22	09:48
	X	01:45	19:26	01:39	22:50
27-09-2023	X	00:03	05:20	04:21	09:49
03-10-2023	X	02:22	03:36	11:30	17:28
	X	01:38	05:54	01:44	09:16
09-10-2023	X	01:17	07:31	01:47	10:35
	X	00:03	08:25	01:46	10:15
10-10-2023	Y	00:03	08:31	02:05	10:39
	Y	00:03	14:30	01:38	16:11
11-10-2023	Y	01:02	06:11	03:39	10:52

TABLE III: Percentage of detection rate in each node and in all nodes, and frequency of lamps that tested OK or KO.

Average detection rate in				Average number of products detected	
Node 1	**Node 2**	**Node 3**	**All nodes**	**OK**	**KO**
64%	40%	52%	13%	39	1

We observe an average detection rate for each node between 40 and 64%, and of 13% in all nodes. This rate appears low, but it is important to consider that a single misalignment of a *gabarit* with a node can result in it not being included in this statistic. We assume that the sample retrieved is a random set of *gabarits*. Hence, we predicted that the higher the detection rate on a given date, the higher our system approximates the real measures.

We may attempt to relate the average production time with the number of manufactured products by that line. For this matter, we present Table IV

In this analysis, we use a table to assess the impact of individual stage durations on monitored production time and its correlation with the average time differences between tests and the total output of the production line. While the time difference between two tests may not directly correlate with total production time, its average can provide a metric of production speed, helping us determine if the sample of *gabarits* detected that day is representative.

For example, on October 27th (in green), a shorter stage 1 duration resulted in a decreased average time for product manufacturing and one of the highest production outputs. Confronting with Table III, although only having a detection rate of 9%, provides a reasonably accurate representation of overall production.

Conversely, on October 11th (in red), the lowest total product count and longer time differences indicate slower production due to increased stage durations. Despite a 9% detection rate, the data still offers a reasonable representation of the production.

However, there is an exception on October 6th (depicted in blue) where the time difference suggests fast production and high output, but the system measured a longer production

TABLE IV: Summarization of average stage times and total monitored production time. In addition, the average of the time differences between tests as well as the total output of the line.

		Time in (mm:ss)					
Date	**Model**	**Stage 1**	**Stage 2**	**Stage 3**	**All stages**	**Average test time difference (mm:ss)**	**Total products manufactured**
25-09-2023	X	00:33	04:22	01:30	06:25	01:27	304
27-09-2023	X	00:06	03:34	01:35	05:15	01:19	333
28-09-2023	X	00:14	03:47	01:27	05:28	01:23	316
02-10-2023	X	00:26	04:26	01:28	06:20	01:28	300
03-10-2023	X	00:26	03:42	01:40	05:48	01:24	313
06-10-2023	X	00:20	04:55	01:19	06:34	01:16	346
09-10-2023	X	00:33	03:48	01:50	06:11	01:29	295
10-10-2023	Y	00:38	03:30	01:52	06:00	01:24	314
11-10-2023	Y	00:58	04:01	01:59	06:58	01:37	270

time. With only a 4% detection rate, this data may not be representative, potentially affecting the overall results.

B. Employee's Proximity Node

In this subsection, our objective is to extract insights about the relative locations of employees. Initially, we discuss the approach used to get the system that classifies a specific data point. Subsequently, we conduct an analysis of metrics that we consider pertinent for a particular set of days and, in more detail, for an individual day.

The factory is replete with metal components and machinery, which can significantly impact Bluetooth signal propagation. Thus, it would be helpful to establish baseline RSSIs for each station, taking the environmental conditions into account. These are presented in Table V. For instance, the first row corresponds to the received values from each node when a beacon is placed in station 1.

TABLE V: Average RSSI values received in every node when the beacon is placed in each station.

	RSSI of node (dB)		
Beacon placed in station	**1**	**2**	**3**
1	-66.4	-74.3	-73.9
2	-60.9	-43.4	-66.0
3	-68.7	-56.4	-61.5

Beacon placement in each station typically results in a stronger signal at that station. However, station 3 deviates from this pattern due to factors like its height and metal structure. To study the impact of human bodies on the signal, an employee's route, represented in Figure 7 was simulated through stations 1 to 3, recording the time spent at each station.

Using RSSI data from each node, we adopted a weighted approach to determine the person's location. Weight combinations, of stations 1 to 3 and the maximum RSSI, were iteratively tested to maximize classification accuracy. Note that transition areas between stations were not included in the training data due to data limitations. The selected weight combination was 0.1 for station 2, 0.9 for the maximum RSSI, and 0 for the remaining stations.

We eliminated data points that fell between other classified stations, as they were likely the result of momentary signal

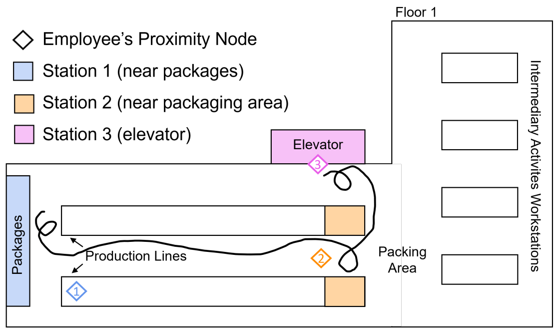

Fig. 7: Plot depicting the trajectory of a person with a beacon, simulating an employee's route, as they move from station 1 to station 2, and ultimately to station 3.

redirection and did not signify actual station transitions. This allowed us to reach an accuracy for stations 1 to 3 in comparison to the ground truth of 97%, 93%, and 94%, respectively. Subsequently, the system is applied to real-life data, while comparing with the data provided by Arquiled.

On the days analyzed, both production lines, A and B, are in operation, and the employees from each line are associated with beacons A and B. Gathering the results, it is beneficial to compare the metrics of both employees such as the relative time spent in each station, the number of times a station was visited, and the average time a person remained in a station (average dwell time). In Table VI, we present the average of each key metric for all available dates.

As expected, due to station 2 being the base station, both employees spend the most time there. When leaving for other locations, employee A spends more time in 1 and with a higher average dwell time, compared to B, while employee B spends more time in station 3, as well as a higher average dwell time, compared to A. The observation may suggest that person A handles package boxes for both production lines, and, in turn, person B takes on the responsibility of transporting final products from both production lines to the elevator.

Furthermore, employee B makes more frequent visits to stations, and their production line is associated with a higher production output. This disparity between both lines could be attributed to several factors, including differences in the workforce in each production line, and the production time,

TABLE VI: Summary of average key metrics for beacon A and beacon B across analyzed dates, including relative time spent, station visits, average dwell times and, in addition, the amount of product manufactured by line A and B.

	Percentage of time in station				**Station counts**			**Dwell time in station (mm:ss)**			**Products manufactured by line**
Beacon	**1**	**2**	**3**	**ND**	**1**	**2**	**3**	**1**	**2**	**3**	
A	11%	70%	11%	8%	57	134	100	01:03	03:13	00:35	308
B	6%	63%	18%	13%	82	203	168	00:22	01:48	00:48	332

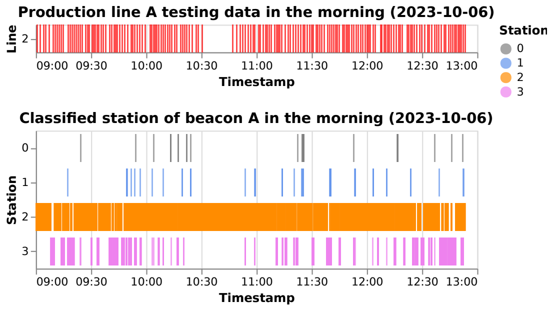

Fig. 8: Classified stations for beacon A during a morning shift on the October 6th. The top part of each figure presents tests provided by Arquiled, with missing data indicating employees' breaks.

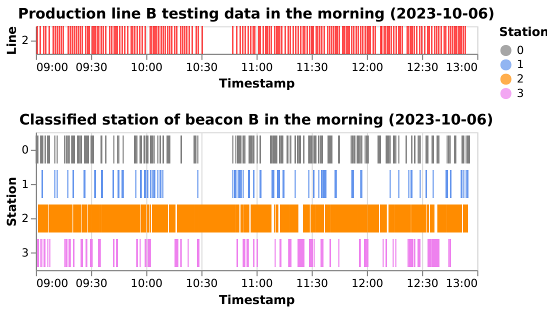

Fig. 9: Classified stations for beacon B during a morning shift on the October 6th. The top part of each figure presents tests provided by Arquiled, with missing data indicating employees' breaks.

analyzed in subsection IV-A, among many others. However, having access to the employees' relative location, we may conjecture that when the employee moves more frequently to other stations, their line displays a higher output. To confirm this, further investigation is needed in the factory.

The analysis of a specific day is of great importance since some circumstances may be particular to certain days. Furthermore, this approach enhances our capacity to visualize and comprehend the patterns discovered in the averaged analysis, as we present Figures 8 and 9.

This figure displays the classified station for beacon A (Figure 8) and beacon B (Figure 9) during a morning shift on the 6th of October. Furthermore, the tests provided by Arquiled are presented in the top part of each figure, with the absence of data indicating the employees' breaks. These points are excluded from the metrics calculations. It is also relevant to mention that timestamps with the color gray denote instances when the person was considered in transit between two stations.

As it can be seen in Figure 8, the visits of beacon A to station 1 are more scarce, resulting in a less filled line at this station, in contrast to Figure 9. This graph not only offers a visual representation of the overall allocation of employees to each station but, if seen in an amplified view, the sequence taken: which station was visited after the other.

We acknowledge that the data is likely adjacent to classification errors, due to the swift movements of employees, contrary to the set trained. Nevertheless, the metrics presented and the visual representation of employee's station allocation may offer valuable insights to managers, especially given their lack of prior knowledge of this information.

C. *Task Monitoring Node*

This analysis aims to extract insights from the activities performed by employees. On dates available, the factory mainly manufactured the product associated with the tasks previously mentioned: "Tekox - red and black cables", "Tekox - blue and brown cables", "MFALG assembly" and "Screws in board". Additionally, when prompted by demand for an order, employees use an additional card labeled as "Extra Activity 1".

The average times of each task, as provided by managers at Arquiled, are presented in Table VII. These values are taken as a reference for the duration of a task, though not as the ground truth.

TABLE VII: Arquiled's data with average duration of task and quantity of tasks produced per hour.

Task	**Mean (s)**	**Quantity (per hour)**
Tekox red and black cables	12	300
Tekox blue and brown cables	12	300
MFALG assembly	144	25
Screws in board	40	90
Extra activity	60	60

The node collected data about the employee, task identification, and the duration of the task performed. We observed that some points deviated significantly from the remaining. To address this, we adopted a strategy based on the median duration of each task measured for each day. We considered it reasonable that a person could complete a task in up to 50% of the median duration and take a maximum of double the median duration to complete it. This approach was confirmed as valid by Arquiled managers. Any task duration falling outside this range was treated as an outlier and removed from the dataset.

TABLE VIII: Summarization of the task dataset per date, specifying the employee, the task they performed, the total duration, the median of the time per task, and in addition, the deviation between the quantity per hour provided by Arquiled and the same value retrieved from the node.

Date	Task(s) performed	Employee	Total duration (hh:mm:ss)	Median (s)	Quantity per hour (node)	Quantity per hour (Arquiled)	Deviation to Arquiled
31-07-2023	Screws in board	Felix	03:22:48	26	133	90	+47%
01-08-2023	Screws in board		00:39:52	25	135	90	+50%
	MFALG assembly		01:28:10	112	31	25	+24%
02-08-2023	MFALG assembly		03:33:34	120	29	25	+16%
06-09-2023	Extra activity		01:16:34	22	150	60	+150%
07-09-2023	Extra activity		03:00:38	24	147	60	+145%
27-09-2023	Tekox red and black cables	Lucas	00:07:40	12	290	300	-3%
	Screws in boards		00:11:33	40	88	90	-2%
	Tekox blue and brown cables		00:05:25	12	299	300	0%

The activities performed at this workstation vary based on specific order demands. Therefore, the node data is exclusively collected on days when the workstation is in operation. The data is summarized in Table VIII, providing details for each day, including the tasks performed, the employee responsible, total duration, median duration, average tasks completed per hour, and standard deviation.

Using the median aims to represent the central tendency in the dataset, while the number of tasks per hour already includes the task average, making it a suitable metric for comparison with Arquiled's reference values.

According to Table VIII, concerning the employees, Felix predominantly occupied this workstation, and on the final day, Lucas took over. Furthermore, each of the mentioned tasks is executed at least once. The gathered data allowed us to retrieve some insights:

- **Tasks are often continued on the following day.**

This could imply that it was not feasible to complete the required work on the first day alone. Alternatively, it is possible that, since the workstation was already set up for a specific task, the employee leveraged that readiness and continued the task.

- **Committing more time to the same task on a given day causes an increase in the median time and a decrease in the quantity per hour.**

Notably, when Felix spent more time on a specific task, it led to a slightly longer median duration and a lower average number of tasks completed per hour. This pattern suggests that productivity may be influenced by a loss of focus.

- **Unstandardized work methods might lead to increases in the completion time of each individual task.**

In the task "Screws in board", despite Lucas operating on the task for a shorter period, the mean task duration is longer than Felix's. This suggests that Felix may be more adept at performing the task quickly, potentially employing a specific technique or having more experience.

Now comparing the quantity per hour measured by the node with the reference given still in Table VIII, we may see that there is a considerable difference between these measures in some of the values.

Indeed, particularly on the days when Felix executed the tasks, there was a notably higher quantity produced per hour. A closer examination of the "Extra activity" reveals a deviation of 145% and 150%. This drives the deduction about misclassification of the task (i.e., they were performing a different task than the one mentioned), or the average provided by the company for this particular task might be inaccurate. Either way, further investigation may be needed to clarify these discrepancies.

Further, the deviations observed in the remaining activities performed by Felix (ranging from 16% to 50%) are higher in comparison to those by Lucas, suggesting that the tasks may not be entirely standardized. It is possible that the reference values were measured during one of Lucas' shifts, highlighting potential variations in task execution between different employees.

Taking a closer look at September 7th, Figure 10 illustrates task duration in each timestamp, during Felix's shift. In addition, the distance sensor data is also shown, classifying the employee's presence. It is worth distinguishing two main types of breaks: unscheduled and scheduled breaks.

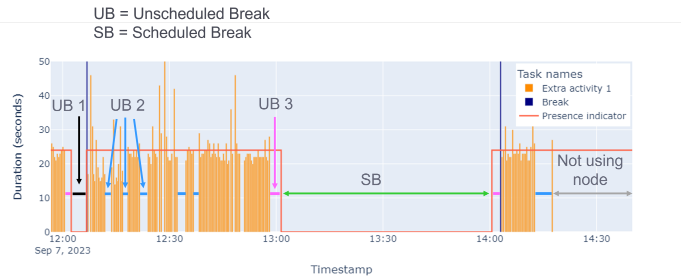

Fig. 10: Task duration and employee's presence on September 7th. Provide details if an employee is present and engaging with the node, and when not, these periods are classified as scheduled or unscheduled breaks or as periods of non-interaction.

Starting at approximately 12:00, we observe a break (shown in black), as confirmed by both the lack of task performance and presence data. At this time, we deduct that the employee might have left the workstation either for personal rest or material handling.

Following this, the employee works for a consistent period, engaging with the node until around 13:00. However, there are often periods when the employee is not detected performing tasks (shown in light blue). These shorter intervals, which we refer to as "micro-stops," may occur due to countless factors. While each micro-stop may be brief, the cumulative time spent can add up significantly.

From 13.00, the employee is not at the workstation and is expected to be on their lunch break (depicted in green). This period is stipulated by the factory and, thus is considered as a scheduled break.

Taking a closer look, there is a third type of unscheduled break (shown in pink), distinguished by their specific timing. Just before leaving the workstation, when finishing a set of tasks, and right after arriving, when starting a new one, there are moments when the employee is at the workstation but not operating on any task. During this time, employees may engage in activities such as tidying up their workspace before leaving or gathering their tools from their workspace when arriving, for example.

Overall, the study of breaks, especially the micro stops due to their high frequency, can be quite valuable. Recognizing their existence is the first step, and might lead the factory to examine their root causes, and mitigate issues encountered.

Finally, at 14:00, the employee resumes task performance (and interaction with the node). Notably, there is a selected section in the Figure (highlighted in gray) where the node detects the employee's presence, but they do not interact with the node. In fact, the trustworthiness of these insights is, of course, dependent on the engagement with the node.

We assess the reliability of our task duration data by examining how often employees interact with the node in relation to their presence near the workstation. This data provides a reference point for the reliability of the task data, as there is no confirmation of employee shifts from the factory. On days when Felix operates at the workstation, he consistently demonstrates a strong commitment to interacting with the node, with usage percentages ranging from 64% to 96%. These values contribute to the reliability of the task data for those specific dates.

V. Conclusions

This thesis had the initial goal of developing methodologies for the implementation of IoT solutions in industrial settings, specifically to support Lean professionals in conducting Lean Diagnosis. The objective was put to the test through the creation and deployment of a set of IoT nodes to collect data on production times, employee locations and their relation to the workstation, and task allocation.

With the increasing adoption of IoT in industrial settings and the need for data on employee interactions within workstations, our solution was designed with a focus on non-intrusiveness, easy installation, and cost-efficiency. We conducted a case study in an LED street light factory with two production lines where manual tasks were performed. Leveraging data from an existing testing system, our proposed solution involved tracking nodes for products, employee location, and manual tasks in a workplace lacking real-time data collection.

In summary, product tracking nodes allowed us to pinpoint stages that resulted in a high production time. Regarding the proximity system, comprised of receivers and beacons, it was possible to track employee's relative locations, revealing distinctions in behavior and responsibilities. Task monitoring data provided insights into employee performance. Adding the detection of employee presence data, we highlighted patterns and classified breaks as scheduled and unscheduled.

While not providing the exact causes behind the pain areas, the study successfully pinpointed them. At this point, the company has the option to install a permanent solution that includes more accurate measurements to address the identified problems or they might be satisfied with the insights obtained and choose to accept them without further action.

The proposed solution, tailored for this case, has broader applications in settings with similar underlying processes. It is open-source, ensuring accessibility for interested parties. By sharing our findings and tools, we hope to contribute to positive transformations, not only in industry but also in other domains where similar challenges persist.

References

[1] R. A. Martinho, "IIoT for Lean Production Systems," Master's thesis, Instituto Superior Técnico, Oct. 2021.

[2] M. A. Lewis, "Lean production and sustainable competitive advantage," *International Journal of Operations & Production Management*, vol. 20, no. 8, pp. 959–978, Jan. 2000, publisher: MCB UP Ltd. [Online]. Available: https://doi.org/10.1108/01443570010332971

[3] S. M. Asaad and H. S. Maghdid, "A Comprehensive Review of Indoor/Outdoor Localization Solutions in IoT era: Research Challenges and Future Perspectives," *Computer Networks*, vol. 212, p. 109041, Jul. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128622001918

[4] R. Martinho, J. Lopes, D. Jorge, L. C. de Oliveira, C. Henriques, and P. Peças, "IoT Based Automatic Diagnosis for Continuous Improvement," *Sustainability*, vol. 14, no. 15, p. 9687, Jan. 2022, number: 15 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2071-1050/14/15/9687

[5] R. Drath and A. Horch, "Industrie 4.0: Hit or Hype? [Industry Forum]," *IEEE Industrial Electronics Magazine*, vol. 8, no. 2, pp. 56–58, Jun. 2014, conference Name: IEEE Industrial Electronics Magazine.

[6] M. Hermann, T. Pentek, and B. Otto, "Design Principles for Industrie 4.0 Scenarios," in *2016 49th Hawaii International Conference on System Sciences (HICSS)*, Jan. 2016, pp. 3928–3937, iSSN: 1530-1605.

[7] M. Oliveira, "Apresentação lean," Instituto Superior Técnico, 2021.

[8] S. Gupta and S. K. Jain, "A literature review of lean manufacturing," *International Journal of Management Science and Engineering Management*, vol. 8, no. 4, pp. 241–249, Nov. 2013, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/17509653.2013.825074. [Online]. Available: https://doi.org/10.1080/17509653.2013.825074

[9] B. A. d. O. Madaleno, "Lean Production in SMEs - Diagnosis and Implementation Plan, a case study," Ph.D. dissertation, Instituto Superior Técnico, Nov. 2018.

[10] S.-V. Büer, J. O. Strandhagen, and F. T. S. Chan, "The link between Industry 4.0 and lean manufacturing: mapping current research and establishing a research agenda," *International Journal of Production Research*, vol. 56, no. 8, pp. 2924–2940, Apr. 2018. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/00207543.2018.1442945

[11] H. Boyes, B. Hallaq, J. Cunningham, and T. Watson, "The industrial internet of things (IIoT): An analysis framework," *Computers in Industry*, vol. 101, pp. 1–12, Oct. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166361517307285

[12] S. Accent Advanced Systems, "Connected WorkerAccent Systems," 2021, visited on 29-12-2022. [Online]. Available: https://accent-systems.com/solutions/connected-worker/

[13] S. U. Corporation, "Industrial IoT (IIoT) Services & Solutions - ScienceSoft," 2022. [Online]. Available: https://www.scnsoft.com/services/iot/industrial

[14] J. de Carvalho Silva, J. J. P. C. Rodrigues, A. M. Alberti, P. Solic, and A. L. L. Aquino, “LoRaWAN - A low power WAN protocol for Internet of Things: A review and opportunities,” in *2017 2nd International Multidisciplinary Conference on Computer and Energy Science (SpliTech)*, Jul. 2017, pp. 1–6.

[15] K. E. Jeon, J. She, P. Soonsawad, and P. C. Ng, “BLE Beacons for Internet of Things Applications: Survey, Challenges, and Opportunities,” *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 811–828, Apr. 2018, conference Name: IEEE Internet of Things Journal.

[16] U. Raza, P. Kulkarni, and M. Sooriyabandara, “Low Power Wide Area Networks: An Overview,” Jan. 2017. [Online]. Available: http://arxiv.org/abs/1606.07360

[17] S. J. Hayward, K. van Lopik, C. Hinde, and A. A. West, “A Survey of Indoor Location Technologies, Techniques and Applications in Industry,” *Internet of Things*, vol. 20, p. 100608, Nov. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660522000907