Increasing accuracy RFID Dock Door Discrimination with Naive Bayes Classifier

Bastijn Berenschot

TU Delft

Increasing accuracy RFID Dock Door Discrimination with Naive Bayes Classifier

by

Bastijn Berenschot

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday January 17, 2023 at 15:00.

Student number:	4451899	
Report number:	2022.MME.8752	
Project duration:	May 1, 2022 – January 17, 2023	
Daily supervisor Mieloo & Alexander:	Ir. W. Romijn	
Thesis committee:	Ir. M. B. Duinkerken,	TU Delft, supervisor
	Prof. Dr. R. R. Negenborn,	TU Delft

This thesis is confidential and cannot be made public until January 17, 2023.

TU Delft

Preface

This report has been written to obtain the master's degree in Mechanical Engineering at the Delft University of Technology with a specialization in Multi Machine Engineering. In this report, I present the insight of improvements of accuracy for RFID Dock Door Discrimination. I was engaged in writing this report from May 2022 till January 2023. This report is the final version of my graduation project. I would like to express my gratitude to my TU supervisor Dr. Ir. Mark Duinkerken for mentoring me through this project, and to Prof. Dr. Rudy Negenborn for helping me with insights into the academic world. I am also grateful for the supervision of Ir. Anouk Pelser and Ir. Walter Romijn, who helped me by coaching me through the project day by day. Next to that, I would also like to thank the company Mieloo & Alexander for giving me excess to their resources and facilities in order to conduct my graduation project.

Abstract

Errors during truck loading at dock doors lead to unwanted wrong deliveries in logistics. Due to the falling price of RFID tags, RFID dock door discrimination is now being used for product registration. The problems that arise with the current system for RFID Dock Door Discrimination are cross-reads and miss-reads. The purpose of this research is to increase the accuracy of product registration by the proposed Dock Door Discrimination method with Naive Bayes Classifier (NBC). A hardware design, including 4 RFID antennas at three adjacent dock doors and 1 added antenna at the staging area, and software design, including the implementation of the NBC, are proposed to improve the RFID Dock Door Discrimination. The Experimental Setup and Plan were used to gather data to compare the current RFID transition system with the new proposed NBC system in six scenarios, three with and without noise at other gates. For each individual scenario, the accuracy improved most with NBC with one input feature. The accuracy for all scenarios combined for the collected data improved from 82.1±0.8% (current) to 93.6±0.5% (NBC).These results mean that there is a solid improvement in implementing the Naive Bayes Classifier over the current RFID transition system for Dock Door Discrimination.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Logistics can be found in many places around the world. In particular, many logistics processes take place at large warehouses and distribution centres. One of these logistic processes is the loading of goods on trucks at the so-called dock doors, shown in figure 1.1. It is important to keep a good understanding of where products are and where the products should go, something called stock registration. Errors in the loading of trucks at the dock doors can lead to incorrect deliveries. As a result, customers may claim damages because the delivery is wrong or late, which is not desirable. It is therefore important to register the products in order to have proof of which product went into which truck.

Figure 1.1: Dock door system (loading bay) (Lyzs, 2022)

The hands-free automatic registration solution for identifying these products is Radio Frequency IDentification (RFID) technology. The operation of such an RFID technology relies on Ultra-High Frequency (UHF) radio waves to identify the RFID tags, which are attached to all products. Due to the falling price of RFID tags (Chawla and Ha, 2007), RFID technology is increasingly being used for stock registration when loading different kinds of products on trucks. With the help of an RFID technology at the dock doors, it will be registered which products or crates have been identified and in which trucks they are stashed. Before the implementation of RFID in this logistics process, manual checks by warehouse operators were still used. Since the RFID solution is hands-free, the registration of the products is no longer labour-intensive and the information about the location of the products is stored in the cloud.

1.1. Problem

The problem that arises when installing an RFID system in a multiple dock door environment, is in the proximity of the dock doors. These dock doors are placed with as little space between them as possible at warehouses, because that leads to the (un-)loading processes being most profitable, since you can stack more dock door lanes in the least amount of work floor space. Figure 1.2 shows schematically from a top view how the installation of an RFID system is carried out at three contiguous dock doors. In this figure, the RFID antennas with their reading field are shown in red, the static products stored in the staging areas between the loading lines in yellow and the products in motion through the blue dock doors towards the truck by means of the arrow.

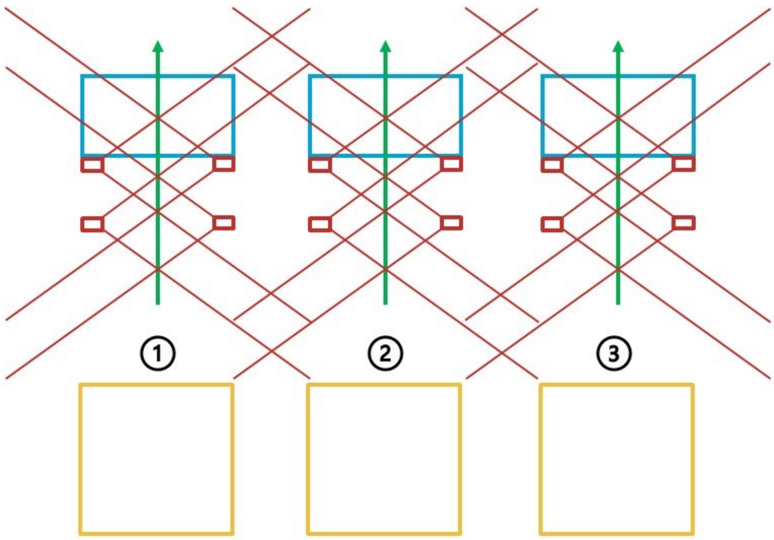


Figure 1.2: Problem of cross-reading (red: RFID antennas and reading range, yellow: static RFID tags, blue: dock doors, green arrow: possible direction loading of RFID tags

At each dock door, the RFID tags on products passing by are registered by the RFID antennas, communicating via Ultra High Frequency radio waves. Besides these moving products through the different dock doors, many other products with RFID tags are stored between the supply lines in the designated staging areas, these products are not moving and therefore called the static products. However, this form of communication through radio waves is not limited to just the dock door area where the antennas are installed. So not only the moving products with RFID tags at the respective dock door are registered, also the moving products at adjacent dock doors. In addition, nearby static products in the staging areas are also registered. This means that the RFID system not only registered the correct moving products with RFID tags as being loaded in the truck, but also the wrong moving and static products. In the case of the shown figure, this would mean that dock door 2 registers the products that are routed through dock door 2, as well as products moving though dock doors 1 and 3 and static products in proximate staging areas. This causes problems with the registration of the products, as the products are registered by multiple dock doors as being loaded into the truck, which makes it impossible to distinguish through which dock door the RFID tags went. These RFID tag reads that are not distinguishable are called **cross-reads**.

Besides the crossreads there is another phenomenon that is unwanted during the registration of RFID tags at dock doors. When the dock doors are installed with an RFID system, it might also happen that some products with RFID tags are not read by an RFID antenna while being loaded into a truck. In this way it is not registered that the product concerned has gone into the truck. When this happens, a product is missing, which in turn takes time and money to find back, which is not desirable. These RFID tag reads that are missed are called **miss-reads**.

1.2. Research gap

Dock Door Discrimination (DDD) methods are used as a solution for the problems of cross-reads and miss-reads in a multiple dock door environment, a warehouse for example. Such a DDD method aims at improving the discrimination of the RFID tags at the dock doors, which means the products can be better distinguished from each other. This means that the DDD method also contributes to the registration and give better insight in the location of the goods, i.e. if and where products are loaded on the trucks.

Various methods for RFID Dock Door Discrimination exist, such as the Satellite portal method (Keller et al., 2012), Zone Discrimination (Mojix, 2016) and Metal shielding (Ahson and Ilyas, 2017,Yuan and Yu, 2011). But because these are expensive alternatives (Krishna and Husak, 2007), the logistics industry is looking for cheaper alternatives. A preliminary literature study revealed that a DDD method based on Bayes' probability theory has the most potential to make a positive impact as a solution to cross-reads and miss-reads in a multiple dock door environment. Other studies focused on movement discrimination at a single dock door using various DDD methods, which leaves a gap for using DDD methods in a multiple dock door environment. In this study, a DDD method using Bayes theory is proposed to ensure that there is cooperation and fusion between the readings at multiple dock doors to achieve an accurate determination of location by implementing the Naive Bayes Classifier. In practice, a DDD method based on the Bayes theorem has been shown to work. However, what is lacking is a scientific foundation on which the method is based. This study will do that by providing the scientific background, a proposed design and the data gathered during experiments to measure the impact of the Naive Bayes Classifier as a DDD method in an RFID multiple dock door environment.

1.3. Research objective

The objective of this study is therefor to increase the accuracy of RFID Dock Door Discrimination with the Naive Bayes Classifier in a multiple dock door environment.

1.4. Research scope

Items that are **inside the scope** of this study:

- **Outbound process flow**
  The purpose of this study is to avoid wrong deliveries in the outbound flow of goods. The warehouse itself is only responsible for the outbound process flow, since this is the only flow of goods that the warehouse itself can influence. The inbound flow of goods depends on the shipment, so the responsibility of wrong delivery does not lie with the warehouse, and therefore will be left out of the scope during this study.

- **Radio Frequency IDentification (RFID)**
  In this study, RFID technology is used for product registration of goods in a warehouse. Other forms of product registration, for example barcode scanning, are excluded from the scope of this study.

- **Dock Door Discrimination**
  This study is concerned with distinguishing which dock door the products (with RFID tags) move through, Dock Door Discrimination. Other processes within the outbound process flow in a warehouse, such as order picking, are excluded from the scope of this study.

- **Data gathering using experiments**
  The comparison between the current RFID system for Dock Door Discrimination and the newly proposed method with the Naive Bayes Classifier shows whether the accuracy is improved. To compare the methods, a data set is needed for the application. This data set is gathered during this study with an Experimental Setup and Plan. Only this data set is used for comparison.

- **Naive Bayes Classifier**
  The new method for Dock Door Discrimination is based on the Naive Bayes Classifier. The conclusion from a previous literature review with comparison between different methods found that it is the most promising. Therefore, the other methods are excluded from the scope of this study.

Items that are **outside the scope** of this study:

- **Inbound process flow**
  Correct inbound process flow is not the responsibility of the receiving warehouse. The sender of the goods has this responsibility as it has influence in what goods leave in which trucks. Therefore, inbound process flow is kept out of scope.
- **Environment**
  The environment around the RFID Dock Door Discrimination system is not included in the scope of this study. Due to the complexity of environmental influences, this is kept out of scope.
- **Topology setup variation**
  The topology setup for the Dock Door Discrimination is a standard configuration typically used for implementing an RFID system, according to experts at Mieloo & Alexander. The topology setup is the configuration and orientation of the RFID antennas. Since this is already a standard, the topology setup variation is kept out of scope.
- **Reader settings alteration**
  The reader setting alteration are can affect the detectability of the RFID tags by the RFID system. Transit power, reading talking time, channel switching and session alternation are examples of setting of the RFID reader. Since this is too complex for the time available for this study, this is also kept out of scope.

1.5. Research questions

The main research question that is posed to support the goal of the research assignment is: **What impact has implementing the Naive Bayes Classifier on RFID Dock Door Discrimination?**

Sub-questions that support this main research question are:

1. What is the current RFID system used for RFID Dock Door Discrimination?
2. What is RFID technology and how does it relate to the Bayes theorem?
3. What design is proposed for RFID Dock Door Discrimination with Naive Bayes Classifier?
4. What experimental setup and plan are used for gathering data on Dock Door Discrimination?
5. What is the performance of the Naive Bayes Classifier compared to current RFID system?

1.6. Approach

Steps in the process of this study.

1. Process performance analysis of current RFID dock door system
2. Literature study on RFID technology and relation to Bayes theorem
3. Design of RFID Dock Door Discrimination method with Naive Bayes Classifier
4. Experiment scaled design to collect data on RFID Dock Door Discrimination
5. Comparing the performance of RFID Dock Door Discrimination with current RFID dock door system and Naive Bayes Classifier

6. Conclusion on impact of the Naive Bayes Classifier on RFID Dock Door Discrimination and Future works

In chapter 2 the current RFID system of registering products by being loaded on trucks is analysed. First, the whole process of the outbound flow of products at a warehouse is mapped, together with the current RFID Dock Door Discrimination system at the dock doors. Subsequently, the Key Performance Indicators are given, in order to be able to compare the current system performance with the new Naive Bayes Classifier system later on in this study. In doing so, it also explains which data sets are used to make this comparison between the systems.

In chapter 3 there is a deepening into the literary background of RFID technology. First of all, the functioning of RFID technology is discussed. Then it looks at related works where RFID in combination with Bayes theorem is used for other applications in the world of RFID technology. At last, it discussed why Bayes method has been chosen to solve the problem of dock door discrimination. The further continuation of this study is then based on this method.

Chapter 4 is used to explain the design of the proposed RFID system for Dock Door Discrimination with the Naive Bayes Classifier. Here a distinction is made between two aspects belonging to the design of an RFID DDD method. First, the hardware design is discussed, followed by the software design based on the Naive Bayes Classifier.

Chapter 5 explains the experimental setup and plan to collect data for comparing the DDD methods. It also explains how this data set is transformed before it is used for the Naive Bayes Classifier.

Chapter 6 compares the results in performance of the current RFID DDD system with the Naive Bayes Classifier method for Dock Door Discrimination. A conclusion can then be drawn about the impact of the Naive Bayes Classifier.

Chapter 7 provides the conclusion of this study, answering the main and sub-questions. Recommendations are also made for future work within the topic of this study.

Chapter 2

Analysis

This chapter answers the sub-question: **What is the current RFID system used for RFID Dock Door Discrimination?** In order to strengthen the reason for this research, the current system of product registration in logistics is analysed. First, the two related processes affected by this study are discussed in section 2.1. The first is the outbound process flow in a warehouse, where the introduction of RFID is explained, and later zooms in more on how the current RFID system for Dock Door Discrimination works when loading products in a multiple dock door environment. Then, the Key Performance Indicators are determined in section 2.2, which should quantify the performance of the current system and the proposed system in this study. Besides the KPIs, the way data is collected to compare the current and the new Dock Door Discrimination system is also determined. Later in section 2.3, the way to determine the accuracy of the current RFID system for Dock Door Discrimination is discussed. The margin of error involved in determining accuracy is also discussed in section 2.4.

2.1. Process

In this study, RFID technology is used to make an impact on two related processes. The first is the outbound process flow in a warehouse, and is followed by the second process, being the current RFID system for Dock Door Discrimination. The layout of such a warehouse is shown in Figure 2.1. There are a number of components that appear in all warehouse layouts. The Inbound and Outbound Docks is the loading bay, where trucks are loaded and unloaded through the outbound and inbound docks respectively. The Staging and Shipping areas are where the products that have just arrived or are about to leave are stored. In many warehouses, the two areas together are called the staging area. The Storage and Packaging Areas are where products are stored and packaged for the long term. Often no Packaging Area occurs in a warehouse, so this is referred to as the Storage Area.

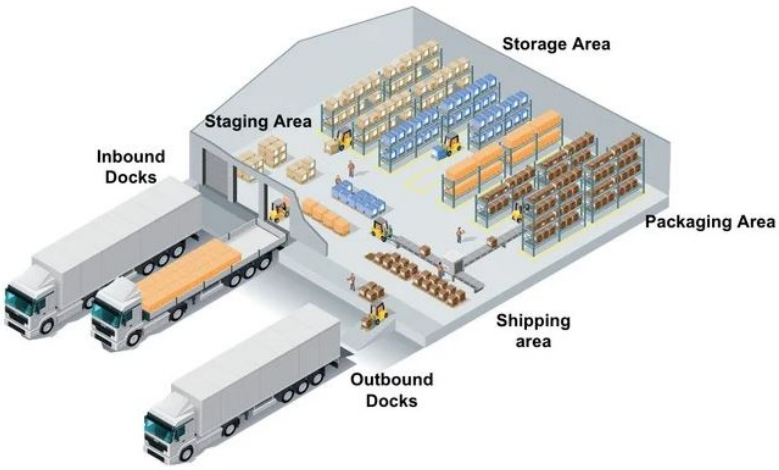


Figure 2.1: Warehouse layout (Sunol, 2022)

2.1.1. Outbound process flow warehouse

The outbound process flow is the outgoing goods flow in a warehouse, and can be seen in figure 2.2. This can occur after inbound products are initially unloaded from their truck and stored in the warehouse's storage area. After the customer orders are processed, warehouse operators receive a checklist of crates to be picked and moved to the staging area. The staging area is an area that has been specially designated for pre-processing all the products that have to end up in the designated truck. Since this process is carried out by human operators without extra checks, human errors may occur. This can result in the wrong products ending up in the truck. Also, sometimes the staging area is too small for all the products, so some pallets are placed next to it, making it difficult for the next operator to determine which staging area the pallets belong to. This is very susceptible to errors occurring in the process. As mentioned before, the products/pallets in the staging area are loaded into the truck by another operator. This operator simply ensures that all products in the staging area are neatly placed in the truck. But if an error is made at the start of the outbound process, this will count all the way through to the delivery, as there is no warning of an incorrect pallet in the meantime.

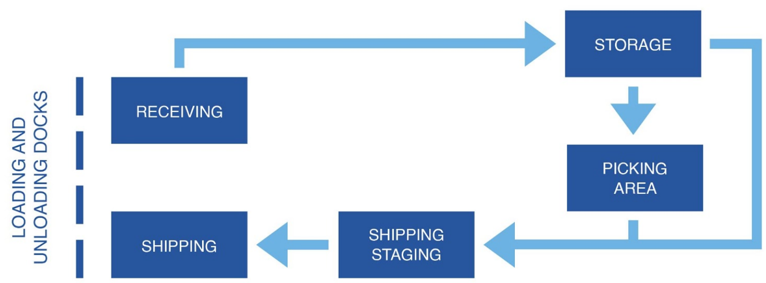


Figure 2.2: Outbound process flow (Johnson, 2022)

The impact that the use of RFID technology at the dock doors can have is to ensure that products are better registered. In this way, a dock door can be set to only accept the correct product tags. When loading the wrong product, a warning in the form of an alarm will ensure that only the right products are loaded. It also improves the speed of the process, as there is no need to keep comparing a paper checklist with the product tag in order to get the right pallets.

2.1.2. Current RFID system for Dock Door Discrimination

Zooming in on the moving of products between staging and shipping area, the current RFID transition system is reached. In the case where an RFID system is installed in a multiple dock door environment such as a warehouse, the default for Dock Door Discrimination is an RFID transition system. The layout of this RFID transition system is shown in figure 2.3. Four RFID antennas are installed per dock door, which are responsible to register the moving products provided with RFID tags (green arrow).

To determine a transition, a distinction is made between the first two antennas (light green) and the back two antennas (red). In this way, a distinction can be made between the number of reads ”at the front” of the dock door and ”at the back” of the dock door. Two factors play a major role here, the time window and the number of reads.

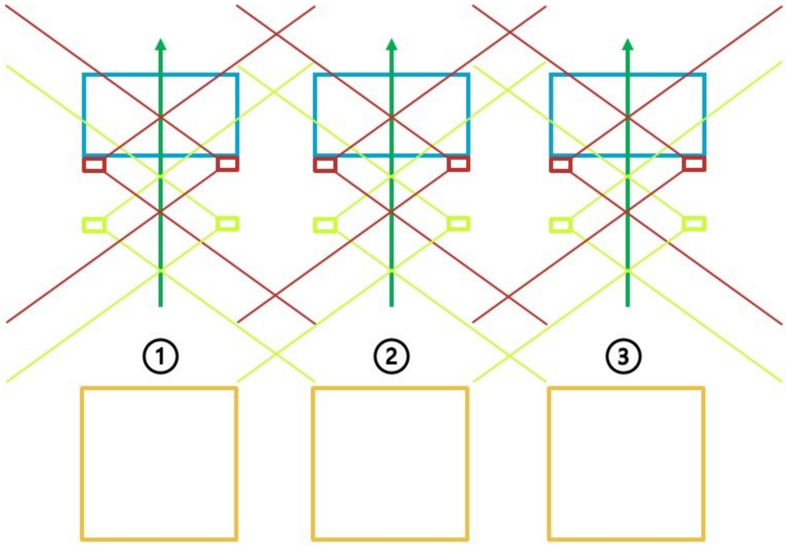


Figure 2.3: Current RFID transition system

The time window is the time used for the transition through the dock door from inside the warehouse to the truck. Here a time window of x seconds is used, this may differ according to the application of the current RFID transition system. This time window determines how long the antennas on one side of the dock door are ”on”. This means that the antennas on the front side of the dock door are ”on” for the first 2/3rds of the time window $[0\ s : \frac{2}{3}x\ s]$ and the back side for the second 2/3rds of the time window $[\frac{1}{3}x\ s : x\ s]$. This makes the middle 1/3rd of the time window the transition area, where the RFID tags are registered by both sets of antennas.

The number of reads detected at the dock door in this time window is decisive for Dock Door Discrimination. Because cross-reads occur between dock doors, the RFID is not only observed at one dock door. Practice did show that an RFID tag is most likely to have passed through a dock door where the number of reads at the front and back of the dock door is greatest. The number of reads is thus used as a filter for cross-reads. On the other hand, this does make the system susceptible to miss-reads, as the RFID tag is not considered to have moved through the dock door if there are insufficient reads on both sides of the dock door.

2.2. Performance (KPI)

In order to make the impact of the later proposed method compared to the current RFID transition system for Dock Door Discrimination quantifiable, a Key Performance Indicator (KPI) is proposed. The KPI is the accuracy of the method to assign RFID tag movements through the correct dock door and is compared by using the same data set for both Dock Door Discrimination methods. By looking at these values, it can be determined whether the implementation of the proposed DDD method is worthwhile.

Accuracy
The accuracy can be determined with the help of performance classifications. This entails that after classifying the data sets with their extracted features in different tag movement directions, a number of performance indicators are used to evaluate the results. First of all, the results are evaluated in four groups:

- True Positives (tp_i) - Correctly classified to be into a specific classification C_i
- False Positives (fp_i) - Incorrectly classified to be into a specific classification C_i
- True Negatives (tn_i) - Correctly classified to be out of a specific classification C_i
- False Negatives (fn_i) - Incorrectly classified to be out of a specific classification C_i

After the results have been divided into the four groups, the test is evaluated in terms of accuracy. This makes it possible to find out which method is the best based on the same quantity. These key performance indicator for accuracy is determined as follows (Alfian et al., 2020):

Accuracy

$$Average\ accuracy = \frac{\sum_{i=l}^{l} \frac{tp_i+tn_i}{tp_i+fn_i+fp_i+tn_i}}{l} \tag{2.1}$$

Data
To compare the accuracy of the current RFID transition system for Dock Door Discrimination with the method with Naive Bayes Classifier proposed in this study, Data of transitions of products (RFID tags) through dock doors is needed. The Data will be collected through an experimental setup and plan, which is discussed in detail in Chapter 5. The results of both DDD methods are then discussed and compared in Chapter 6 to determine the impact of the Naive Bayes Classifier.

2.3. Determining Accuracy for current RFID DDD method

This section explains how to determine the accuracy of the current RFID transition system for Dock Door Discrimination. As mentioned in section 2.1.2, the current RFID system distinguishes between the sum of the number of reads of the RFID tag on the front and back side of each dock door. When the two front RFID antennas of the dock door read the RFID tag, the reading is assigned to the ”Front” group. When the two rear RFID antennas of the dock door read the RFID tag, the reading is assigned to the ”Back” group. These reads are summed to the ”n” number of read counts in the ”Front” group and the ”n” number of read counts in the ”Back” group. In case there are three adjacent dock doors (Gates), this is done for both Gate 1, Gate 2 and Gate 3.

The sum of the number of reads ”n” is then recorded in six columns, one ”Front” and one ”Back” per Gate. To perform Dock Door Discrimination, a threshold value for the minimum number of reads is applied. When the number of reads ”n” meets this threshold value, i.e. when the number of reads is equal to or greater, the presence of the corresponding RFID tag is confirmed. When the ”Front” group and the ”Back” group of a Gate satisfy this, the name of the Gate (Gate 1, Gate 2 or Gate 3) is added to a new column ”total score”.

It may happen that the name of two or three Gates are added to this column. This means that the current RFID system thinks that the RFID tag has passed through several dock doors at the same time, in which case ”cross-reads” occur. On the other hand, it can also happen that none of the Gates meets enough reads to meet the threshold value. This means that the current RFID system thinks that the RFID tag has not passed through any Gate, in which case ”miss-reads” occur. In both cases, there is a failed Dock Door Discrimination. Examples of cross-reads, good Dock Door Discrimination and miss-reads are given in tables 2.1, 2.2 and 2.3 respectively.

Class	gate1_f	gate1_b	gate2_f	gate2_b	gate3_f	gate3_b	col1	col2	col3	total_score	Successful
Gate1	7	6	3	6	8	3	Gate1	Gate2	Gate3	Gate1Gate2Gate3	No
Gate1	12	6	2	4	8	2	Gate1	Gate2	Gate3	Gate1Gate2Gate3	No
Gate1	9	8	7	2	2	5	Gate1	Gate2	Gate3	Gate1Gate2Gate3	No
Gate1	14	7	4	0	4	2	Gate1		Gate3	Gate1Gate3	No
Gate1	9	10	2	2	0	0	Gate1	Gate2		Gate1Gate2	No

Table 2.1: Current RFID system for Dock Door Discrimination with n=1 (cross-reads)

Class	gate1_f	gate1_b	gate2_f	gate2_b	gate3_f	gate3_b	col1	col2	col3	total_score	Successful
Gate1	7	6	3	6	8	3	Gate1			Gate1	Yes
Gate1	12	6	2	4	8	2	Gate1			Gate1	Yes
Gate1	9	8	7	2	2	5	Gate1			Gate1	Yes
Gate1	14	7	4	0	4	2	Gate1			Gate1	Yes
Gate1	9	10	2	2	0	0	Gate1			Gate1	Yes

Table 2.2: Current RFID system for Dock Door Discrimination with n=6 (good dock door discrimination)

Class	gate1_f	gate1_b	gate2_f	gate2_b	gate3_f	gate3_b	col1	col2	col3	total_score	Successful
Gate1	7	6	3	6	8	3					No
Gate1	12	6	2	4	8	2					No
Gate1	9	8	7	2	2	5					No
Gate1	14	7	4	0	4	2					No
Gate1	9	10	2	2	0	0					No

Table 2.3: Current RFID system for Dock Door Discrimination with n=11 (miss-reads)

To find out the accuracy of the current RFID system for the moving tags, the column with the names of the Gates is compared with the Class. The Class is the Gate through which the RFID tag actually passed. When these two values match, there is correct Dock Door Discrimination. To indicate the correct Dock Door Discrimination, ”Yes” is then added to the column ”Successful”. In addition to the Dock Door Discrimination at the three dock doors, the static tags at Gates 1, 2 and 3 are also discriminated. Because the antennas are only pointed towards the static tags, it is easy for the current system to distinguish between moving and static tags, leading to 100% accuracy for static tags. The proportion of static tags to moving tags is s to m tags. The total accuracy for the Dock Door Discrimination with the current RFID system is then determined using the equation 2.2 below.

$$Accuracy = \left(\frac{\frac{Number\ of\ "Yes"\ values}{Total\ number\ of\ values}}{m} + \frac{1}{s}\right) * Total\ number\ of\ tags * 100\% \tag{2.2}$$

2.4. Margin of Error

A data set reflects a subset of all situations that can occur in reality. Therefore, some inaccuracy in the results must always be taken into account. To clarify this, this section first discusses the Confidence Interval. It then shows how the inaccuracy is calculated using the Margin of Error.

2.4.1. Confidence Interval

A confidence interval is the mean of the estimate plus or minus its variance. Within a certain confidence level, this is the range of values in which you expect your estimate to fall when you redo the test. Another term for probability is reliability. In this study a confidence interval is assumed with a 95% confidence level, this indicates that 95 percent of the estimates are convinced to fall within the upper and lower bounds of the confidence interval (Scharwächter, 2022), as shown in figure 2.4.

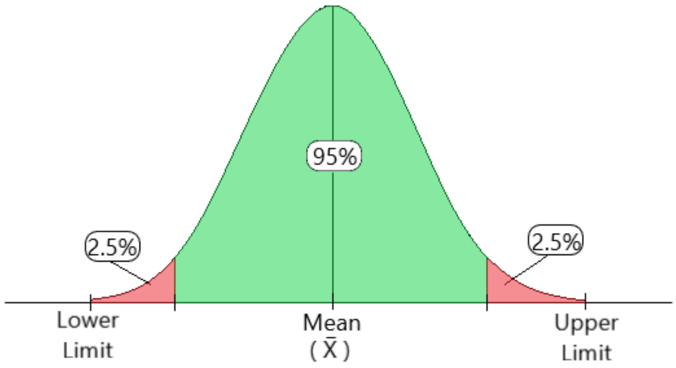


Figure 2.4: 95 % Confidence Interval (Prak, 2020)

2.4.2. Margin of Error

The inaccuracy is determined from the confidence interval using the Margin of Error. The formula for calculating the Margin of Error using accuracy proportions is given in equation 2.3 (Qualtrics, 2022). z_γ is the critical value from the Z-table (Gerstman, 2021) belonging to the selected Confidence Level, in this study it is the critical value 1.96 for a Confidence Level of 95%. p is sample proportion, in this case also called accuracy, which indicates what percentage of the measurements belong to the correct group. n is the sample size used to determine accuracy for the sample proportion.

$$Margin\ of\ Error = z_\gamma * \sqrt{\frac{p(1-p)}{n}} \tag{2.3}$$

2.5. Conclusion

In this chapter the following sub question will be answered: **What is the current RFID system used for RFID Dock Door Discrimination?**. There are two related processes considered. The first is the outbound process flow in a warehouse, this is the process of storing the products that via picking end up in the staging areas and then loaded onto the truck. The second process, when zooming in on truck loading between staging and shipping, is the current RFID system for Dock Door Discrimination. To recognise a transition, a distinction is made between the two antennas at the front of the dock door and at the back of the dock door. A time window per antenna set can be used to determine whether the product (RFID tag) has actually transitioned through the dock door. The number of reads observed at the front and back of the dock door applies as a way to prevent the cross-reads, but undesirably creates more miss-reads.

The Key Performance Indicator makes the methods quantifiable for comparison. Accuracy is used as a KPI in this study to compare the newly proposed method with the current method for Dock Door Discrimination. To do this, Data is collected throughout this study through an experimental setup and plan.

The current RFID system determines Dock Door Discrimination by setting a threshold value for the number of reads ”n”. When there are equal or more than ”n” readings at the front and back, the tag is considered to have passed through that dock door. The accuracy of the current RFID transition system for Dock Door Discrimination is determined by the number of clear transitions of a RFID tag out of the total number of RFID tags.

To determine the inaccuracy of the results, a Margin of Error is used. In this study a confidence interval is assumed with a 95% confidence level, this indicates that 95 percent of the estimates are convinced to fall within the upper and lower bounds of the confidence interval. With this information the Margin of Error can be determined.

Chapter 3

Literature

The aim of this chapter is to answer the following question: **What is RFID and how does it relate to the Bayes theorem?** This is to ensure that readers can read the information in this study and the current state of RFID technology without prior knowledge. To first understand how RFID works, the working principle, the tags and the frequency bands of RFID technology are discussed in section 3.1. Then the low-level read data that emerges when using RFID and what can be done with it is discussed in section 3.2. After that, several methods from related works are revealed in which RFID technology in combination with the Bayes theorem has been used for applications in section 3.3. Finally, the principle of the Bayes method used for Dock Door Discrimination is explained in section 3.4, on which this study will build further.

3.1. RFID

This section contains the background information about RadioFrequency IDentification to give more insight to the reader about the working principle, RFID Tags and frequency bands used in RFID technology.

3.1.1. Working principle

As the name suggests, RFID is based on radio waves. In figure 3.1, such an RFID system is shown. An RFID system consists of a number of components that are in contact with each other. A reader or transceiver, a tag or transponder and an antenna (Want, 2006). In addition, the reader is in contact with a host computer.

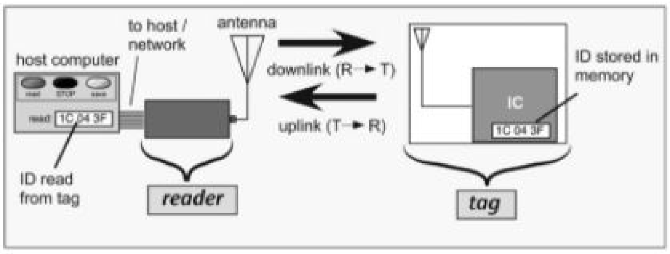


Figure 3.1: Overview RFID

The antennas are connected to the tags or the reader. In the case of the tag, the antenna is physically integrated. In addition, a tag also has an integrated circuit to provide the tag with its own identification and logic. In turn, the reader is either integrated with the antennas or is connected separately to the reader via cables. The antennas of the readers give energy to the tags via radio waves, this is called the downlink. The other way round, where the tag sends its energy together with its identification to the reader, is called uplink. The information that reaches the reader is often passed on to a connected computer. This computer is part of a communication network that makes it possible to process the data coming from the RFID system.

3.1.2. Tags

RFID tags come in three different types; active, semi-passive and passive. These different types are shown in figure 3.2 below.

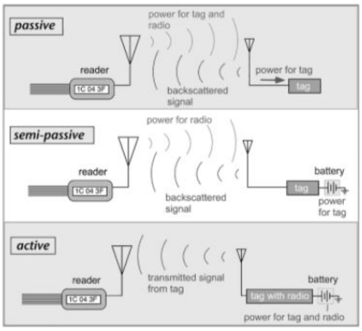


Figure 3.2: Types of RFID tags (Dobkin, 2012)

Passive tags
The passive tag does not have its own power supply or radio transmitter (Dobkin, 2012). Furthermore, it consists of a microchip for memory and logic and an antenna. The operating principle of a passive tag is based on amplitude modulation, which is discussed further in the next section. The main advantage over the active tag is the cost for a single tag, being €0.08-0.15 compared to €30 for an active tag ("How much does an RFID tag cost", n.d. This is of course due to the fact that it does not have its own power supply, and according to (Chawla and Ha, 2007) they will only get cheaper in the immediate future.

Semi-passive tags
The semi-passive tag has its own power supply, just like the active tag. The difference is that this power supply is only used for the auxililiary electronics circuit, being sensors or user-interface (Khan et al., 2009). But just like with the passive tags, the signal is reflected to the reader via the backscattering method.

Active tags
As shown in the figure, the active tag has its own energy source in the form of a battery (Weinstein, 2005). Besides this power supply, the active tag has control ciruitry and a transmitter to receive and send it (Khan et al., 2009). By using a local oscillator, the active tag can send a stronger signal to the readers and can therefore be detected from further away. This can be done in the following four ways: (Dobkin, 2012)

- Amplitude shift keying (ASK)
- Phase shift keying (PSK)
- Frequency shift keying (FSK)
- Quadrature amplitude modulation (QAM)

In addition, the energy is also used for other energy-consuming electrical components of the tag, such as sensors or user interface.

3.1.3. Frequency bands

RFID transmits radio waves with a large difference in frequency, depending on the application. The whole spectrum that can be reached with an RFID system is from 100 kHz to more than 5 GHz (Chunli and Donghui, 2012). Different applications require different frequency bands. The frequency bands that are most known for an RFID system are the low-frequency (LF), high-frequency (HF), ultra-high-frequency (UHF) and microwaves. The corresponding frequency bands are respectively 125/134 kHz (LF), 13.56 MHz (HF), 860-960 MHz (UHF) and 2.4 GHz (microwave). The working principle can be divided into two parts. First, the near-field RFID is discussed and then the far-field RFID.

Near-field
The operation of near-field RFID is based on the principle of magnetic induction, as shown in figure 3.3. An alternating current runs through the coil of the reader, creating an alternating magnetic field near the reader (Kaur et al., 2011). When a tag is held close to the reader, it captures the magnetic field through the coil of the tag. This magnetic field is converted into power for the tag through the coil and capacitor of the tag. In turn the tag generates a smaller magnetic field opposite to the magnetic field of the reader. The reader captures a slightly larger magnetic field, causing the coil of the reader to detect a larger alternating current. This difference is equal to the charge that is transferred to the tag, which is why this principle is called load modulation. Via load modulation different encodings can be passed on depending on the number of ID bits, data transfer rate and redundancy bits to remove errors from the code.

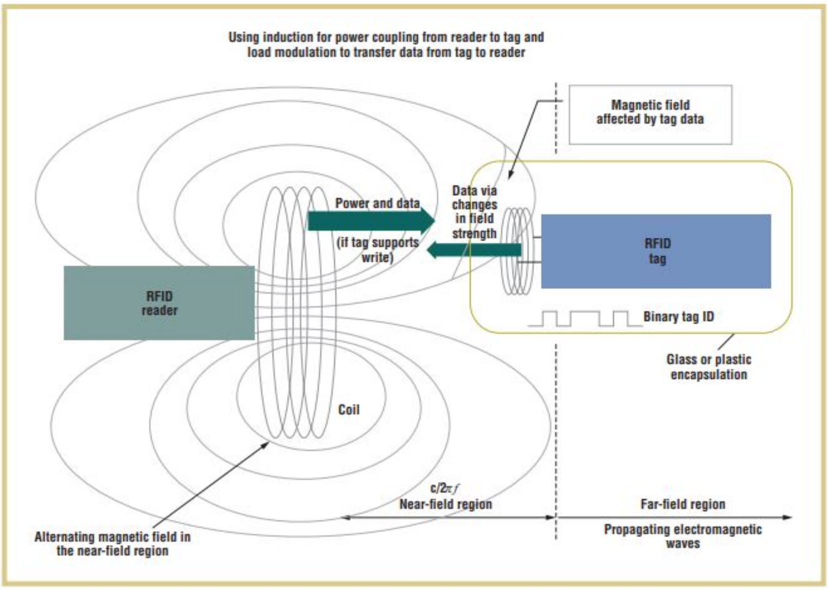


Figure 3.3: Near-field (Want, 2006)

The operating frequencies that belong to the group of near-field RFID and therefore use magnetic induction are up to 100 MHz (Kaur et al., 2011). This means that the RFID categories LF and HF fall within the near-field RFID group. This is because the range of magnetic induction depends on equation 3.1. This means that the operational distance of near-field RFID becomes smaller and smaller as the frequency increases. In addition, the energy that is extracted from the magnetic field of the reader is reduced by equation 3.2, where r is the distance between the reader and the tag. At higher frequencies, those of UHF and Microwaves, another operating principle must be used.

$$c/2 * \pi * f \qquad (3.1)$$

$$1/r^3 \qquad (3.2)$$

Far-field
In contrast to near-field RFID, far-field RFID relies on electromagnetism, as shown in figure 3.4. Here, the dipole antenna of the reader propagates electromagnetic waves that are received by the smaller dipole antenna of the tag. In the tag, the alternating potential difference is converted via the capacitor back into energy for the tag. A small fraction of this energy is then reflected back to the reader, this is

called back-scattering. To be able to reflect that signal, a so-called impedance mismatch occurs, where the frequency of the reader is not absorbed but reflected. With the help of a sensitive radio receiver, this energy is captured by the reader. By adjusting the impedance of the tag antenna, the tag ID can be encoded.

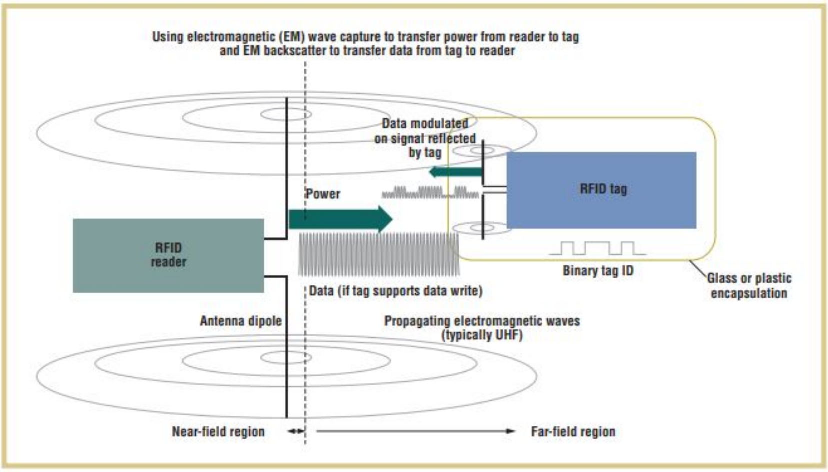


Figure 3.4: Far-field (Want, 2006)

There is a far-field RFID system when the operating frequency is more than 100 MHz (Kaur et al., 2011). This means that UHF and microwaves fall under this group of RFID systems. Because the energy first has to go from the reader to the tag and then is reflected back to the reader, a lot of energy is lost. Therefore the remaining energy is subject to equation 3.3. But since the power a tag needs and the radio receivers of the readers are getting better and better, these systems are getting better over time.

$$1/r^4 \tag{3.3}$$

Applications
The different types of RFID, their coupling type and their applications in the real world are shown in table 3.1 (Duroc and Tedjini, 2018).

	Frequency range	Read range	Coupling type	Applications
LF	125 kHz 134 kHz	~0.1 m	Magnetic Near field	Smart card, ticketing, access, animal tagging, laundry...
HF	13.56 MHz	~1 m	Magnetic Near field	Small item management, supply chain, anti-theft, library...
UHF	860-960 MHz	~2-20 m	Electromagnetic Far field	Transportation vehicle ID, access, security, supply chain, large item management...
Microwaves	2.4 GHz	~10 m	Electromagnetic Far field	Transportation vehicle ID, road toll, access, security, supply chain, large item management...

Table 3.1: RFID Characteristics and applications (Duroc and Tedjini, 2018)

3.2. Low-level read data

RFID reads, where RFID tags are read by the RFID antennas, involve low-level read data. First, the information contained in this data is explained. Then how this is used to determine the direction of the RFID tag. Finally, the low-level read data is used to find out extracted features, which are needed for further research with Dock Door Discrimination.

3.2.1. Low-level read data

During a tag read, the reader's antenna receives information that depends on a variety of factors. These factors consist of both the underlying information that belongs to a tag read and the environmental factors that can influence it. The figure shows an example of such a tag read.

EPC	Timestamp	Antenna	RSSI	Phase angle
3032...7D	1,453,989,765.31	15	-59.0	3.50
3032...D1	1,453,989,765.31	15	-56.0	2.91
3032...7D	1,453,989,765.34	4	-69.0	2.72
3032...7D	1,453,989,765.34	17	-56.0	3.07

Table 3.2: Example of lower-level features from tag read (Hauser et al., 2019)

EPC
The Electronic Product Code (EPC) indicates the identity of the corresponding tag (Keller et al., 2010). This code is divided into four parts to provide each tag with its own identity; Header, EPC Manager, Object Class and Serial Number. The header indicates the type of EPC format used by the tag. The second part of the number is the EPC Manager, which indicates the manufacturer's company. Then there is the Object Class, which indicates the product class to which the tag belongs. Finally, there is a Serial Number, which is unique for each tag.

RSSI
The Received Signal Strenth Indication (RSSI) is a measure of the strength of the signal received from the tag by the antenna of the reader (Keller et al., 2014). This value is expressed in dBm and the closer the tag gets to the reader, the stronger the signal and therefore the higher the RSSI.

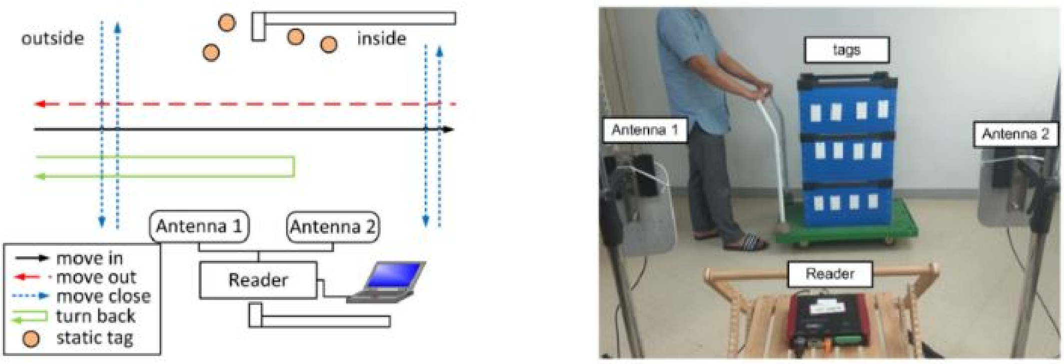
Timestamp
The value of Timestamp is determined by the time the tag is read (Hauser et al., 2019), this can be used to determine the difference in RSSI of a tag over a certain period of time. This value is sometimes indicated with SinceStart in order to make the results more comprehensible for the reader, this means the time from the start of the gathering cycle.

Antenna
A reader is connected to one or more antennas, depending on the application. To get a better idea of the position of the tag in relation to the antennas, it is indicated by which antenna the tag is read. By combining this information with for example the timestamp, one gets a better idea of the tag's completed path.

Phase angle
In addition to the strength of the signal, the phase angle of the tag read is also retrieved when it is detected (Hauser et al., 2019). This means the current phase angle of the backscattered sinusoidal wave. This value is often used in combination with RSSI to get an accurate determination of the tag's location (Buffi et al., 2019).

3.2.2. Tag Direction

A distinction is made between the tags based on their direction. In (Alfian, Syafrudin, Yoon, et al., 2019) a distinction is made between tag movements through the gate, near the gate and static tags. On this basis, in (Alfian, Syafrudin, Lee, et al., 2019) on further distinction in movements through the gate, namely outward (towards the truck) and inward (towards the warehouse). (Alfian et al., 2020) is the most recent study on tag movement direction. In this study, an additional movement pattern is added, namely turning around halfway through the gate, as shown in figure 3.5a.
The test set-up, shown in figure 3.5b, consists of a reader with two antennas next to each other, through which a stack of crates and tags are led. The reader antennas then transmit a UHF radio wave and receive a back-scattered signal back from the tags. This signal is then decomposed into all kinds of low-level features, previously indicated in the previous section.



(a) Schematic (b) Test setup

Figure 3.5: Tag movement detection (Alfian et al., 2020)

3.2.3. Extracted features

Based on the low-level features of a tag read, discussed in section 3.2.1, certain relations can be determined, so-called extracted features (Ma et al., 2018). In table 3.3 lists these features and their description.

Features	Description
Max RSS	The maximum of the received signal strength during a gathering session. When the moved tags move through the portal, they tend to have a larger RSS because the distance to the antenna is closer than the static tags.
Mean RSS	The average of the received signal strength during a gathering session. When the tags move through the portal, they tend to have a larger mean RSS because the distance to the antenna is closer than the static tags during most of the gathering session.
RSS variance	Because the distance of the moved tags to the antenna changes continuously when the moved tags pass through the portal, they tend to have a larger variance.
RSS range	The difference between maximum received signal strength and minimum Received signal strength collected during a gathering session.
RSS Skewness	Represents asymmetry of the RSS's histogram with respect to its mean value. Static tags tend to have a symmetrical RSS distribution.
RSS Kurtosis	Measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution. Moved tags tend to have a heavy-tailed RSS distribution.
Phase variance	Because the distance of the moved tags to the antenna changes continuously, they tend to have a larger variance.
Phase range	The difference between the maximum and minimum measured phase values during a gathering session
Phase Skewness	Asymmetry of the phase's histogram with respect to its mean value. Static tags tend to have a symmetrical phase distribution.
Phase Kurtosis	Measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution. Moved tags tend to have a heavy-tailed phase distribution.
Count	The total number of interrogation times during a gathering session. Moved tags tend to have a higher number of reads because they pass through the portal and are closer to the antenna.

Table 3.3: Extracted features (Ma et al., 2018)

When looking at a tag that is moving passed two separate antennas, the ideal RSSI distribution would be as given in figure 3.6.

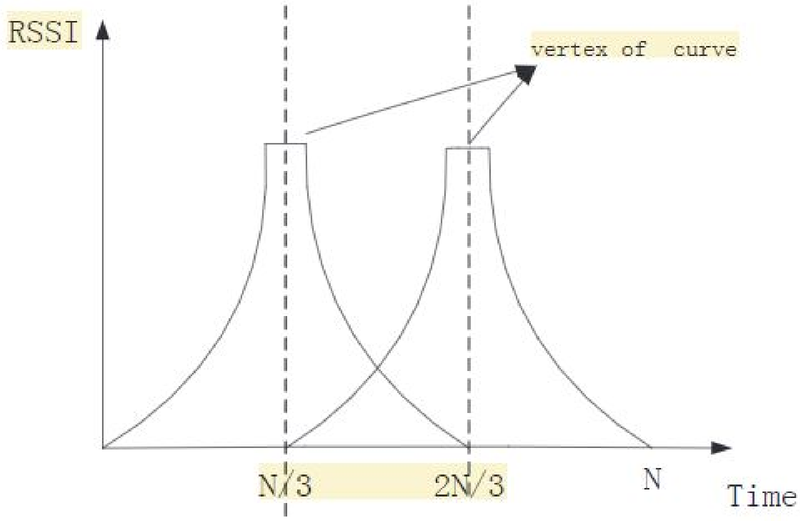


Figure 3.6: RSSI detection (Jie et al., 2018)

There are two extracted features which are not based on the distribution of the RSS or phase, the skewness and kurtosis, which are calculated as follows:

$$kurtosis = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2)^2} \tag{3.4}$$

$$skewness = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2})^3} \tag{3.5}$$

In figure 3.7 an example is given for the RSS and phase distribution and histograms of RFID reads from moved and static tags. It can be seen in figure 3.7a that the value for Max RSS, Mean RSS, RSS variance and RSS range will be the highest for the moved tags. This can be explained by the fact that the RSS value of the tag increases as it gets closer to the reader and decreases as it moves away from the reader. In figure 3.7b a clear distinction can be made between the moving and static tags, since the phase values change with the changing distance between tag and antenna. Because of this, the phase variance and range will also be larger for the moving tags than for the static tags. Figures 3.7d and 3.7f furthermore show to be less symmetrical and heavy-tailed distribution compared to the static tags of figure 3.7c and 3.7e, because there is a larger distribution in different tag read values.

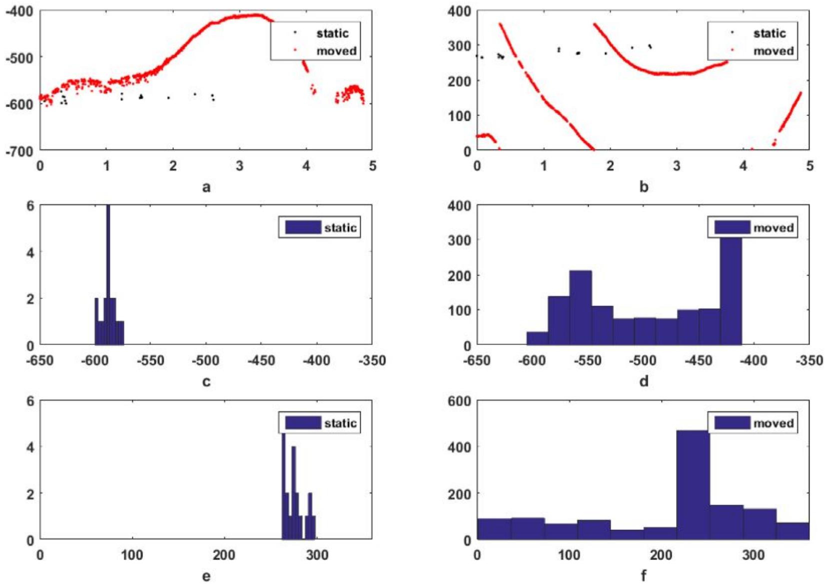


Figure 3.7: Example of extracted features (a) RSS distribution (b) Phase distribution (c) RSS histogram static tag (d) RSS moved tag histogram (e) Phase histogram static tag (f) Phase moved tag (Ma et al., 2018)

3.3. Related works on RFID with Bayes theorem and probability

This section discusses how methods based on the Bayes theorem (section 3.3.1 and 3.3.3) and other probabilistics (3.3.2 and 3.3.4) are applied in RFID technology. It discusses the different purposes that RFID has and how the Bayes theorem or other probabilistics ensures that the RFID system works properly. This can be divided into process models, tracking, localization and inference.

3.3.1. Process model localization

A first model applying Bayes theorem to an RFID application has been proposed by (Goller, 2013) and (Goller and Brandner, 2012). Here the location where the RFID tags are read is combined with the business process model of the supply chain dynamics. Using the Coninuous Time Morkov Chain as probabilistic framework (or Hidden Markov model in (Goller and Brandner, 2011)), first the probability of transition between business-process states is determined. For this purpose the temporal behaviour in terms of a dwell time parameter for every state is considered. It is then assumed that the dwell time in a given state is exponentially distributed with a mean. This gives an insight into the flow of all business processes. The probability that a read event z_j has subsequently led to a correct state transition is computed by Bayes law in the following way:

$$P(t; Detection|z_j) = \frac{P(z_j|t; Detection)xP(Detection)}{P(z_j)} \tag{3.6}$$

Thus, the model can estimate the location of a product in the supply chain at a given point in time.

3.3.2. Tracking

The tracking of RFID tags is not only done on the basis of the probability that a tag is correctly read at the current moment. (Kelepouris et al., 2011) propose a model that can accurately determine the future location of the tags based on past and present locations. Given a location probability distribution (based on Hidden Markov Model), the future location distribution can be predicted in the following way:

$$P(X_{t+1}) = \sum_{x_t/inS} P(X_{t+1})p(x_t) \tag{3.7}$$

3.3.3. Indoor localization

Another application of the Bayes theorem in RFID is an indoor positioning method proposed by (Xu et al., 2017). Here the location, together with the accuracy of this location determination, is determined by a Gaussian filter for abnormal RSSI values and Bayes probability in combination with a k-Nearest Neighbor algorithm. This results in a good improvement in the accuracy of the indoor positioning of the RFID tags. The experimental setup for the determination of the tag locations is shown in figure 3.8.

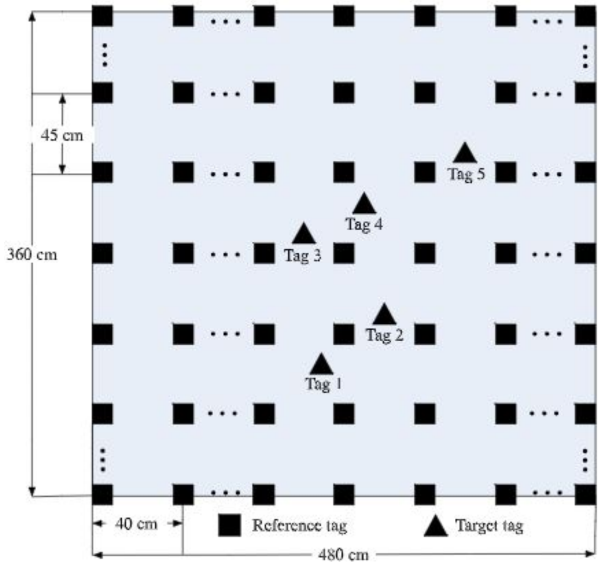


Figure 3.8: Indoor localization (Xu et al., 2017)

With Z_k representing the sets of all measured tags k coordinates and X_k being the unknown tag position, the posteriori probability of X_k being under the known set Z_k is given by $p(X_k|Z_{k-1})$ and the prior probability of the position estimation of the unknown tag in the unknown set by $p(X_k|Z_{k-1})$. These are used for determining the probability distribution of $p(X_k|Z_k)$ by means of a Bayesian positioning algorithm (Xu et al., 2017):

$$p(X_k|Z_k) \propto p(Z_k|X_k) \times p(X_k|Z_{k-1}) \tag{3.8}$$

1. Determine prior probability: $p(X_k|Z_{k-1})$

$$p(X_k|Z_{k-1}) = \frac{1}{\sigma_1\sqrt{2\pi}} \exp -\frac{D_1^2}{2\sigma_1^2} \tag{3.9}$$

2. Determine posteriori probability: $p(X_k|Z_{k-1})$

$$p(X_k|Z_{k-1}) = \frac{1}{\sigma_2\sqrt{2\pi}} \exp -\frac{(D_2 - Q)^2}{2\sigma_2^2} \tag{3.10}$$

3. Calculate probability distribution function $p(X_k|Z_k)$ of unknown tag's position.

$$p(x, y) = \prod_{i=1}^{4} p_i(X_k|Z_k) = C(\frac{1}{\sigma_1\sigma_2\sqrt{2\pi})^4 \exp f(x, y)} \tag{3.11}$$

where,

$$f(x, y) = \Sigma_{i=1}^4 f_i(x, y) = -\frac{1}{2}\Sigma_{i=1}^4\left\{\frac{1}{\sigma_1^2}\left[(x - x_{k-1})^2 + (y - y_{k-1})^2\right] + \frac{1}{\sigma_2^2}(\sqrt{(x - a_i)^2 + (y - b_i)^2} - q_i)^2\right\} \tag{3.12}$$

3.3.4. Inference

Another application of the probabilistics in RFID is proposed by (Tran et al., 2009) to determine reader mobility, object dynamics and noisy readings. This mainly concerns a probabilistic data generation model. In this research, a distinction is made between evidence variables, which are observed from the collected data, and hidden variables, which reflect the actual locations of the reader and the objects. These are represented by the shaded and unshaded nodes respectively in figure 3.9. The aim of the proposed model is to predict the locations of the various objects using the conditional distribution of p.

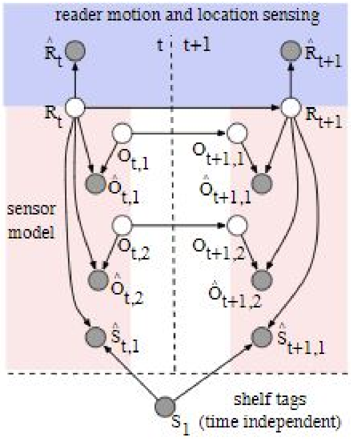


Figure 3.9: Bayesian network model

In order to make a good prediction, it is necessary to divide it into smaller models. Thereby, the reader motion model and reader location sensing model are used to make a good prediction for the evidence value R. In addition, the RFID sensor model in combination with the object location model is used to find out the evidence values for O. The working principle with the different models is summed up as follows:

1. Generate new reader location R_t by sampling the previous location R_{t-1} with the reader motion model $p(R_t|R_{t-1})$.
2. Generate noisy reader location observation R_t from reader location sensing model $p(\hat{R}_t|R_t)$.
3. Generate new object locations O_t from the object location model $p(O_{ti}|O_{t-1,i})$.
4. Determining if object i is observed using the sensor model, each object with a probability $p(\hat{O}_{ti}|R_t, O_{ti})$.
5. Determining if tag is observed using the sensor model, each tag i with a probability $p(\hat{S}_{ti}|R_t, S_t)$

When all probabilities of the different models are combined, one arrives at the following equation for the local probability distribution:

$$P(R, \hat{R}, O, \hat{O}|S) = p(R_1, O_1) \prod_t p(R_t|R_{t-1}) p(\hat{R}_t|R_t) \times \prod_{i \in O} p(O_{ti}|O_{t-1,i}) p(\hat{O}_{ti}|R_t, O_{ti}) \prod_{i \in S} p(\hat{S}_{ti}|R_t, S_t) \tag{3.13}$$

3.4. Bayes theorem for RFID dock door discrimination

The application of the Bayes theorem for dock door discrimination in RFID technology is identified as promising from prior literature research. The method is briefly explained in this one from a white paper, which unfortunately lacks the scientific background.

3.4.1. Bayes method

The Bayes method (Reva, 2006) revolves around the concept of filtering the crossreading with their Location Virtualization (LV) logic. The logic bases its findings on, among other things, the spatial relationships between antennas and locations, which are based on antenna placement and orientation. In addition, the read rates and tag observations of the antennas are also used. It is also possible to use external sensors for the direction of the tags and tag association information to verify the pallets. The location determination of the tags by the LV logic is based on Bayes likelihood theorem (Joyce, 2003) as follows:

$$P(Loc < T,x,t > |Obs < T,R,t >) = \frac{P(Obs < T,R,t > |Loc < T,x,t >)xP(T \in x)}{P < T,R,t >} \tag{3.14}$$

where,

- $(Obs < T,R,t >)$: Tag T observed by reader R at time t
- $(Loc < T,x,t >)$: Tag T in location x at time t
- $(T \in x)$: Tag T in location x
- $P(T \in x)$: Prior likelihood estimate of the tag T in location x
- $P(Obs < T,R,t > |Loc < T,x,t >)$: Likelihood of the tag T being observed by reader R when the tag is in location x at time t. This information is computed based on the RF link settings used by the reader during that read cycle.
- $P < T,R,t >$: The probability that the reader R observed the tag T at time t.

The probability that a tag is in a specific location is calculated by the observations of the multiple readers together. For this, the reader uses both the spatial and temporal observations of the tag.

Example
An example of a practical application is shown in figure 3.10. It can clearly be seen that there are three dock doors in a row, close to each other, so cross reading can occur here. At the dock doors, two pairs of antennas are directed towards each other per doorway. Per dock door, the antennas are connected to a single reader. There are also photo sensors on both sides of the doors. In addition, there are also four antennas directed towards the passageway locations to register the products that are put away.

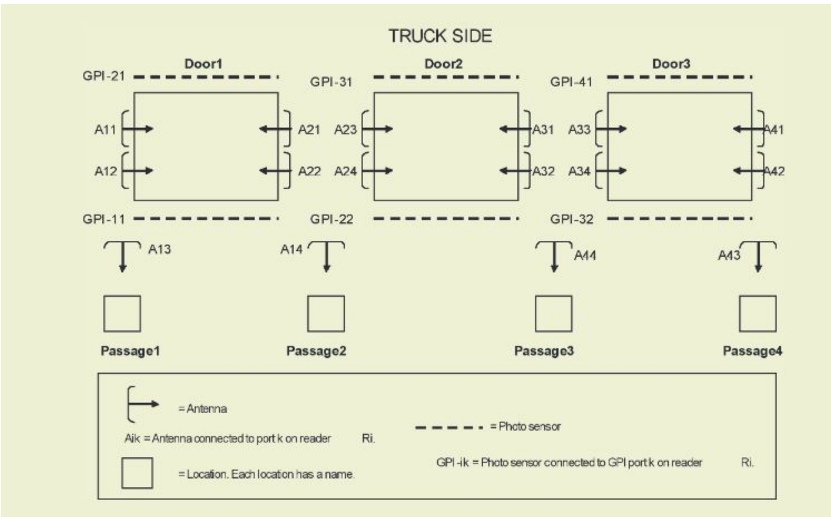


Figure 3.10: Bayes method (Reva, 2006)

The intention is that the tag pallets that are moved through a dock are also categorised as such by the algorithm. Thus, a distinction must be made between the movement directions, either loading or unloading of the truck. During this loading and unloading process, the algorithm uses probability theory to filter out crossreads. The system worked while also mimicking external influences. These include tag activity at another dock through portals, tag activity in the passage way and tags still in the truck.

3.5. Conclusion

This chapter answers the question posed in the intro: **What is RFID and how does it relate to the Bayes theorem?** RFID systems consist of a reader, tag, antenna and host computer. The reader gives energy to the tag (if passive) by radio waves (downlink), the tag gives energy back together with its identification (uplink). A distinction is then made between three types of tag: active, passive and semi-passive. The active tag has its own power supply, the passive none and the semi-passive only for the auxiliary electronics circuit. The coupling method depends on the frequency of the operation. Low frequency and high frequency fall under near-field RFID, which is based on magnetic induction. Ultra-high frequency and microwaves, on the other hand, fall under far-field RFID, where the working principle is based on electromagnetism.

Low-level read data is retrieved at a tag read, this is subdivided into EPC, RSSI, Timestamp, Antenna and Phase angle per tag read. When the tags are detected, the direction of movement of the tags needs to be determined. Also, based on the low-level read data, extracted features are determined in order to be able to recognise the patterns and distinguish between different tag reads.

Related works methods based on Bayes theorem and other probablistics are applied with RFID technology. Four methods are described. First the Process model localization method, which can estimate the location of a product in the supply chain at a given point in time. Then the Tracking method, which can accurately determine future location of the tags based on past and present locations. The indoor localization method determines the location of the target tags by the use of reference tags and by a Gaussian filter for abnormal RSS values and Bayes probability in combination with a k-Nearest Neighbor algorithm. The method of Inference combines multiple models in order to discribe the local probability distribution for the RFID tags.

The Bayes method is the foundation of this research because it has the most potential for a Dock Door Discrimination method. Using Bayes' theorem, it is then determined at which location or through which door the RFID tags have passed. The probability that a tag is in a specific location is calculated by the observations of the multiple readers together. For this, the reader uses both the spatial and temporal observations of the tag. Using this practical example as a starting point, this study examines the further optimization of the detection and discrimination of RFID tags at the dock doors in the following chapters.

Chapter 4

Design

This chapter answers the sub-question: **What design is proposed for RFID Dock Door Discrimination with Naive Bayes Classifier?** To answer this question, the proposed design for Dock Door Discrimination with the Naive Bayes Classifier (NBC) is determined step by step. After the problem definition and analysis of the current RFID system are determined in Chapter 1 and Chapter 2 respectively, the first section 4.1 determines the requirements for the design, both functional and non-functional. Then, in the system design sections, the design of the Dock Door Discrimination method is explained. First the hardware design in section 4.2 and later in section 4.3 the software design, which is based on the Naive Bayes Classifier. Then, in section 4.4, the configuration to collect data and transformation of the collected data are discussed. The last section 4.5 of this Chapter describes how the performance of the design will be evaluated, which is based on the KPI accuracy given in Chapter 2.

The goal of the design is to increase the accuracy of the current RFID transition system for Dock Door Discrimination. The new Hardware and Software design have to make this possible, without the use of extra non-RFID hardware.

4.1. Requirements

To find out whether the Naive Bayes Classifier design of the Dock Door Discrimination method meets expectations, this section provides requirements that the design should meet. The functional requirements are set out first, followed by the non-functional requirements.

4.1.1. Functional requirements

The functional requirements of the design should ensure that the design ultimately gets the features that reflect the intended functionality. For this design, the main functionalities have been converted into six requirements (between brackets for changed/additional requirements for scaled design):

- **Standard dimensions (scaled)**
  The dimensions used in this study to simulate the dock doors are the standard dimensions. The standard dimensions are 2.44 x 2.74 m (8 x 9 ft) and the distance between the center lines of the dock doors is at least 3.70 m (12 ft) (Stertil, 2022), otherwise multiple trucks cannot be loaded at once. (The dimensions used in this study to simulate the dock doors are the standard dimensions on a scale of 1:2. The standard dimensions are then 1.22 x 1.37 m and the distance between the center lines of the dock doors is at least 1.85 m, which means the space between adjacent dock doors is 0.63 m.)
- **Standard antenna configuration (scaled)**
  The antenna configuration was determined in consultation with experts from Mieloo & Alexander as it is applied in reality, to avoid too much complexity. This involves a standardised antenna configuration at the following points:
  - Positioning (4 antennas at dock door, 1 at staging area)
  - Orientation (All antennas are under a 45 degree angle)
  - Number of antennas (Per dock door 5 antennas, which comes to 15 antennas in total)
- **Dock door discrimination by software**
  The principle of dock door discrimination should eventually be achieved via software, specifically a Bayes classifier to distinguish the RFID tags from each other.

- **No extra hardware**
  Additional hardware is not required for the design of the dock door discrimination method based which is based on Bayes theorem.

- **Moving to static**
  Besides distinguishing which dock door the RFID tags move through, the design should also be able to distinguish moving tags from static tags.

- **Better performance than current RFID system**
  The design should ultimately ensure that the performance of dock door discrimination using this method improves in performance of the currently used RFID system for DDD.

4.1.2. Non-functional requirements

Non-functional requirements ensure that functional requirements are enabled. In this case, these are the requirements to enable the design. A distinction is made here between the following non-functional requirements (between brackets for changed/additional requirements for scaled design):

Non-functional requirements (scaled)	Description (scaled)
Zebra FX9600 readers	Type of RFID reader used in the design.
Mojix UHF RFID antennas	Type of RFID antennas used in the design.
Antenna cables	Cables used to connect the RFID antennas to the RFID reader.
RFID tags	RFID tags to identify the products.
Internet connection	Internet connection required for the design to function.
Node-RED	Application used to convert the Tag Data Events into a CSV file.
Python	Application to append the Bayes theorem for dock door discrimination to the data in the CSV file.
Pallets (boxes)	Pallets for transporting goods. (Boxes for transporting RFID tags)
Forklift (carriers)	Forklift for transporting pallets. (Carriers for transporting boxes)
(ISB-profiles)	(ISB-profiles are used to make the structure from which the RFID antennas are hung to simulate the correct dimensions.

Table 4.1: Non-functional requirements (Non-functional requirements for scaled design)

4.2. System Design - Hardware

The System Design - Hardware section looks at the overall entity to be designed. Therefore the location and layout of the design are discussed in this section. First on an overall design level and later on a detailed level.

4.2.1. Overall design

A Dock Door Discrimination method starts with the hardware design of the RFID system. This RFID system must then be integrated into the layout of the warehouse. An example of a warehouse design is shown in figure 4.1. Here, the multiple adjacent dock door can be seen, this is where the trucks are loaded. Furthermore the staging areas and the picking area are shown, along with people moving around with forklifts and manual stackers to transfer the products to the designated area.

The main objective of the hardware system design is to detect as many RFID tags as possible, so that no inconveniences are caused at a later stage. Besides detecting the tags, the design must also take into account that large warehouses have multiple consecutive dock doors, which makes it harder to discriminate the tags. The fact that the staging areas of the products to be loaded into the truck are close to the dock doors also makes it necessary to devise a way to discriminate between moving tags and static tags in the proposed hardware system design.

The layout of the design to avoid the above phenomena is shown below in Figure 4.2 and is inspired on the method in section 3.4 based on Bayes' theorem. This figure shows the proposed hardware system design of a dock door loading area. It shows the loading area has 15 consecutive dock doors,

Figure 4.1: Warehouse dock doors (Mecalux, 2018)

which makes it required to have a good software system design for dock door discrimination, which is discussed in the next section of this Chapter. The trucks are shown so that they connect directly to the dock doors and can be loaded with products.

At the dock doors, indicated by the cross-shaped symbol, there are four RFID antennas per dock door, shown with red squares. These four antennas are used to distinguish the passing RFID tags in different directions of movement. Different than the current RFID transition system, also one antenna per dock door is added per dock door, on the other side of the passage way, facing the staging area, indicated by yellow. This antenna registers the static tags in the staging areas. This allows better differentiation between the moving RFID tags through the dock doors and the static RFID tags in the staging area, which ensures that now only a Dock Door Discrimination method is needed for the adjacent dock doors. The larger storage, which counts as the picking area for the staging area, is shown in orange.

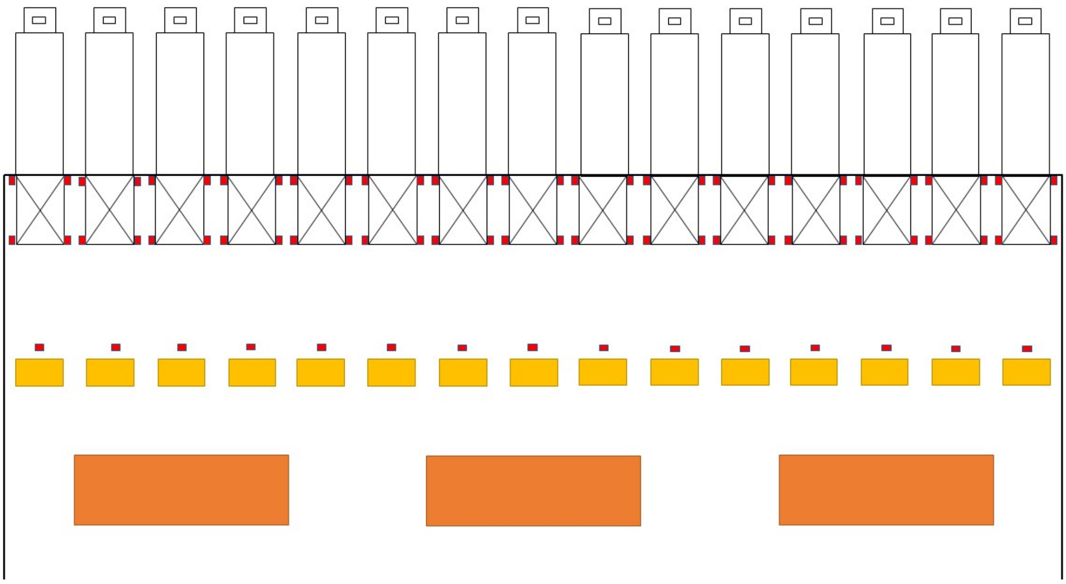


Figure 4.2: Warehouse layout design (red: RFID antennas, yellow: staging area, orange: picking area)

4.2.2. Detailed design

When zooming in more on the overall version of the hardware system design to the detailed design, only three dock doors are considered. This is done because further in this study in Chapter 5 data is collected from a dock door system of three consecutive dock doors to compare the current and the proposed Dock Door Discrimination method. Through this detailed design, any possible situation at a dock door can be simulated. These are the situations where there is another dock door on the left, right or both sides of the dock door.

For the detailed system design, shown in figure 4.3, the standard sizes are considered. This means the dimensions for a dock door are 2.44 m in height and 2.74 m in width, according to (Stertil, 2022). The distance between the centre lines of adjacent dock doors is at least equal to 3.70 m, which means the standard dimension for the space in between two adjacent dock doors is 1.26 m. The figure shows the front view of the three adjacent dock doors in the detailed system design, with the antennas indicated in red and the incoming and outgoing products with RFID tags indicated by the purple cross and dot symbol. By default, the antennas are at a 45-degree angle, as this is the conventional way that RFID systems are implemented in the real world, according to RFID experts at Mieloo & Alexander.

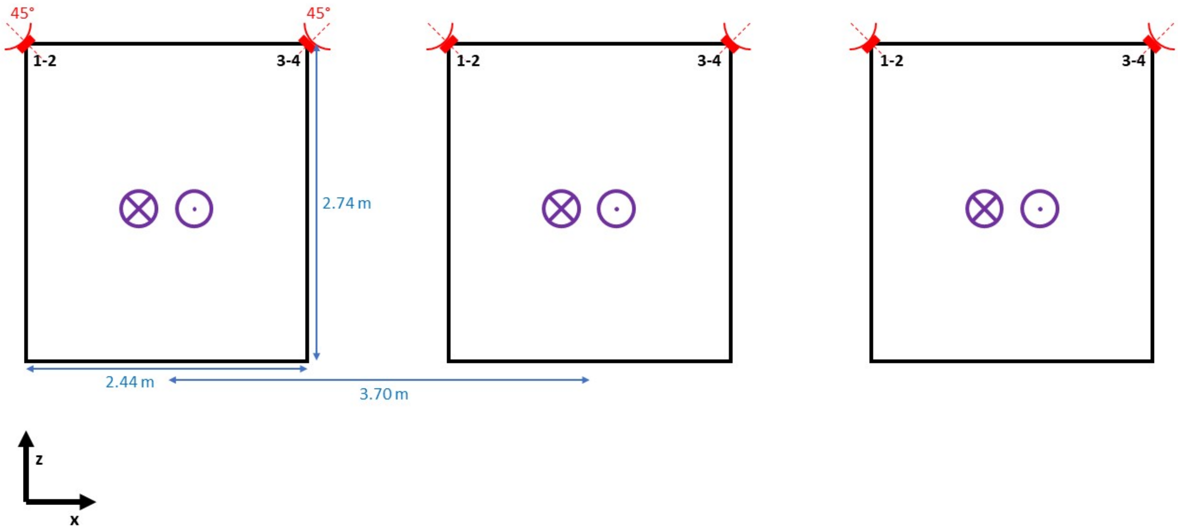


Figure 4.3: Basic configuration and dimensions, front view

Figure 4.4 shows the top view of the detailed system design. Again the dimensions are shown, displaying a width of 2.44 m and the distance between the dock doors 1.26 m. Furthermore, the direction detection of the products (RFID tags), again shown in purple, is done by means of four antennas, to create a high accuracy. For this a margin of 2.5 m (Armo, 2018) between antennas 1-3 and 2-4 are given at the dock doors, which is the standard dimensions for the dock leveller in order to load the products into the truck. Then there is the one antenna installed as an improvement compared to the current RFID transition system towards the staging area, where the static products (RFID tags) are stored. The storage of the static products is shown by means of the yellow lining. These are positioned at 4.57 m from the dock door RFID system to provide enough space for vehicles such as forklifts to move through the passage way and straight into the dock doors to load the trucks (Nova, 2022), which is represented by the green arrows.

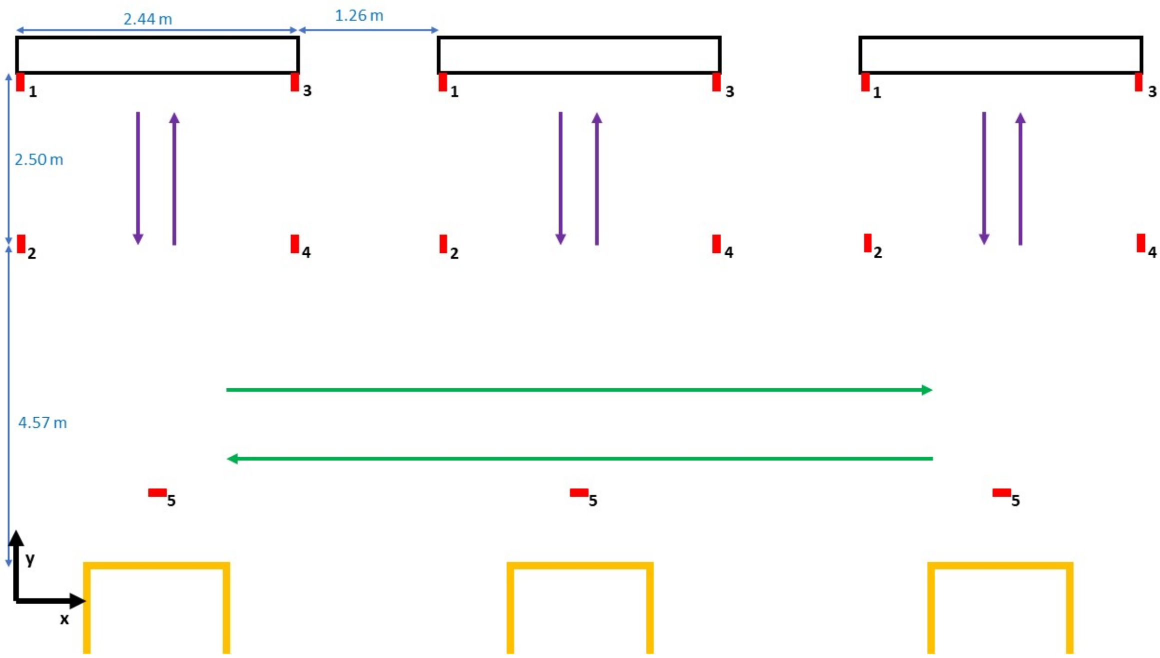


Figure 4.4: Basic configuration and dimensions, top view

4.3. System design - Software

Besides the hardware design, a software design is also needed in the proposed Dock Door Discrimination method. For the software design in this section, first the working principle of the Naive Bayes Classifier (NBC), the foundation of software design, is discussed. Later the implementation of the NBC in the software design is given.

4.3.1. Working Principle Naive Bayes Classifier

The Naive Bayes Classifier is a form of machine learning that uses the Bayes theory in order to predict to which class the data points in a data set belong. The classes in this study entail the dock doors through which the RFID tags move. The advantages of Naive Bayes are that it is easy to use and the Classifier only needs one training data set to generate the class probability (Farid et al., 2014).

The assumptions to be made with the Naive Bayes Classifier is that the Classes are independent and equal. This means that there is no dependency between the different features used in the NBC. This is quite possible in the case of Dock Door Discrimination since the features for Read Counts, RSSI and Phase of the reads are independent of each other. All features also contribute equally to the final outcome since no features are considered irrelevant.

Bayes theory underlies the Naive Bayes Classifier. When there is a dataset $D = \{X_1, X_2, ..., X_n\}$ containing features X_i with classes $C_i = \{C_1, C_2, ..., C_m\}$, the Bayes theorem applies according to equation 4.1. In this case, the probability that X belongs to a particular class C_i is calculated. For the classes, the different dock doors the tags passed through are used. X is the feature implemented to base the Dock Door Discrimination on.

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \tag{4.1}$$

where $P(C_i|X)$ is the Maximum posteriori hypothesis, $P(X|C_i)$ is the likelihood and $P(C_i)$ and $P(X)$ are the probabilities of the class and the feature respectively. Since the assumptions are there of independent and equal features for the Naive Bayes Classifier, makes that the likelihood $P(X|C_i)$ of the feature X in a given Class C_i can be computed with equations 4.2 and 4.3.

$$P(X|C_i) = \prod_{k=1}^{n} P(x_k|C_i) \tag{4.2}$$

$$P(X|C_i) = P(x_1|C_1) \times P(x_2|C_2) \times \dots \times P(x_n|C_n) \tag{4.3}$$

Since the features contain continuous-valued attributes, the data set is assumed to have a Gaussian distribution with a mean μ and standard deviation σ, defined respectively by the following two equations:

$$P(X|C_i) = g(x_k, \mu_{C_i}, \sigma_{C_i}) \tag{4.4}$$

$$g(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-p)^2}{2\sigma^2}} \tag{4.5}$$

In equation 4.4, μ_{C_i} is the mean and σ_{C_i} is the standard deviation of the values of the features for all training instances in the class C_i. To predict the class label of instance $X, P(X|C_i)P(C_i)$ is evaluated for each class $C_i \in D$. The Naive Bayes classifier predicts that the class label of instance X is the class C_i, if and only if

$$P(X|C_i)P(C_i) > P(X|C_j)P(C_j) \tag{4.6}$$

In Equation 4.6, $1 \leq j \leq m$ and $j \neq i$. That is the predicted class label is the class C_i for which $P(X|C_i)P(C_i)$ is the maximum probability.

An example of classification using the Gaussian Naive Bayes Classifier is shown in figure 4.5. At each data point in this case, a z-score is given for each Class that is available. This z-score is based on the distance between that data point and the class-mean divided by the standard deviation of the class.

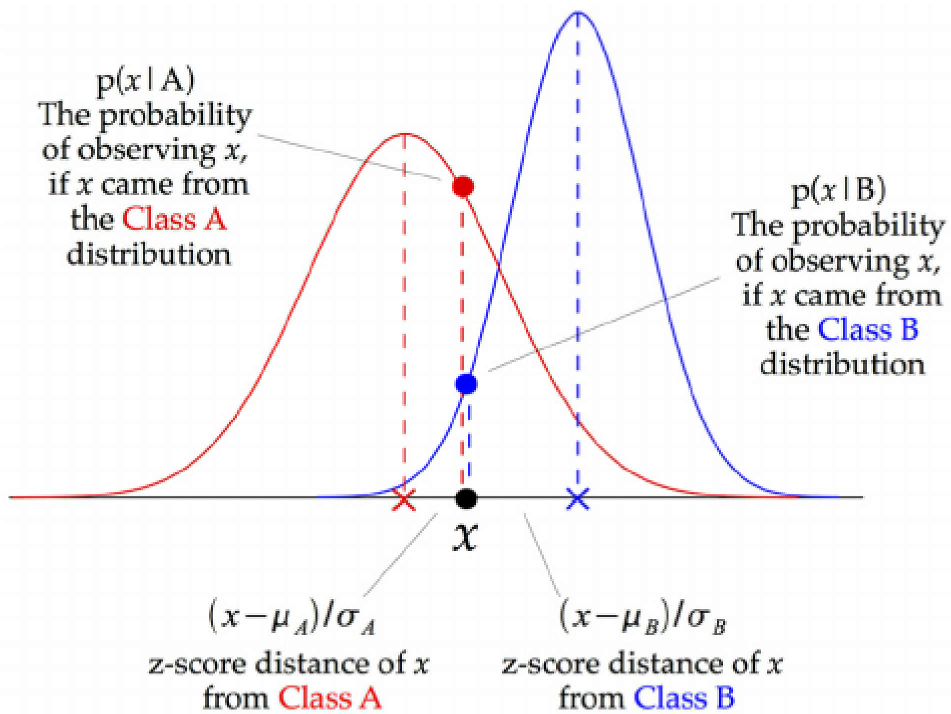


Figure 4.5: Gaussian Naive Bayes Classifier (Majumder, 2020)

4.3.2. Implementation Naive Bayes Classifier

To apply the Naive Bayes Classifier proposed in Chapter 4 to the data, an algorithm is written in Python. A summary of the functions needed to implement the Naive Bayes Classifier is as follows(Phrasant, 2020):

1. **Import Libraries**
   Libraries in Python contains the functions for various analytical functionalities, which are used in the code. In this case, there are 3 libraries that are installed in the code. 1 - Numpy library for linear algebra analytics, 2 - Pandas library for data processing, 3 - Matplotlib.pyplot for data visualisation.

2. **Import Dataset**
   The Dataset obtained after the data transformation is used for the application of the Naive Bayes Classifier. This data set is called with the command; **pd.read_csv()**.

3. **Exploratory Data Analysis**
   During Exploratory Data Analysis, the data set is checked for compliance and missing values are checked and filled. The data set is analysed in two steps. 1 - Categorical variables, these are the variables of type object, which means that a text is used as value, this is the case for the Class column in this data set. 2 - Numerical variables, these are variables of type int64, which means that the values in these columns are numeric, examples of these columns are ”Gate1, Gate2, Gate3, Static”.

4. **Declare feature vector and target variable**
   The data set is split into 2 parts, X and y. X is the feature vector, the data set that is used to recognise patterns that ensure to which y the row of data belongs. y is the target variable, this is the column that should eventually be predicted by the model, in this case the ”Class” to which the particular feature vector belongs.

5. **Split data into separate training and test set**
   Now both the feature vector X and target variable y are split into a training and a test set, in this case under a 75%-25% distribution. The training set is used to train the model and the test set is used to validate the model.

6. **Encode categorical variables**
   To prepare the categorical variables for the NBC, the variables are encoded. This assigns a value ”1” to cells where the value matches the encoder and ”0” if the value of the cell does not match the encoder.

7. **Feature Scaling**
   The numeric variables are scaled times feature scaling. The values for the NBC must be between 0 and |1|, so the values in the cells are scaled to the maximum value in that column.

8. **Model training**
   The Naive Bayes Classifier is now trained from the data. The training set of feature vector X and target variable y are put into the model to recognise patterns in the data. The model is run on the Gaussian Naive Bayes algorithm.

9. **Predict the results**
   The NBC model is used to predict the expected results. Based on the test set of the feature vector X, the values of the target variable y are predicted.

10. **Check accuracy score**
    The accuracy of the NBC model is then determined. In it, the prediction of the target variable y is compared with the test set of the target variable y. The degree to which the results match indicates how accurate the model is.

4.4. Configuration and Transformation

The RFID reader needs to be configured before the RFID antennas can read the tags and collect data. The data obtained is then transformed into a set usable for the application of the Naive Bayes Classifier.

4.4.1. Reader configuration

The reader configuration ensures that the RFID antennas will transmit and capture the radio waves. The antennas are connected to the reader via antenna cables and can thus transmit the information. To get the RFID reads required in a design, attention must be paid to setting up the RFID reader. To carefully set up a reader, the following points must be carefully determined:

- **Endpoint configuration**
  The endpoint configuration determines where the information of the RFID reads from the antennas ends up. These Tag Data Events are then forwarded to a port in the network, which can then be read in Node-RED.
- **Mode configuration**
  The mode configuration deals with the settings of the reader. This involves reading 'mode', the filter of Tag IDs, the antenna configuration and the Data configuration, which determines what information the Tag Data Events include.
- **Node-RED to data set**
  Using Node-RED, the Tag Data Events extracted from the endpoint configuration are converted via a flow configuration to the desired type of data set. In this case, the Tag Data Events are converted to a CSV file, after which the dock door discrimination method can be applied to the data set using python, the full conversion from RFID reader to python is shown in figure 4.6.

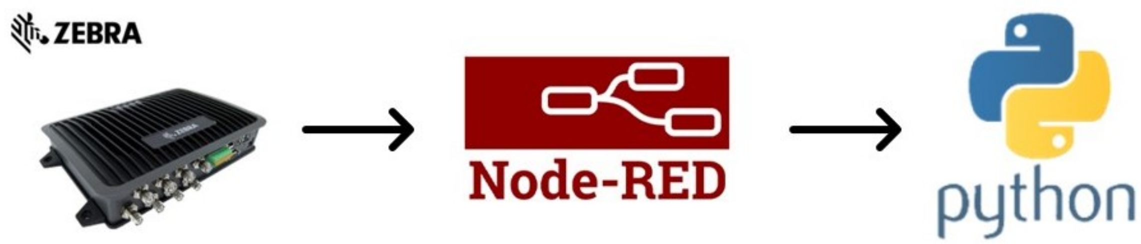


Figure 4.6: Conversion to programming tool

4.4.2. Data-set transformation

From the reader configuration comes a data set that is loaded into a CSV file, but before the dock door discrimination method can be applied, a first transformation has to be done, which is discussed in this section.

The information coming out of the CSV file has to do with the information sent via the Tag Data Events sent by the reader. These Tag Data Events are determined in the mode configuration of the RFID reader, as discussed in the previous section. The information about the reads consists of the Tag Data shown in table 4.2 below:
But before the data set is usable for the Dock Door Discrimination method, it is already further processed. The data of a test is divided into the number of runs the test has gone through. The observations are then divided into the gates where they are located, being Gate 1, Gate 2, Gate 3 or Static for static tags. Next, the variable applicable at that time is examined. For this, the options are; read counts, RSSI and phase of the observations. For the RSSI and Phase values of the observations,

Tag Data	Description
Timestamp	Time of Tag reading
Antenna	Antenna number of Tag reading
EPC	Identification code of the Tag reading
RSSI	Signal strength of Tag reading
Phase	Phase angle of Tag reading
Seen count	Count of Tag reading
Reader	Reader identification of tag reading
Endpoint	Tag Data Event endpoint

Table 4.2: Tag data

different background features can be accessed to obtain inputs with the highest accuracy for the NBC. The values for these are added together in a new DataFrame. This DataFrame is then used in evaluation to find out the accuracy of the RFID system, which will later be shown in section 5.4. It will contain the following values:

Data	Description
EPC	Identification code of RFID Tag
Class	The Gate through which the RFID tag actually passed.
Gate 1	The value of the variable read by antennas from Gate 1.
Gate 2	The value of the variable read by antennas from Gate 2.
Gate 3	The value of the variable read by antennas from Gate 3.
Static	The value of the variable read by static tags directed to staging area.

Table 4.3: Data set information

4.5. Performance Evaluation

After obtaining the wanted data set after the transformation, the set is divided into train data and test data. This division is equivalent to 75% training and 25% test data. The rows chosen for train and test data are completely random. After the Naive Bayes classifier learns from the train data, the test data is then used to determine accuracy.

To maximise the accuracy of the Naive Bayes Classifier, the different background features are used to determine the impact. These background features are shown below in table 4.4. When there is a substantial difference between the values in the data set information of a background features, it is included to obtain the highest accuracy.

Variable/Background feature	Sum	Max	Min	Mean	Median	Std dev.	Variance	Mean Absolute Deviation
Read Count	X							
RSSI		X		X	X			
Phase						X	X	X

Table 4.4: Data evaluation

As can be seen in the table, only the sum of the read counts has an impact, since the value of a read count is always 1, the other background features are of no use here. RSSI reflects the proximity of the RFID tag to the RFID antenna, so only the Max, Mean and Median may have an influence. As a result, the other background features that give the minimum value or the difference between the values are also difficult to distinguish between the different dock doors and are not used for NBC accuracy. For Phase, it is all about the relative angle to the RFID tag with respect to the antenna and therefore the difference is precisely of great influence, as shown by the Standard deviation, Variance and Mean Absolute Deviation. Usually, the tag that passed through the relevant Gate indicates the highest values for these background features.

4.5.1. Determining Accuracy

Several background features can have an affect in determining the accuracy of the Naive Bayes Classifier with respect to the data set for Dock Door Discrimination. Table 4.5 shows once again which features can possibly influence the accuracy of the NBC.

Variable	Read Count	RSSI	Phase
Feature	Sum	Max, Mean, Median	Standard Deviation, Variance, Mean Absolute Deviation

Table 4.5: Features for Naive Bayes Classifer

For each background feature, a 5-column matrix is created, as shown in table 4.6. The first column contains information about the Class, i.e. which Gate the tag actually passed through. This section distinguishes between 4 classes, which can occur. These are the classes Gate 1, Gate 2, Gate 3 and Static. Gate 1, Gate 2 and Gate 3 are the dock doors 1, 2 and 3 respectively, as was used to gather the data. Static is the designation for the static tags, which are thus read by the RFID antenna that is not at the dock door, but is directed towards the staging area.

Class	**gate1**	**gate2**	**gate3**	**static**
Gate1	18	8	13	0
Gate1	17	13	9	7
Gate1	21	6	12	1
Gate1	19	5	0	0
Gate1	25	2	2	0

Table 4.6: Matrix Sum Read Counts for scenario Dock Door 1 w/o noise

The accuracy of the Naive Bayes Classifier depends on input variable, the background feature, which is used for Dock Door Discrimination. Through data analysis of the highest accuracies per input variable, the best configuration for the Naive Bayes Classifier in the case of this data set is determined. After this is figured out for a single input feature, combinations of two or three input features are also created. This is to see if the Naive Bayes Classifier for Dock Door Discrimination becomes even more accurate in that case.

Based on the designated input feature, the Naive Bayes Classifier predicts a class at the corresponding tag, i.e. Gate 1, 2, 3 or Static. Afterwards, the prediction is compared with the real class, which is known beforehand. An example of this is shown in the confusion matrix set up per scenario. Obviously, when the tag is assigned to the right class, there is good Dock Door Discrimination. When a wrong class is assigned, there is wrong Dock Door Discrimination. The accuracy of the Dock Door Discrimination method with Naive Bayes Classifier is then determined by dividing the number of good predictions by the total number of predictions, as shown in equation 4.7 below.

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} * 100\% \tag{4.7}$$

4.5.2. Margin of Error

Paired with the determination of accuracy, the Margin of Error is calculated, similar to section 2.4. The formula for calculating the Margin of Error using accuracy proportions is shown as repition in equation 4.8. z_γ is the critical value from the Z-table (Gerstman, 2021) belonging to the selected Confidence Level, in this study it is the critical value 1.96 for a Confidence Level of 95%. p is sample proportion, in this case also called accuracy, which indicates what percentage of the measurements belong to the correct group. n is the sample size used to determine accuracy for the sample proportion.

$$Margin\ of\ Error = z_\gamma * \sqrt{\frac{p(1-p)}{n}} \tag{4.8}$$

4.6. Conclusion

This chapter has answered the sub-question: **What design is proposed for RFID Dock Door Discrimination with Naive Bayes Classifier?** To design a Dock By Discrimination method, the requirements are first defined. The most important are the functional requirements, which can be listed as follows:

- Standard dimensions
- Standard antenna configuration
- Dock door discrimination by software
- No extra hardware
- Moving to static
- Better performance than current RFID system

The non-functional requirements must contribute to achieving the functional requirements. Then the overall hardware design is started in an overall entity, being a warehouse. 15 consecutive dock doors are fitted with 4 antennas at the dock door to register the moving tags and the 5th antenna is added compared to the current RFID transition system, to register the products (RFID tags) in the staging area, otherwise known as the static tags. The detailed design is limited to three consecutive dock doors according to industry standard sizes. In this way, three different situations can be simulated; a dock door adjacent to the left, right and both sides of the respective dock door.

For the software design the Naive Bayes Classifier is implemented. This classifier is entirely based on the Bayes theorem and compares the probability of a given tag being at a dock door with the probability of the tag being somewhere else and then assigns the value with the highest probability to it. Later, the reader is configured by endpoint and mode, and Node-RED ensures that the reads are converted to a CSV file. This CSV file is then transformed into a data set that is usable to apply the Naive Bayes Classifier to. Once the train data has first been used by the model to train, the test data allows the accuracy of the NBC to be determined.

The accuracy of Dock Door Discrimination using the Naive Bayes Classifier is approximated per background feature. For this, the data set is first split into 75% train data and 25% test data. The NBC learns from the train data and is tested for accuracy using the test data. Through data analysis of the highest accuracies per input variable, the best configuration for the Naive Bayes Classifier in the case of this data set is determined. In order to determine this the good predictions are divided by the total number of prediction done by the Naive Bayes Classifier. Based on the Confidence level, sample proportion (accuracy) and sample size the Margin of Error is determined.

Chapter 5

Experimental Setup and Plan

To compare the performance of both Dock Door Discrimination methods, data needs to be gathered. Both designs are put to the test with an experimental setup and plan in this Chapter. This answers the sub-question: **What experimental setup and plan are used for gathering data on Dock Door Discrimination?** To collect the data, an experimental setup is first designed in section 5.1. This experimental setup is a scaled-down version of the hardware design from the previous chapter. Next, the test plan is created in section 5.2, in which different situations are simulated to generate a diverse data set. But before running these tests, the sample size of these tests is determined in section 5.3. Then, reader configuration and data transformation are discussed in section 5.4, this time applied to the proposed tests to collect data.

5.1. Experimental setup

The experimental setup used during data collection is scaled at 1:2 of the hardware design and thus reality. This makes the dimensions of the test setup twice as small as in the hardware design. The scaled design is as shown in both figure 5.1 and 5.2. In figure 5.1, three dock doors are shown side by side from a front view, the yellow, blue and green zone indicate a specific dock door being gate 1,2 and 3 respectively. This setup of three adjacent dock doors allow for all possible situations of cross-reads to occur at the dock doors. These situations are considered to be with an adjacent dock door to the left, right or on either side of the respective dock door.

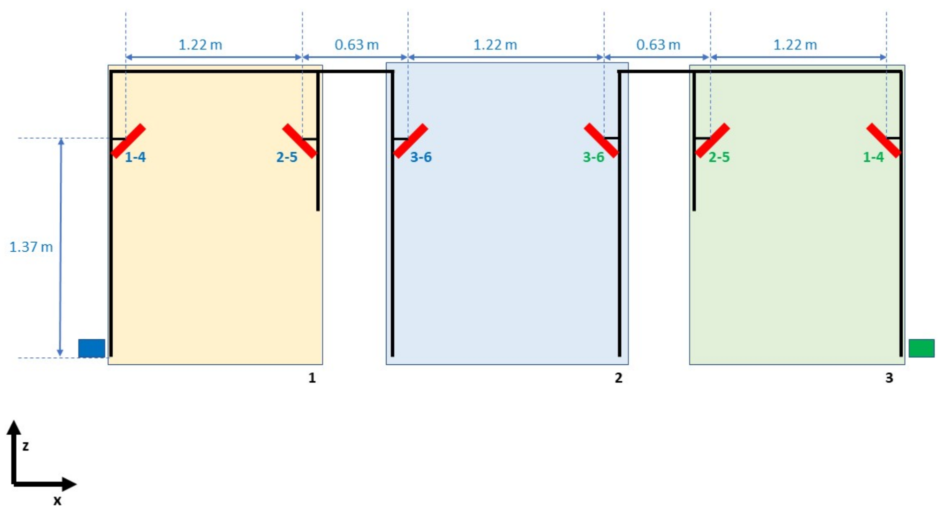


Figure 5.1: Experimental setup, front view

At each dock door, four RFID antennas are placed in the top corners of the dock door, having a 45 degree angle. These are then spaced 1.22 m apart, the width of the scaled dock door. Between adjacent dock doors there is a 0.63 m space. The height at which the antennas are attached also corresponds to the height of the dock door, which is 1.37 m in the scaled design. The blue and green square depict the RFID readers, which are used to drive up to eight associated antennas. The RFID antennas are numbered up to six and belong to the RFID reader of the same colour. This means the following for the experimental setup; Gate 1 is blue antennas 1-2-4-5, Gate 2 is antennas 3-6 of both colours and Gate 3 is green antennas 1-2-4-5.

In figure 5.2 the top view of the experimental setup is given. Here, it can be seen that the distance between the front and back two RFID antennas of a dock door is 1.25 m (for instance for Gate 1 between antennas 1-2 and 4-5). At 2.30 m from the dock doors in the negative y direction are the RFID antennas which are meant to be reading the static tags, which are in the staging area. This fifth antenna is placed there to properly distinguish between the moving and static RFID tags. The antenna with blue number 8 belongs to Gate 1, while the antennas with the green numbers 7 and 8 belong to Gate 2 and 3 respectively.

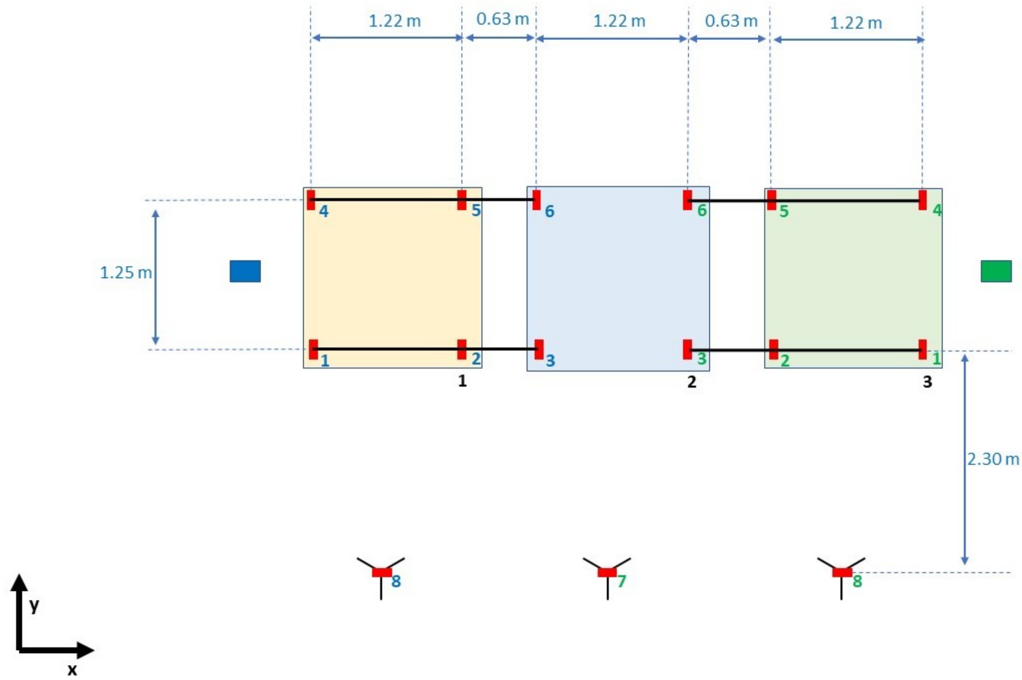


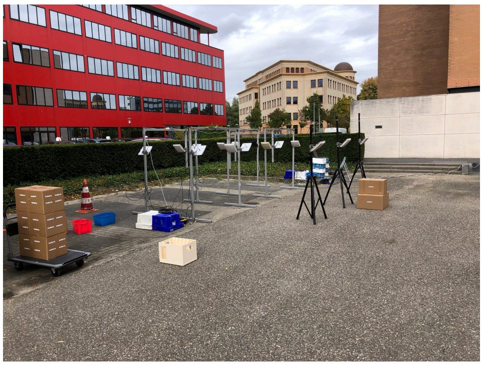
Figure 5.2: Experimental setup, top view

5.1.1. Experimental setup in reality

The experimental setup is implemented in reality. The result of the setup to simulate the three dock doors is shown in figures 5.3a,b and 5.4a,b. The skeleton of the setup are the ISB profiles that guarantee the height and width of the antennas relative to each other, just like in the experimental setup in figures 5.1 and 5.2.

In total 15 UHF RFID antennas from Mojix are used to register the RFID tags at the dock doors. These antennas are connected via antenna cables to the RFID reader, Zebra's model FX9600. The configuration of these readers is discussed in the next section. One more antenna is placed per dock door at 2.30 m using a tripod in order to distinguish the static RFID tags. The boxes with the tags are walked through the dock doors using manual carriers. This is to create as straight a walkway as possible. A slight deviation in walkway is also not so there, as forklifts in reality do not go through the Gates perfectly straight either.

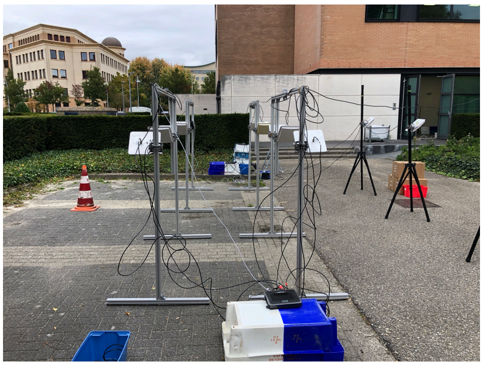
(a) Diagonal view on setup



(b) Diagonal view on setup including boxes

Figure 5.3: Test setup

(a) Diagonal view on setup from behind



(b) Side view on setup

Figure 5.4: Test setup

5.2. Experimental Plan

The experimental plan to collect the data to compare the new and old design for Dock Door Discrimination is discussed in this section. This will first explain how the testing plan works, then what different test/situations will be simulated and finally what that looks like in reality.

5.2.1. Testing plan

An initial test plan failed because there was an incorrect route for the RFID tags through the gates. This caused a lot of unwanted interference between the RFID tags and Gate 1. To correct the faulty setup of the initial test, a new experimental plan was used to gather data for the Dock Door Discrimination methods. The front view of this new experimental setup is shown in figure 5.5. Here, it is shown that there are three possible routes, through Gate 1, Gate 2 and Gate 3 respectively.

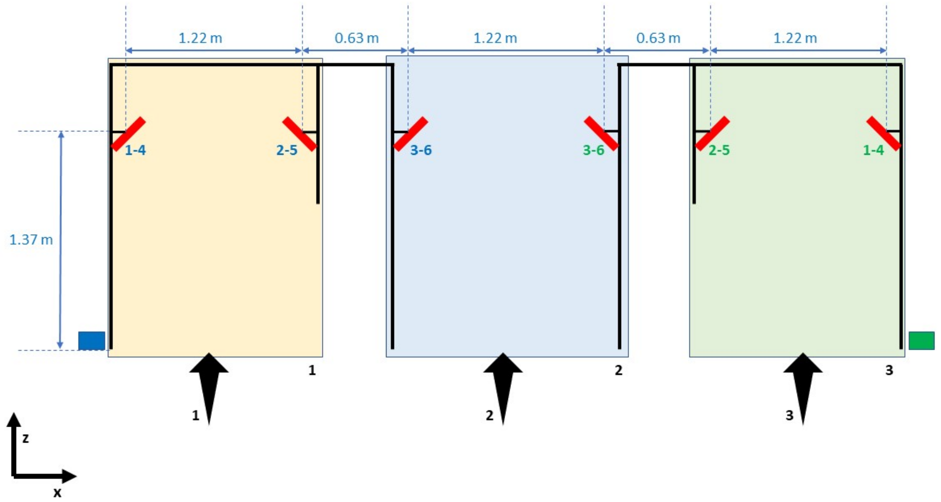


Figure 5.5: Experimental plan, Front view

To get a better idea of the actual route taken, the testing plan is also shown from a top view in figure 5.6. The start area is in front of the gates, but behind the antennas of the staging area. To start the test, a mobile User Interface was used. The stopping area is behind the gates, after the tags have moved through the gates. Stopping the test is also done via the mobile User Interface to ensure the route stayed strictly focused on the testing plan. With this User Interface the RFID readers are turned on and off whenever a button is pushed. When the reader turns on the RFID antennas switch on and start reading the RFID tags passing through the gates. These antennas also include the antennas aimed at the staging areas, shown as the orange boxes numbered 1,2 and 3. These antennas read the RFID tags in these boxes and make it easier to separate the moving and static tags.

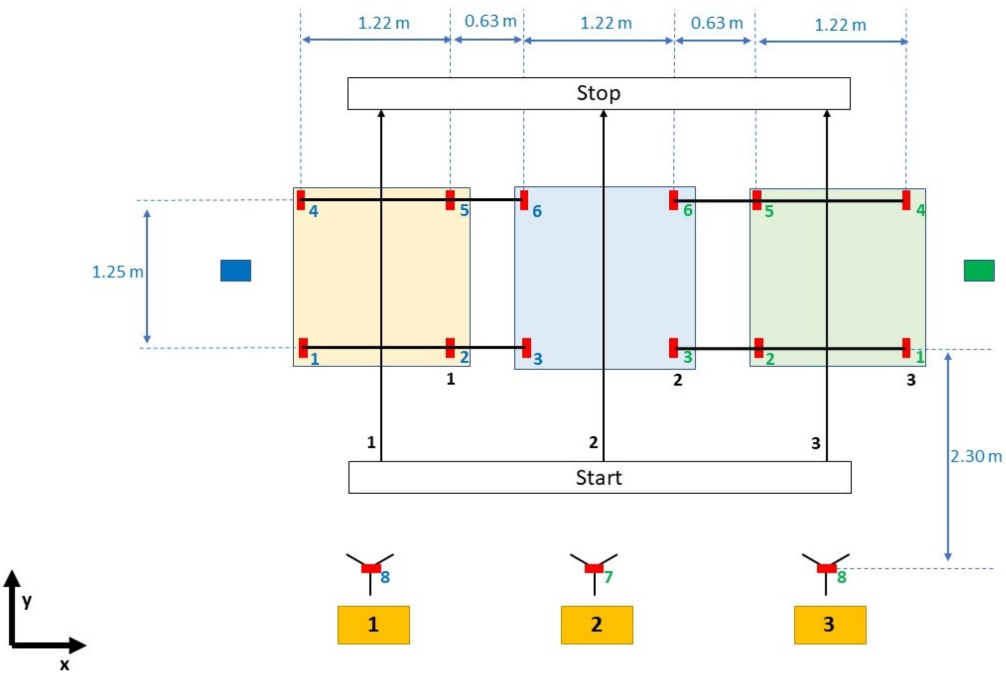


Figure 5.6: Experimental plan, Top view

5.2.2. Types of tests

To generate various data, several test scenarios are carried out, all of which simulate different situations at the dock doors. First, the RFID tags are distributed across boxes and are then moved through the dock doors in different ways to create different scenarios.

Boxes
Boxes are used in conducting the test scenarios. A number of RFID tags are stuck on these boxes to mimic a pallet of products. In total, there are 6 boxes, 3 of which are boxes for the moving RFID tags and 3 for the static tags. The moving boxes are provided with 12 RFID tags per box, 4 on the left, 4 in the middle and 4 on the right. These boxes are referred to as boxes 1, 2 and 3. The static boxes are provided with 4, 4 and 5 tags at boxes S1, S2 and S3 respectively. One tag left and right and 2 or 3 tags in the middle.

Scenarios
Different scenarios are tested to obtain a diverse data set and to properly compare the Dock Door Discrimination methods. So for this purpose, different combinations of possible scenarios that could happen in reality are simulated. An overview of these can be found in table 5.1. During the first three tests, boxes 1,2 and 3 all go through Gate 1, Gate 2 or Gate 3 at the same time. This builds up a data set of three situations without noise at other dock doors.

Scenario	Dock Door 1	Dock Door 2	Dock Door 3	Staging Area
Gate 1 w/o noise	Box 1, 2, 3	-	-	Box S1, S2, S3
Gate 2 w/o noise	-	Box 1, 2, 3	-	Box S1, S2, S3
Gate 3 w/o noise	-	-	Box 1, 2, 3	Box S1, S2, S3
Gate 1 w/ noise	Box 1 (moving)	Box 2 (static)	Box 3 (static)	Box S1, S2, S3
Gate 2 w/ noise	Box 1 (static)	Box 2 (moving)	Box 3 (static)	Box S1, S2, S3
Gate 3 w/ noise	Box 1 (static)	Box 2 (static)	Box 3 (moving)	Box S1, S2, S3

Table 5.1: Test scenarios

Next, three tests are done where each time only one box moves through the equally numbered Gate and the other boxes are static in the Gate to create noise. This ensures that it will be more difficult to perform Dock Door Discrimination as the tags are read more frequently and more strongly. But ultimately, the data set is meant to reflect these more difficult situations as well. In all the different tests, static boxes S1, S2 and S3 are at the antennas. This allows the Dock Door Discrimination method to also be tested to distinguish between moving and static tags.

5.3. Determining sample size

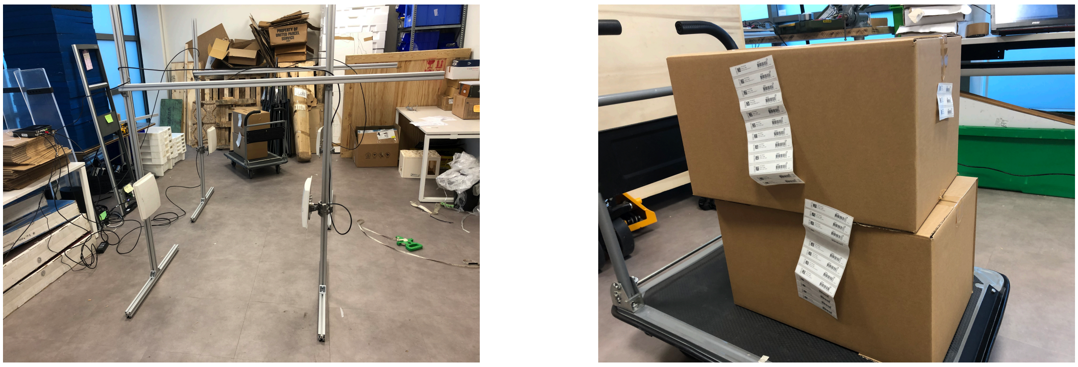
Before the different scenarios of testing can be carried out, the sample size has to be determined. This sample size n is the minimum number of times one test must be run, otherwise the test is not sufficient.

To determine the sample size of the number of tests that need to be carried out per scenario, the following function is used(Dekking et al., 2005):

$$n \geq \left(\frac{2z_{a/2}\sigma}{w}\right)^2 \tag{5.1}$$

In this equation the minimal sample size is dependent on; the confidence interval α, the standard deviation σ, the z-score $z_{\alpha/2}$, which is determined by the confidence interval, and the confidence level w. A confidence interval is a percentage of the chance that a sample falls in the confidence level, being $w = 1 - \alpha$. The standard deviation σ is the square root of the variance, being the mean deviation of the data points. The z-score z has a constant value depending on the confidence interval.

Calibration experiment
An experiment was conducted at an earlier stage to find out the sample size of the tests. For this, one dock door was used separately with the antennas facing each other, as shown in figure 5.7a. The aim of the experiment is to read as many RFID tags as possible moving through the dock door. These RFID tags were driven through the dock door on two boxes on a cart, shown in figure 5.7b. After each measurement, it is recorded how many RFID tags out of the total 49 RFID tags are read by the readers. This test was conducted 30 times to generate results.



(a) Dock door (b) Box with 49 RFID tags

Figure 5.7: Sample size test setup

Results and sample size
The results that follow from the 30 tests are shown in table 5.2 below.

Test number	RFID tags read	Test number	RFID tags read	Test number	RFID tags read
1	45	11	36	21	44
2	41	12	47	22	40
3	42	13	37	23	38
4	47	14	45	24	42
5	38	15	37	25	41
6	44	16	49	26	49
7	35	17	41	27	44
8	45	18	48	28	45
9	42	19	46	29	45
10	45	20	46	30	47

Table 5.2: Results sample size test

To determine the minimum number of tests, the formula is filled. For a confidence interval α of 95%, it means that the confidence level $w = 1 - 0.95 = 0.05$ and the z-score is 1.96. The standard deviation σ depends on the mean deviation, and then the root of that again, which comes out to 6.73% and thus σ = 0.0673.

$$n \geq (\frac{2 \cdot 1.96 \cdot 0.0673}{0.05})^2 = 26.4 \tag{5.2}$$

It follows that the sample size $n \geq 26.4$ must be adhered to. To be on the safe side, sample size 30 will be used for running the different scenarios of the testing plan.

5.4. Configuration and Transformation

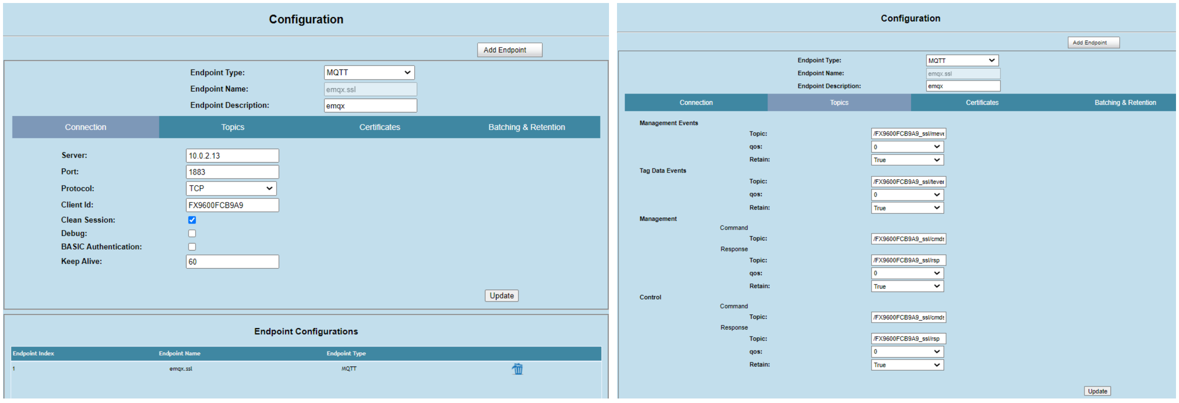
The reader configuration ensures that the antennas will read the RFID tags, so they must be set correctly. When finally the data is received from the readers, it has to be transformed before the Naive Bayes Classifier can be applied.

5.4.1. Reader configuration

To extract data from the tests, the reader must be configured. Those settings can be used to ensure that the RFID reads are converted to data points in a CSV file, making it usable for computer modeling (*FX SERIES RFID READER INTEGRATION GUIDE*, 2022).

Endpoint configuration

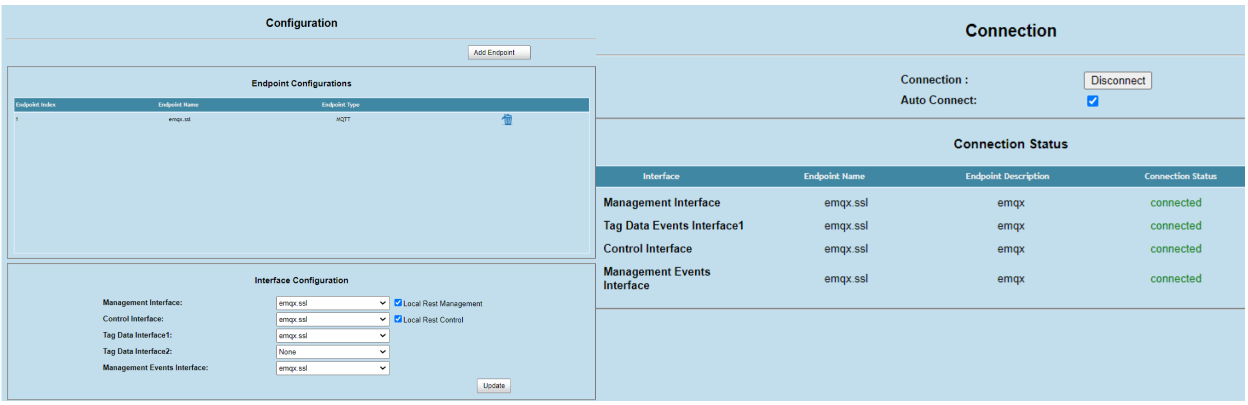
First, an endpoint must be given to the RFID tag reads. The reads are sent to an MQTT broker, which can transport data between different devices. This is important since the reads are eventually sent to a CSV file. First, the connection is configured in figure 5.8a. Here it is shown that the endpoint type is indeed MQTT, the description is chosen according to protocol of Zebra but is arbitrary. The server indicates the IP address to which the reads are sent, which is the same as the PC/laptop that is connected. Port 1883 is connected to the Node-RED, where the reads are finally loaded. Client ID indicates the corresponding reader, during this study FX9600FCB9A9 and FX9600FCB1EA are used.



(a) Connection (b) Topic

Figure 5.8: Configuration

In figure 5.8b it is shown how the reads are passed. Here a distinction is made between Management Events, Tag Data Events, Management and Control. The one used in this study are the Tag Data events, this is where the data in background features of the RFID reads is transmitted.



(a) Endpoint (b) Connection status

Figure 5.9: Configuration

In figure 5.9a is then shown that the endpoint is configured and in figure 5.9b is shown whether the endpoint is connected to the MQTT broker, which is now the case.

Mode configuration

Next, the FX9600 Zebra reader is configured, which is shown in figure 5.10. First, the mode "Inventory" with a reporting interval 0 s ensures that each tag read is included in the data set, so it remains pure raw data and nothing is lost. Indeed, if the reporting interval were larger, a summary would be made about the tag reads in that interval. The configuration also adds a filter so that only the tags that belong to this study are recorded. The tag Metadata indicates what background features are passed, for this study RSSI, Phase, tag seen count and antenna are of interest. In the antenna configuration, the transmit power of the antennas can be set, which is the power at which the tags can be read. The highest power would be 29.2 dBm, but because this study is done at 1:2 scale, 3 dBm is subtracted. In fact, a deduction of 3 dBm causes the transmit power to be halved. This means the readers' transmit power is now on 26 dBm.

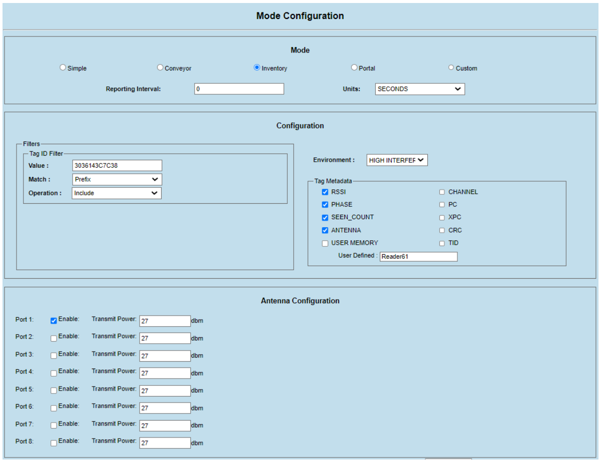


Figure 5.10: Mode configuration

Node-RED to CSV file

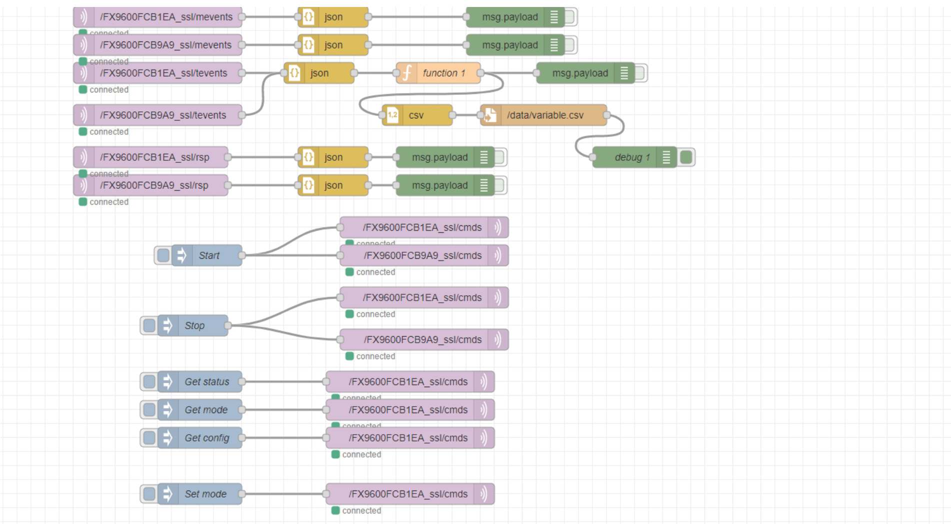
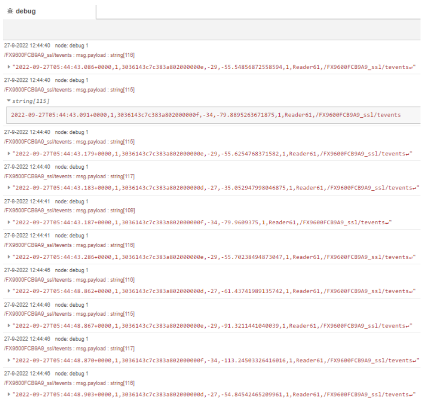


Figure 5.11: Node-RED configuration

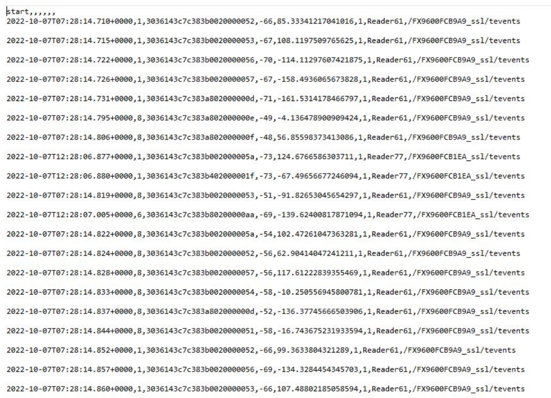
As mentioned earlier, the reader is sent to network address 10.0.2.13/1883. Using Node-RED, see figure 5.11, the tag reads can be converted into usable data for this study. In fact, the Tag Data events from both RFID readers are converted into a CSV file here. The tag reads come out in Node-RED as shown in figure 5.12a and the CSV file is shown in figure 5.12b. The data that eventually all comes out of the Tag Data Events is as follows:

timestamp	antenna	epc	rssi	phase	reads	readername	reader
2022-11-10T15:33:48.403+0000	1.0	0012	-69.0	78.27997589111328	1.0	Reader77	/FX9600FCB1EA_ssl/tevents
2022-11-10T15:33:48.405+0000	1.0	0019	-63.0	-4.394665241241455	1.0	Reader46	/FX9600F445D8_ssl/tevents
2022-11-10T15:33:48.487+0000	4.0	005b	-74.0	-146.78732299804688	1.0	Reader77	/FX9600FCB1EA_ssl/tevents
2022-11-10T15:33:48.492+0000	4.0	001c	-73.0	-47.6656379699707	1.0	Reader77	/FX9600FCB1EA_ssl/tevents
2022-11-10T15:33:48.584+0000	7.0	0057	-48.0	76.1265869140625	1.0	Reader77	/FX9600FCB1EA_ssl/tevents
2022-11-10T15:33:48.570+0000	8.0	0010	-70.0	99.38536071777344	1.0	Reader46	/FX9600F445D8_ssl/tevents
2022-11-10T15:33:48.710+0000	3.0	001a	-72.0	-63.17331314086914	1.0	Reader46	/FX9600F445D8_ssl/tevents

Table 5.3: Tag Data Events



(a) Node-RED



(b) CSV file

Figure 5.12: Data results

5.4.2. Data transformation

The Tag Data Events in table 5.3 are transformed to the specific value that can affect the Dock Door Discrimination method. An example of this is shown in table 5.4. Here, the first 5 observations are listed, where the sum of the read counts is the background feature. The Class indicates that the tag passed through Gate 1. Based on the values in the other columns, the Naive Bayes Classifier tries to find patterns to distinguish between the different Gates. Another example is shown in table 5.5. Here, the same has been done, but for the RSSI max. Combinations of background features can also be made. In that case, all columns except Class are concatenated into a new input data set.

Class	gate1	gate2	gate3	gate0
Gate1	13.0	11.0	17.0	2.0
Gate1	18.0	8.0	13.0	X
Gate1	17.0	13.0	9.0	7.0
Gate1	21.0	6.0	12.0	1.0
Gate1	19.0	5.0	X	X

Table 5.4: Data Read Counts

ClassRSSImax	Gate1RSSImax	Gate2RSSImax	Gate3RSSImax	Gate0RSSImax
Gate1	-50.0	-64.0	-63.0	-65.0
Gate1	-49.0	-58.0	-66.0	X
Gate1	-49.0	-56.0	-60.0	-61.0
Gate1	-51.0	-61.0	-63.0	-78.0
Gate1	-55.0	-61.0	X	X

Table 5.5: Data RSSI max

5.5. Conclusion

This chapter has answered the sub-question: **What experimental setup and plan are used for gathering data on Dock Door Discrimination?** The experimental setup is derived from the detailed design of the previous chapter, but scaled by a factor of 1:2. Four antennas at the dock door keep track of the moving tags, while the fifth antenna, supported by a tripod record the static tags in the so-called staging area. The experimental plan is to move the moving tags through the dock doors from start to stop. The RFID readers can be switched on and off using a mobile User Interface. Six different test scenarios will then be carried out to make the data set diverse. First, three tests without noise at other dock doors, then with noise.

Before the tests are done, the sample size is first determined using a calibration test. This shows that the sample size per test is 26.4, but to be on the safe side, 30 samples are done per scenario. To finally get a usable data set, the RFID readers have to be configured so that the RFID readings of the antennas are displayed in the data set. This data is converted via Node-RED to a CSV file, which in turn needs to be further transformed to be used for the Naive Bayes Classifier. This transformation means that a data set can be created per background feature with values per Gate, which the NBC then needs to start recognising patterns for Dock Door Discrimination.

Chapter 6

Results and Discussion

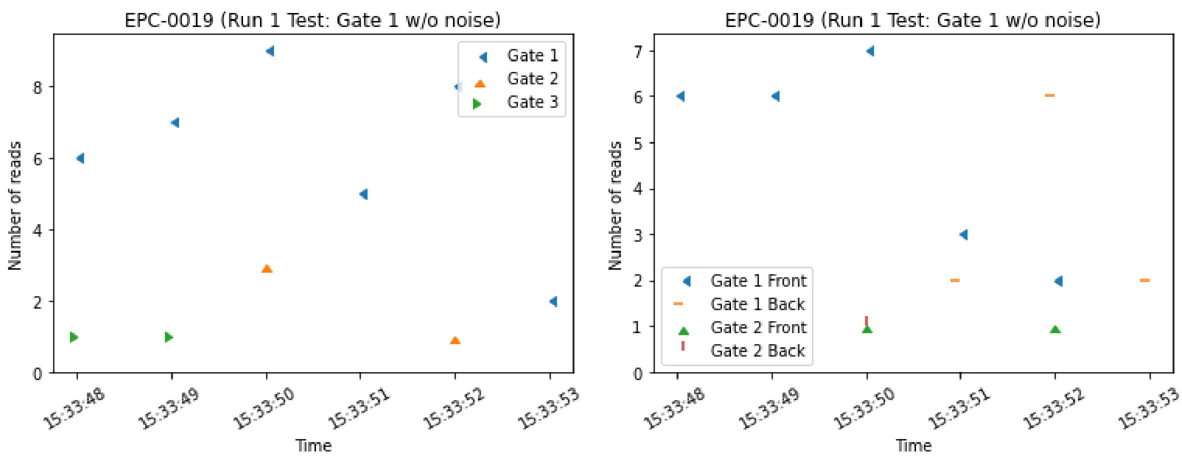
The purpose of this chapter is to answer the following sub-question: **What is the performance of the Naive Bayes Classifier compared to current RFID system?** To answer this question, a better idea behind the data collected during the tests is first generated by scatter plotting the data without Dock Door Discrimination method and with the current RFID transition system for Dock Door Discrimination in section 6.1, to give a better idea of how it works. Then, the accuracy of this current RFID system for Dock Door Discrimination is determined in section 6.2, per scenario and for all scenarios combined. Later in section 6.3, the accuracy of Dock Door Discrimination with Naive Bayes Classifier is also determined, again per scenario and for all scenarios combined. Then, a comparison between the accuracy of the current and new system for Dock Door Discrimination is made in section 6.4 to see which system perform better. Last, a sensitivity analysis is conducted on the Margin of Error the find out the robustness of the results in section 6.5.

6.1. Data plots without Dock Door Discrimination method

This section shows the scatter plots for the data without using a Dock Door Discrimination method and compares them with the scatter plats for the data using the current Dock Door Discrimination method. A distinction is made between Gate 1, 2 and 3 in the number of reads of an RFID tags over the period of one second. In the case of the current RFID system an extra distinction is made, splitting the RFID tag reads at the front- and backside of the dock door. Figures are shown and discussed for each scenario individually.

6.1.1. Gate 1 without noise

The first test simulates the scenario where boxes 1,2 and 3 (total 36 RFID tags with 12 per box) pass through Gate 1 simultaneously. In addition, there is no movement of RFID tags through Gate 2 and 3. Through a scatter plot of the RFID read points per EPC code (RFID tag), a better understanding of the trajectory of such an RFID tag is obtained, as shown in figure 6.1a.



(a) Tag EPC 0019 during run 1 through Gate 1 without noise

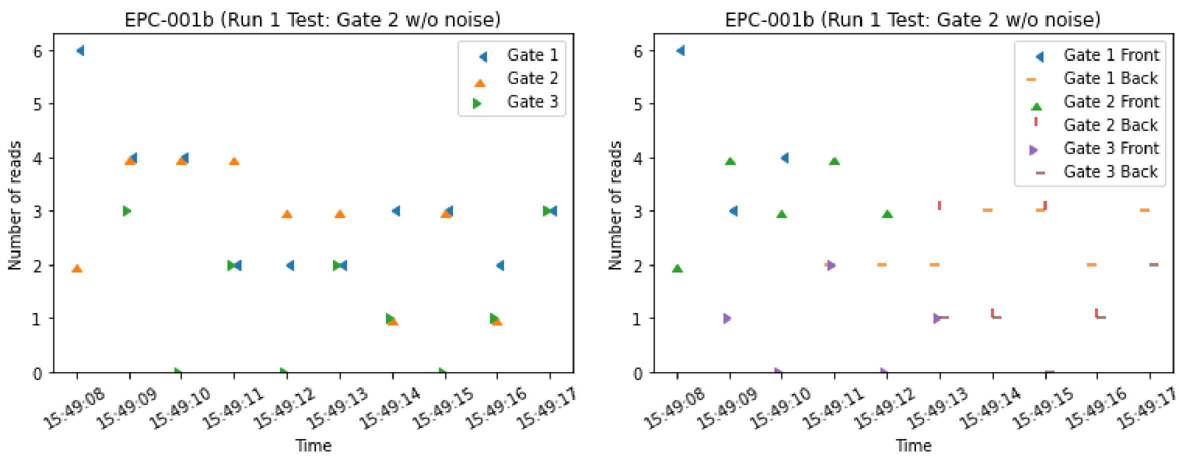
(b) Tag ending 0019 during run 1 through Gate 1 without noise

The figure shows how many times the RFID tag with EPC code 0019 is read per second by the antennas of Gate 1, 2 and 3. At all times, the sum of the number of reads is greatest at Gate 1. This can be explained by the fact that the RFID tag is closest to the antennas of this gate when the tag passes through Gate 1. The increasing number of reads per second is explained by the RFID tag being closer to the antennas during the first half of the run. Similarly, the number of reads per second also decreases again during the second half of the run.

Figure 6.1b shows what the current RFID system captures for Dock Door Discrimination. The two readings for Gate 3 are dropped as they are done by the rear two antennas of Gate 3. These are filtered out using the timewindow. Furthermore, it is clear that the tag went through Gate 1, since both Gate 1 Front and Back have the most reads. It is also noticeable that these readings also increase to a peak and then decrease again. This is because the tag first passes close to the Front and then the Back RFID antennas.

6.1.2. Gate 2 without noise

The second test simulates the scenario where boxes 1,2 and 3 (total 36 RFID tags with 12 per box) pass through Gate 2 at the same time. In addition, there is no movement of RFID tags through Gate 1 and 3. By means of a scatter plot of the RFID read points per EPC code (RFID tag), insight is given into the trajectory of such an RFID tag, as shown in figure 6.2a.



(a) Tag ending 001b during run 1 through Gate 2 without noise

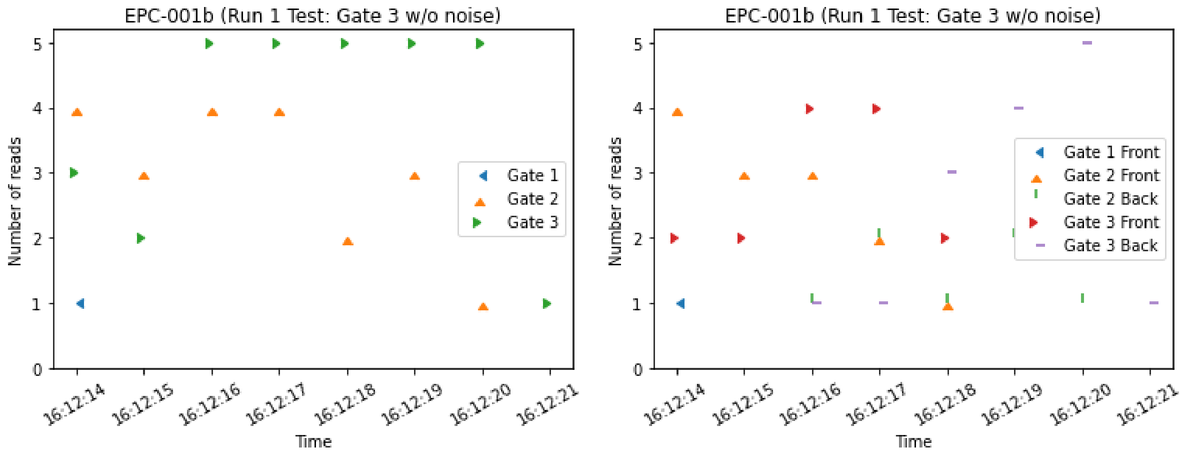
(b) Tag ending 001d during run 1 through Gate 2 without noise

The figure shows the sum of the number of read counts per second of an RFID tag with EPC code 001b read by the RFID antennas of Gate 1, Gate 2 and Gate 3. It is clearly visible that in this case it is more difficult to distinguish that this RFID tag actually passed through Gate 2. With such trajectories, a Dock Door Discrimination method therefore has more difficulty in determining which Gate the tag has passed through. This is because the tag is also often read by the other Gates, this is because both gates are now on both sides of Gate 2.

Figure 6.2b shows what the current RFID system captures for Dock Door Discrimination. In this case, it is difficult for the current RFID system to perform proper Dock Door Discrimination. This is because the tag is read frequently at both Gate 1 and Gate 2. This is an example of cross-reading, two Gates assume the tag has passed.

6.1.3. Gate 3 without noise

The third test simulates the scenario where boxes 1,2 and 3 (total 36 RFID tags with 12 per box) pass through Gate 3 simultaneously. In addition, there is no movement of RFID tags through Gate 1 and 2. By means of a scatter plot of the RFID read points per EPC code (RFID tag), insight is given into the trajectory of such an RFID tag, as shown in figure 6.3a.



(a) Tag ending 001b during run 1 through Gate 3 without noise

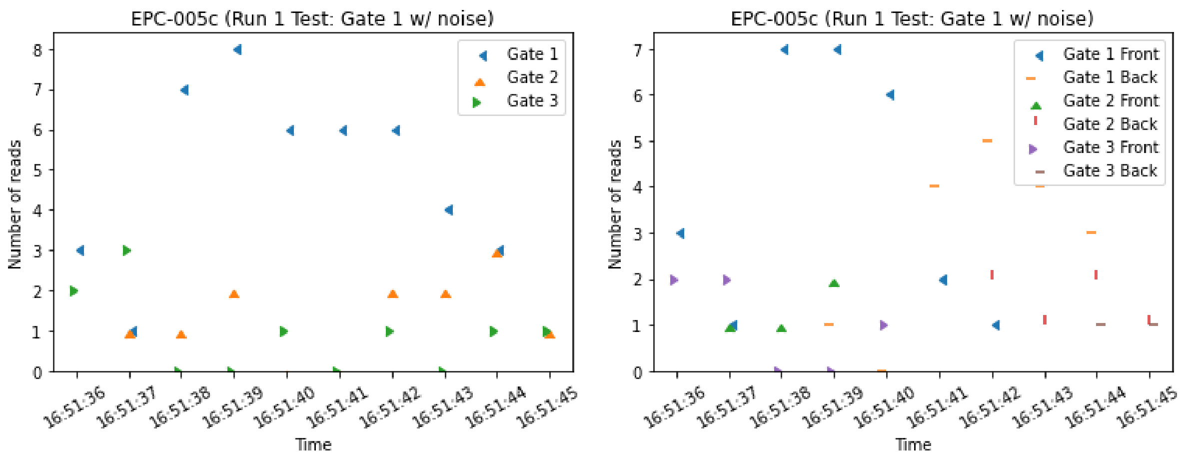
(b) Tag ending 001d during run 1 through Gate 3 without noise

The figure shows the sum of the number of read counts per second of an RFID tag with EPC code 001b read by the RFID antennas of Gate 1, Gate 2 and Gate 3. Visually, it is clearly visible in this case that the RFID tag has passed through Gate 3. In fact, the tag is most often read in the middle section, when the tag passes close to the antennas. The sum of reads at Gate 3 is lower at the beginning and end of the path as that tag is then further away from the antennas. The tag is also often, but less than at Gate 3, read at Gate 2. This is because Gate 2 is located right next to Gate 3. To a lesser extent, the tag is read at Gate 1, since it is furthest away from Gate 3.

Figure 6.3b shows what the current RFID system captures for Dock Door Discrimination. Initially, it is difficult to discriminate the tag from Gate because the tag is often read by the front of Gate 2 and Gate 3. Then, when the number of reads at the back of the Gate are also considered, it becomes clear that the tag has passed through Gate 3. Because the tag is read more frequently at the back, the current RFID system in this case has no problem with Dock Door Discrimination.

6.1.4. Gate 1 with noise

The first test with noise simulates the scenario where box 1 (total 12 RFID tags) passes through Gate 1. Further, boxes 2 and 3 are in Gate 2 and 3 respectively, to cause noise when reading the RFID tags. In this way, the impact of noise at other Gates on Dock Door Discrimination is found out. By means of a scatter plot of the RFID read points per EPC code (RFID tag), insight is given into the trajectory of the RFID tag, as shown in figure 6.4a.



(a) Tag ending 005c during run 1 through Gate 1 with noise

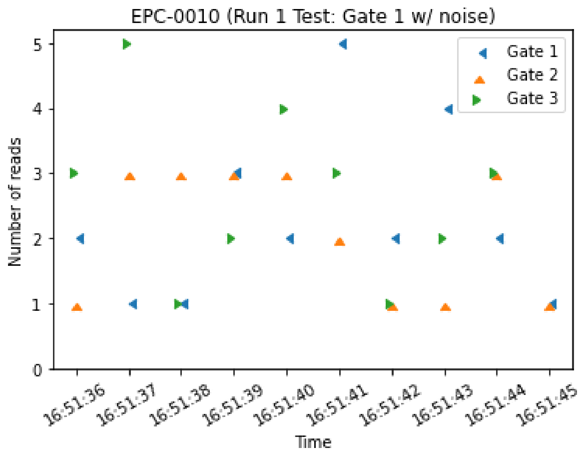
(b) Tag ending 005c during run 1 through Gate 1 with noise

The figure shows the sum of the number of read counts per second of an RFID tag with EPC code 005c (placed on box 1) read by the RFID antennas of Gate 1, Gate 2 and Gate 3. The figure clearly shows that the tag passed through Gate 1, as that is where the tag is read most often. What is also noticeable is that the sum of the number of reads first increases and then decreases again, clearly showing that Gate 1 has seen the tag first approach and then leave and thus has passed. The tag was also read at Gates 2 and 3, but not enough to question the actual direction.

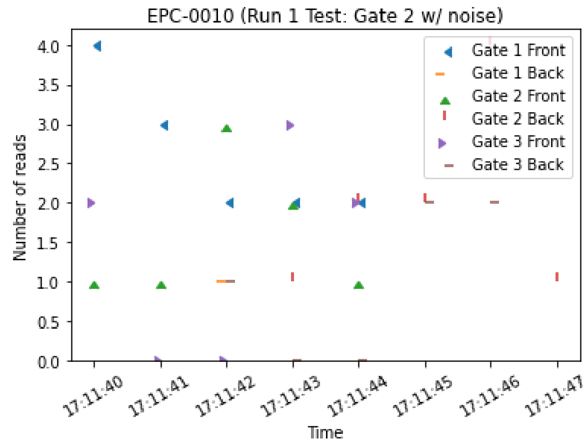
Figure 6.4b shows what the current RFID system captures for Dock Door Discrimination. The figure clearly shows that the tag has moved through Gate 1, as there is a rising and falling line recorded at the front and back of the Gate. The large difference in number of reads with the other gates makes for good Dock Door Discrimination.

6.1.5. Gate 2 with noise

The second test with noise simulates the scenario where box 2 (total of 12 RFID tags) passes through Gate 2. Furthermore, boxes 1 and 3 (also containing 12 RFID tags) are in Gate 1 and 3 respectively, to cause noise when reading the RFID tags. In this way, the impact of noise at other Gates on Dock Door Discrimination is found out. By means of a scatter plot of the RFID read points per EPC code (RFID tag), insight is given into the trajectory of the RFID tag, as shown in figure 6.5a.



(a) Tag ending 0010 during run 1 through Gate 2 with noise



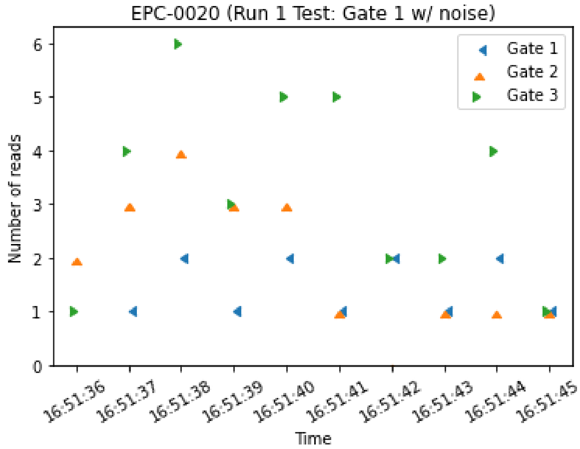
(b) Tag ending 0010 during run 1 through Gate 2 with noise

The figure shows the sum of the number of read counts per second of an RFID tag with EPC code 0010 (placed on box 2) read by the RFID antennas of Gate 1, Gate 2 and Gate 3. It is clearly visible that in this case it is more difficult to distinguish that this RFID tag actually passed through Gate 2. A Dock Door Discrimination method can therefore have more difficulty with such trajectories to determine which Gate the tag has passed through. In this case, the tag is more often read by Gate 3 in the first half and more often read by Gate 1 in the second half of the trajectory. This is because both gates are now on both sides of Gate (2).

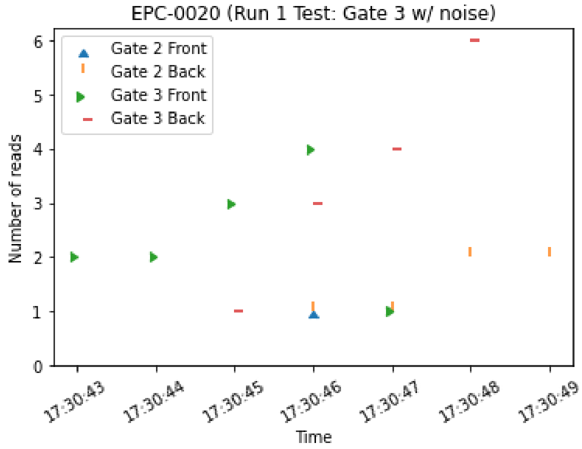
Figure 6.5b shows what the current RFID system captures for Dock Door Discrimination. In this case, it is again more difficult to properly discern the tag direction. Initially, the tag appears to pass through Gate 1, since the tag is often read at the front. When the number of reads at the back are also taken into account, it becomes already more clear that the tag goes through Gate 2, since this is where the tag is read most often. It does remain difficult for this current RFID system to perform proper Dock Door Discrimination in this case.

6.1.6. Gate 3 with noise

The third test with noise simulates the scenario where box 3 (total of 12 RFID tags) passes through Gate 3. Furthermore, boxes 1 and 2 (also containing 12 RFID tags) are in Gate 1 and 2 respectively, to cause noise when reading the RFID tags. In this way, the impact of noise at other Gates on Dock Door Discrimination is found out. By means of a scatter plot of the RFID read points per EPC code (RFID tag), insight is given into the trajectory of the RFID tag, as shown in figure 6.6a.



(a) Tag ending 0020 during run 1 through Gate 3 with noise



(b) Tag ending 0020 during run 1 through Gate 3 with noise

The figure shows the sum of the number of read counts per second of an RFID tag with EPC code 0020 (placed on box 3) read by the RFID antennas of Gate 1, Gate 2 and Gate 3. The figure clearly shows that the tag passed through Gate 3, as that is where the tag is read most often. What is also noticeable is that the sum of the number of reads first increases and then decreases again, clearly showing that Gate 1 has seen the tag first approach and then leave and thus has passed. The tag was also read at Gates 1 and 2, but not enough to question the actual direction.

Figure 6.6b shows what the current RFID system captures for Dock Door Discrimination. The figure clearly shows that the tag has moved through Gate 3, as there is a rising and falling line recorded at the front and back of the Gate. The large difference in number of reads with the other gates makes for good Dock Door Discrimination.

6.2. Current RFID system for Dock Door Discrimination: Accuracy

In this section, the accuracy of the current RFID system for Dock Door Discrimination is discussed for the generated data set. First, the accuracy for each individual scenario is considered and then the accuracy when the system is subjected to all scenarios combined.

6.2.1. Accuracy individual scenarios

The results for accuracy of the current RFID transition system for Dock Door Discrimination per individual scenario are divided into six column values in this section. The first three accuracies are the scenarios without noise and the second three accuracies of scenarios with noise. The accuracies are determined per threshold value for the number of reads ”n”, as shown in table 6.1.

n = ...	Gate 1 w/o	Gate 2 w/o	Gate 3 w/o	Gate 1 w/	Gate 2 w/	Gate 3 w/
1	35.4±1.7	27.7±0.6	41.6±2.1	38.2±1.9	31.6±1.3	32.6±1.4
2	48.4±2.3	32.2±1.4	52.1±2.4	46.7±2.3	38.0±1.9	41.4±2.1
3	61.5±2.6	39.2±1.9	61.7±2.6	57.5±2.5	45.6±2.2	47.8±2.2
4	71.2±2.5	46.1±2.3	70.8±2.5	67.5±2.5	53.7±2.5	54.0±2.5
5	79.9±2.3	50.7±2.4	78.2±2.3	75.4±2.4	61.8±2.6	60.8±2.6
6	86.7±2.0	55.1±2.5	84.3±2.1	80.9±2.2	71.5±2.5	66.7±2.6
7	91.0±1.7	60.0±2.6	89.3±1.8	**83.2±2.2**	76.9±2.4	71.0±2.5
8	94.4±1.4	63.5±2.6	93.3±1.5	82.4±2.2	80.1±2.3	73.5±2.5
9	**95.3±1.3**	**68.2±2.5**	94.6±1.3	79.7±2.3	**80.5±2.3**	**74.1±2.4**
10	94.4±1.4	67.5±2.5	**95.0±1.3**	76.1±2.4	77.0±2.4	72.9±2.5
11	92.0±1.6	61.7±2.6	92.2±1.6	71.1±2.5	71.6±2.5	71.8±2.5
12	87.8±1.9	55.1±2.5	86.4±2.0	66.1±2.6	66.3±2.6	68.8±2.5
13	83.3±2.1	47.8±2.3	78.1±2.3	62.5±2.6	61.9±2.6	65.3±2.6
14	78.6±2.3	39.9±2.0	68.9±2.5	58.4±2.5	57.2±2.5	62.6±2.6
15	72.9±2.5	34.4±1.6	58.4±2.5	55.1±2.5	53.7±2.5	59.7±2.5

Table 6.1: Accuracy current RFID transition system per scenario

Gate 1 without noise
The accuracy of the current RFID transition system for Dock Door Discrimination for testing at Gate 1 without noise at Gate 2 and Gate 3 has the highest value when the threshold value $n = 9$ is applied. This means that a successful Dock Door Discrimination happens when the tag is read at least 9 times at the front and back of Gate 1. At the same time, the tag must not be read 9 times or more at Gate 2 or Gate 3. The accuracy of Dock Door Discrimination for this scenario is **95.3±1.3%**. It is noticeable that the accuracy decreases as the threshold increases or decreases. At a lower threshold, the other Gates seem to confirm that the tag has passed through Gate 2 or Gate 3, causing cross-reading. At a higher threshold, fewer transitions are confirmed by the Gates, particularly Gate 1, as the tag has not been read often enough at the front and back of the dock door, causing miss-reading.

Gate 2 without noise
The accuracy of the current RFID transition system for Dock Door Discrimination for testing at Gate 2 without noise at Gate 1 and Gate 3 has the highest value when the threshold value $n = 9$ is applied. This means that a successful Dock Door Discrimination happens when the tag is read at least 9 times at the front and back of Gate 2. At the same time, the tag must not be read 9 times or more at Gate 1 or Gate 3. The accuracy of Dock Door Discrimination for this scenario is **68.2±2.5%**. It is noticeable that the accuracy for this test is much lower than the same tests at Gate 1 and Gate 3. The reason for this is that Gate 2 is between Gate 1 and Gate 3. This means that RFID tags passing through Gate 2 are also often read by both Gate 1 and Gate 3. In the other tests (Gate 1 and 3), there is only one connecting dock door each time, Gate 2. The significant reduction in accuracy shows that the current RFID system is struggling with two-way cross-reading.

Gate 3 without noise
The accuracy of the current RFID transition system for Dock Door Discrimination for testing at Gate 3 without noise at Gate 1 and Gate 2 has the highest value when the threshold value $n = 10$ is applied. This means that a successful Dock Door Discrimination happens when the tag is read at least 10 times at the front and back of Gate 1. At the same time, the tag must not be read 10 times or more at Gate 1 or Gate 2. The accuracy of Dock Door Discrimination for this scenario is **95.0±1.3%**. It is noticeable that the accuracy decreases as the threshold increases or decreases. At a lower threshold, the other Gates seem to confirm that the tag has passed through Gate 1 or Gate 2, causing cross-reading. At a higher threshold, fewer transitions are confirmed by the Gates, particularly Gate 3, as the tag has not been read often enough at the front and back of the dock door, causing miss-reading.

Gate 1 with noise
The accuracy of the current RFID transition system for Dock Door Discrimination for testing at Gate 1 with noise at Gate 2 and Gate 3 has the highest value when the threshold value $n = 7$ is applied. This means that a successful Dock Door Discrimination happens when the tag is read at least 7 times at the front and back of Gate 1. At the same time, the tag must not be read 7 times or more at Gate 2 or Gate 3. The accuracy of Dock Door Discrimination for this scenario is **83.2±2.2%**. It is noticeable that the accuracy for this test is lower than the test by Gate 1 without noise. Thus, the fact that there is noise at the other Gates has caused the current RFID system to perform less well for this scenario. This may be because the number of tag reads done per Gate are averaged out. The reason is that Gate 1's antennas pick up more signals from the noise at other Gates. This makes the different tag readings less distinguishable.

Gate 2 with noise
The accuracy of the current RFID transition system for Dock Door Discrimination for testing at Gate 2 with noise at Gate 1 and Gate 3 has the highest value when the threshold value $n = 9$ is applied. This means that a successful Dock Door Discrimination happens when the tag is read at least 9 times at the front and back of Gate 2. At the same time, the tag must not be read 9 times or more at Gate 1 or Gate 3. The accuracy of Dock Door Discrimination for this scenario is **80.5±2.3%**. It is noticeable that the accuracy for this test is higher than the test by Gate 2 without noise. Thus, the fact that there is noise at the other Gates has made the current RFID system work better for this scenario. This may be because the number of tag reads done per Gate are averaged out. The reason is that the antennas of Gate 1 and Gate 3 pick up more signals from the noise and therefore are less likely to read the tags going through Gate 2. This makes the different readings of the tags more distinguishable.

Gate 3 with noise
The accuracy of the current RFID transition system for Dock Door Discrimination for testing at Gate 3 with noise at Gate 1 and Gate 2 has the highest value when the threshold value $n = 9$ is applied. This means that a successful Dock Door Discrimination happens when the tag is read at least 9 times at the front and back of Gate 1. At the same time, the tag must not be read 9 times or more at Gate 1 or Gate 2. The accuracy of Dock Door Discrimination for this scenario is **74.1±2.4%**. It is noticeable that the accuracy for this test is lower than the test by Gate 3 without noise. Thus, the fact that there is noise at the other Gates has caused the current RFID system to perform less well for this scenario. This may be because the number of tag reads done per Gate are averaged out. The reason is that Gate 3's antennas pick up more signals from the noise at other Gates. This makes the different tag readings less distinguishable.

6.2.2. Accuracy all scenarios combined

The results for accuracy of the current RFID transition system for Dock Door Discrimination for all scenarios combined are given in this section. This means that the full data set of the six scenarios combined is used for determining the accuracy. The accuracy is determined per threshold value for the number of reads ”n”, as shown in table 6.2.

n = ...	Accuracy
1	34.5±1.0
2	43.1±1.0
3	52.2±1.0
4	60.5±1.0
5	67.8±1.0
6	74.2±0.9
7	78.6±0.9
8	81.2±0.8
9	**82.1±0.8**
10	80.5±0.8
11	76.7±0.9
12	71.8±0.9
13	66.5±1.0
14	60.9±1.0
15	55.7±1.0

Table 6.2: Accuracy current RFID transition system all scenarios combined

The accuracy of the current RFID system for Dock Door Discrimination for all scenarios combined has the highest value when the threshold value n=9 is applied. This means that a successful Dock Door Discrimination happens when the tag is read at least 9 times at the front and back of the Gate, where the tag passes. At the same time, the tag must not be read 9 times or more at the other Gates where the tag does not pass. The accuracy of the current RFID transition system for Dock Door Discrimination for all scenarios combined then becomes **82.1±0.8%**.

6.3. Dock Door Discrimination with Naive Bayes Classifier

In this section, it is discussed what the accuracy of Dock Door Discrimination with the Naive Bayes Classifier is for the same generated data set. As indicated in section 4.5, the Dock Door Discrimination with the Naive Bayes Classifier uses 75% of this data set for training and 25% for testing. This section first looks at the accuracy per individual scenario and then at the accuracy for all scenarios combined.

6.3.1. Accuracy single input features

The accuracy of the Naive Bayes Classifier for Dock Door Discrimination with a single input feature is given in this section, first by individual scenario, then for all scenarios combined.

Individual scenarios
The accuracy of the NBC at input for a single feature is determined per individual scenario, as shown in the table 6.3 below. The results are given per feature, as shown in the left column, and per tested scenario, as shown in the top row.

Feature	Gate 1 w/o	Gate 2 w/o	Gate 3 w/o	Gate 1 w/	Gate 2 w/	Gate 3 w/
Read counts	100	100	100	89.0±1.6	97.1±0.9	91.1±1.5
RSSI max	99.7±0.3	99.2±0.5	100	69.3±2.4	91.9±1.4	81.4±2.0
RSSI mean	99.7±0.3	99.2±0.5	100	70.1±2.3	85.3±1.8	82.9±1.9
RSSI median	99.7±0.3	99.2±0.5	100	70.6±2.3	82.2±2.0	83.5±1.9
Phase std	100	99.7±0.3	99.2±0.5	65.1±2.4	72.2±2.3	71.9±2.3
Phase var	99.5±0.4	99.7±0.3	98.7±0.6	60.6±2.5	70.6±2.3	68.2±2.4
Phase mad	100	99.7±0.3	99.2±0.5	63.8±2.5	69.6±2.4	71.7±2.3

Table 6.3: Accuracy single input per individual scenario

The table indicates that the Naive Bayes Classifier for the individual scenarios achieves the highest accuracy for the input feature Read Counts. The three scenarios without noise the NBC runs through flawlessly in that particular case (100%). The three scenarios with noise have lower accuracy, but still considerably higher than for the other input features. These other input features also have little trouble with Dock Door Discrimination during the scenarios without noise, but are never error-free in all cases there. For the scenarios with noise, the accuracy goes down rapidly, noting that only the input feature RSSI max by itself can amass a score higher than 90% (Gate 2 with noise).

All scenarios combined
The accuracy of the NBC for Dock Door Discrimination when inputting a single feature for all scenarios combined is given in Table 6.4. The left column gives the input feature and the right column gives the corresponding accuracy.

Feature	Accuracy (%)
Read counts	92.6±0.5
RSSI mean	88.0±0.7
RSSI median	86.3±0.7
RSSI max	87.1±0.7
Phase std	71.1±0.9
Phase var	68.4±1.0
Phase mad	71.7±0.9

Table 6.4: Accuracy single input feature for all scenarios combined

The table shows that for all scenarios combined, the Naive Bayes Classifier achieves the highest overall accuracy for the input feature Read Counts (**92.6±0.5%**). The accuracy for the RSSI inputs is higher than for the Phase input features. This is because the RSSI values are absolute versus Phase's relative features, which ensures that the Dock Door Discrimination is better. To see whether a combination of two input features can have an impact on accuracy, the next section combines Read Counts and other input features.

6.3.2. Accuracy double input features

The accuracy of the Naive Bayes Classifier for Dock Door Discrimination with two input features (double) is given in this section, first by individual scenario, then for all scenarios combined. The first feature is Read Counts and then any other input feature is used as a combination to obtain higher accuracy for Dock Door Discrimination with NBC.

Individual scenarios
The accuracy of the NBC when inputting for two features is given per individual scenario, as shown in the table 6.5 below. The results are given per feature (+ Read counts), as shown in the left column, and per tested scenario, as shown in the top row. Each time, one input feature is added to the Read Counts to determine the accuracy.

Feature (Read counts + ...)	Gate 1 w/o	Gate 2 w/o	Gate 3 w/o	Gate 1 w/	Gate 2 w/	Gate 3 w/
RSSI max	100	100	100	82.2±2.0	95.8±1.0	87.4±1.7
RSSI mean	100	100	100	83.2±1.9	96.1±1.0	88.2±1.6
RSSI median	100	100	100	82.9±1.9	95.8±1.0	87.9±1.7
Phase std	100	100	100	83.7±1.9	82.9±1.9	79.8±2.1
Phase var	100	100	100	86.9±1.6	82.9±1.9	78.7±2.1
Phase mad	100	100	100	84.3±1.9	82.9±1.9	80.1±2.0

Table 6.5: Accuracy double input per individual scenario

The table indicates that the Naive Bayes Classifier works perfectly (100%) for any combination of Read Counts and another input feature at the scenarios without noise. For convenience, from now on this test will be disregarded for determining the best configuration, as there is no difference between configurations here. What is noticeable next is that the Phase Variance input feature has the highest accuracy (86.9±1.6%) for Gate 1 with noise. In addition, RSSI mean has the highest accuracy for Gate 2 (96.1±1.0%) and Gate 3 (88.2±1.6%) with noise. The accuracy of RSSI input features combined with Read Counts is generally higher than for Phase input features. This may be because the accuracy with RSSI instead of Phase as a single input feature also gives higher accuracy.

All scenarios combined
The accuracy of the NBC for Dock Door Discrimination when inputting a two features for all scenarios combined is given in Table 6.6. The left column gives the input feature (+ Read counts) and the right column gives the corresponding accuracy.

Feature (Read counts + ...)	Accuracy (%)
RSSI max	93.6±0.5
RSSI mean	93.0±0.5
RSSI median	92.6±0.5
Phase std	92.0±0.6
Phase var	92.3±0.6
Phase mad	92.1±0.6

Table 6.6: Accuracy double input feature for all scenarios combined

The table shows that for all scenarios combined, the Naive Bayes Classifier achieves the highest overall accuracy for the input feature Read Counts + RSSI max (**93.6±0.5%**). The accuracy for the RSSI inputs is higher than for the Phase input features, but the difference is much smaller than for a single input feature. This is because the combination with Read Counts overall provides higher accuracy and thus better Dock Door Discrimination. To see whether a combination of three input features can have an impact on accuracy, the next section combines Read Counts + RSSI max and other input features.

6.3.3. Accuracy triple input features

The accuracy of the Naive Bayes Classifier for Dock Door Discrimination with three input features (triple) are given in this section, first by individual scenario, then for all scenarios combined.

Individual scenarios
The accuracy of the NBC when inputting for three features is given per individual scenario, as shown in the tables 6.7, 6.8 and 6.9 below. The results are given per feature (+ Read counts), as shown in the left column, and the top row of the table. Each time, two input features are added to the Read Counts to determine the accuracy. The accuracy is given only for the three scenarios with noise at other Gates, as the other scenarios all have 100% accuracy.

Feature (Read counts + ...)	RSSI mean	RSSI median	Phase std	Phase var	Phase mad
RSSI max	77.1±2.1	77.4±2.1	78.7±2.1	81.1±2.0	79.0±2.1
RSSI mean		76.9±2.2	78.7±2.1	80.8±2.0	78.7±2.1
RSSI median			78.7±2.1	81.9±2.0	78.5±2.1
Phase std				82.9±1.9	77.7±2.1
Phase var					82.9±1.9

Table 6.7: Accuracy triple input test Gate 1 with noise

Feature (Read counts + ...)	RSSI mean	RSSI median	Phase std	Phase var	Phase mad
RSSI max	94.8±1.1	94.8±1.1	88.7±1.6	86.9±1.7	88.2±1.6
RSSI mean		92.9±1.3	88.7±1.6	86.1±1.8	88.2±1.6
RSSI median			88.5±1.6	84.8±1.8	87.9±1.7
Phase std				82.9±1.9	82.9±1.9
Phase var					82.9±1.9

Table 6.8: Accuracy triple input test Gate 2 with noise

Feature (Read counts + ...)	RSSI mean	RSSI median	Phase std	Phase var	Phase mad
RSSI max	86.4±1.8	86.9±1.7	85.0±1.8	85.6±1.8	85.0±1.8
RSSI mean		87.4±1.7	85.6±1.8	85.6±1.8	85.6±1.8
RSSI median			84.5±1.9	85.0±1.8	84.5±1.9
Phase std				78.0±2.1	76.9±2.2
Phase var					76.9±2.2

Table 6.9: Accuracy triple input test Gate 3 with noise

What is noticeable in the tables above is that the accuracy of the individual scenarios goes down compared to a dual input feature. For instance, in table 6.7, the maximum accuracy of three input features goes up to 82.9±1.9%, while the maximum accuracy for two input features is 86.9±1.6%. In table 6.8, the same thing happens as the maximum accuracy is now only 94.8±1.1% instead of 96.1±1.0%. In table 6.9 ditto, as the accuracy has dropped from 88.2±1.6% to 86.9±1.7%. This is caused by the ”overkill” now taking place. There are too many input features for the Naive Bayes Classifier to generate higher accuracy at the Dock Door Discrimination.

All scenarios combined
The accuracy of the NBC for Dock Door Discrimination when inputting a three features for all scenarios combined is given in Table 6.10. The left column gives the input feature (+ Read counts + RSSI max) and the right column gives the corresponding accuracy. The features Read Counts and RSSI max are already fixed, as they obtained the highest overall accuracy among the double input features.

Feature (Read counts + RSSI max + ...)	Accuracy (%)
RSSI mean	92.2±0.6
RSSI median	92.1±0.6
Phase std	93.3±0.5
Phase var	93.4±0.5
Phase mad	93.0±0.5

Table 6.10: Accuracy triple input feature for all scenarios combined

The table shows that for all scenarios combined, the Naive Bayes Classifier achieves the highest overall accuracy for the input feature Read Counts + RSSI max + Phase variance (**93.4±0.5%**). The overall accuracy for the Phase inputs is higher than for the RSSI input features. This is because the combination with Read Counts and RSSI max has no inputs of Phase yet. What is noticeable is that all accuracies for triple inputs are lower than the highest accuracy for double inputs (93.6%). This is because it gets too much input, causing the NBC to recognise patterns that were not there in the first place. This leads to a reduction in Dock Door Discrimination.

6.4. Comparison between current RFID system and Naive Bayes Classifier for DDD

The comparison between the accuracies of the current RFID system and Naive Bayes Classifier as a Dock Door Discrimination method shows the improvement. The comparison first considers the accuracy at individual scenarios, then for all scenarios combined of the DDD method. This comparison is shown in table 6.11 below.

Method	Current	NBC1	NBC2	NBC3
Gate 1 w/o	95.3±1.3	100	100	100
Gate 2 w/o	68.2±2.5	100	100	100
Gate 3 w/o	95.0±1.3	100	100	100
Gate 1 w/	83.2±2.2	89.0±1.6	86.9±1.6	82.9±1.9
Gate 2 w/	80.5±2.3	97.1±0.9	96.1±1.0	94.8±1.1
Gate 3 w/	74.1±2.4	91.1±1.5	88.2±1.6	87.4±1.7
All scenarios	82.1±0.8	92.6±0.5	93.6±0.5	93.4±0.5

Table 6.11: Comparison accuracy Dock Door Discrimination methods

The table shows the comparison with each scenario shown in the left column and the method used in the top row. In this, Current is the current system and NBC1, NBC 2 and NBC3 are the DDD method using the Naive Bayes Classifier with 1,2 and 3 input features respectively. After the six different scenarios, the overall accuracy of the DDD method is also given.

For the first three scenarios, the current system gives a lower accuracy than DDD methods with a Naive Bayes Classifier. From the high accuracy for the NBC (100%), it can be seen that the Dock Door Discrimination is very efficient in a simple transition, with no noise at other Gates. This is an improvement increasing to 5% for a Gate with one adjacent Gate, and increasing to 32% for two adjacent Gates.

The difference in accuracy between the Gate 1 and Gate 3 scenarios with noise is notable. Both accuracies are expected to be roughly equal, as both scenarios have one adjacent dock door, but mirrored. In the case of the current RFID transition system, there is a big difference, this may be due to a deviation when performing the tests. Only, these deviations also occur in reality, when forklifts do not drive ideally through the dock doors. This means it does represent situations that may occur in reality. The accuracy of both scenarios improves when the Naive Bayes Classifier is applied, especially with one or two input features. It is concluded from this that the Naive Bayes Classifier can better deal with the deviations in the data set and thus is better suited for different situations in Dock Door Discrimination.

Also, the scenario of Gate 2 with noise has higher accuracy than Gate 1 and Gate 3 with noise in the Naive Bayes Classifier, respectively. This means that the Naive Bayes Classifier can perform Dock Door Discrimination better with two adjacent Gates than with one adjacent Gate. This should ensure that the accuracy of the Naive Bayes Classifier for all scenarios combined becomes even higher the more dock doors are adjacent.

The accuracy for all scenarios combined improves in all cases of the different NBC methods compared to the current RFID system for Dock Door Discrimination. The accuracy is highest with the Naive Bayes Classifier with two input features, being Read Counts and RSSI max. The accuracy improves by 11.5% from the current RFID system and 1% compared to the Naive Bayes Classifier with one input feature. This is because with the input feature RSSI max, the proximity of the RFID tag also affects the Dock Door Discrimination, which is especially an advantage when tags need to pass at multiple doors simultaneously.

Remarkably, the NBC with one input feature works better for individual scenarios, but the NBC with two input features works better for overall accuracy. This is because the NBC1 is exposed to only one type of data set, where the same scenario is played each time. This ensures that there are fewer variations in the data set and therefore higher accuracy can be achieved. In the accuracy for all scenarios combined, the data from all scenarios is aggregated and thus the Dock Door Discrimination method should achieve high accuracy. This is found to be highest for two input features, which ensures that the overall accuracy over the current RFID transition system is improved from 82.1±0.8% to 93.6±0.5%.

6.5. Sensitivity Analysis

To find out the robustness of the results, a sensitivity analysis is performed on the Margin of Error. In such a sensitivity analysis, the degree of influence of the input factors on the output factor is determined. In sections 6.2 and 6.3 the Margin of Error is determined along with the accuracy, according to the input values and assumptions used in this study. In this section these input values and assumptions are changed to see what happens to the outcome of the output factor. The common output factor for the different methods is the Margin of Error on the accuracy of Dock Door Discrimination. The common input factors in this case are the accuracy of Dock Door Discrimination, the sample size of the collected data and the confidence interval assumption. All input factors affect the Margin of Error of the results.

6.5.1. Accuracy

The sensitivity analysis of the Margin of Error by the accuracy of Dock Door Discrimination indicates the extent to which accuracy affects the Margin of Error, as shown in figure 6.7. Here it is assumed that the Sample size of individual scenarios 1.470 and the Sample size of all scenarios combined is 8.820, according to equations 6.1 and 6.2 respectively.

$$Sample\ size\ individual\ scenarios = 30\ runs * 49\ RFID\ tags = 1.470 \tag{6.1}$$

$$Sample\ size\ all\ scenarios\ combined = 30\ runs * 49\ RFID\ tags * 6\ scenarios = 8.820 \tag{6.2}$$

Also, it is assumed the Confidence Level is 95%. Because the Sample size of all scenarios is larger than the individual scenarios, the Margin of Error of all scenarios is smaller. Because of the proportion of $p(1-p)$ in the formula of Margin of Error (equation 6.3), a large difference in accuracy between what falls inside and outside the Confidence Interval gives a smaller Margin of Error. As a result, for an accuracy of 50%, the Margin of Error is largest and becomes smaller as the accuracy moves toward 0% or 100%. This can be seen in the fact that the Naive Bayes Classifier has a lower Margin of Error than the current RFID transition system in section 6.4, as the accuracy is higher.

$$Margin\ of\ Error = z_\gamma * \sqrt{\frac{p(1-p)}{n}} \tag{6.3}$$

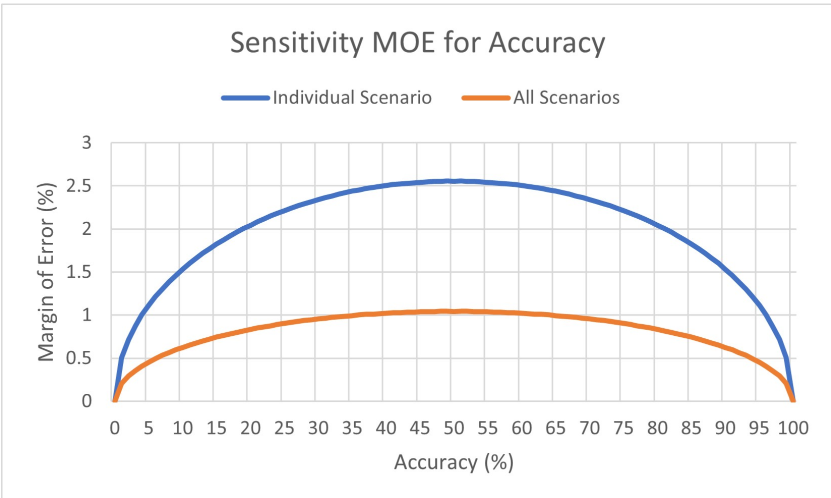


Figure 6.7: Sensitivity Analysis on Margin of Error for input Accuracy

6.5.2. Sample Size

The sensitivity analysis of the Margin of Error by the Sample Size of the collected data indicates the extent to which the Sample Size affects the Margin of Error, as shown in figure 6.8. Here it is assumed that the Sample Size of the individual scenarios is 1.470 at 100% and the Sample Size of all scenarios is 8.820 at 100%. For the accuracy of Dock Door Discrimination, 90% is assumed with a Confidence Level of 95%. Because the Sample Size of all scenarios at 100% is larger than at a Sample Size of 100% for the individual scenarios, the Margin of Error of all scenarios is smaller. Due to the proportion of n in the denominator of the formula for the Margin of Error (equation 6.3), the Sample Size has much influence at the smaller Sample Size of the individual scenarios compared to all scenarios. As the Sample Size becomes much larger, the influence on the Margin of Error becomes smaller and smaller. This can be seen in the fact that the Margin of Error for the Dock Door Discrimination methods is higher for all scenarios combined than per individual scenario, as the Sample Size is larger.

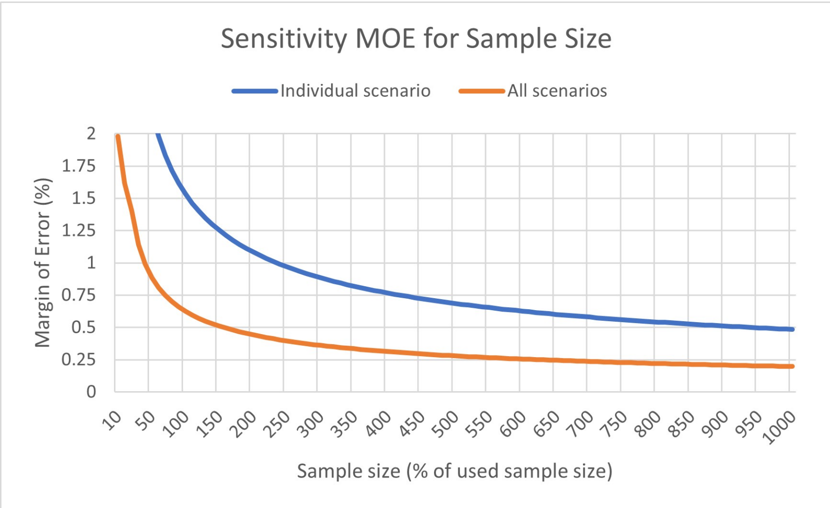


Figure 6.8: Sensitivity Analysis on Margin of Error for input Sample Size

6.5.3. Confidence Level

The sensitivity analysis of the Margin of Error by the Confidence Level of the true value of accuracy indicates the extent to which the Confidence Level affects the Margin of Error, as shown in figure 6.9. Here it is assumed that the Sample Size of the individual scenarios is 1.470 and the Sample Size of all scenarios is 8.820. For the accuracy of Dock Door Discrimination, 90% is again assumed. Because the Sample Size of all scenarios is larger than the Sample Size at the individual scenarios, the Margin of Error of all scenarios is smaller. The Confidence Level affects the critical z-score in the formula for the Margin of Error (equation 6.3). As a higher Confidence Level is required, the Margin of Error increases to ensure that more values lie in the Confidence Interval, which includes the true values of accuracy.

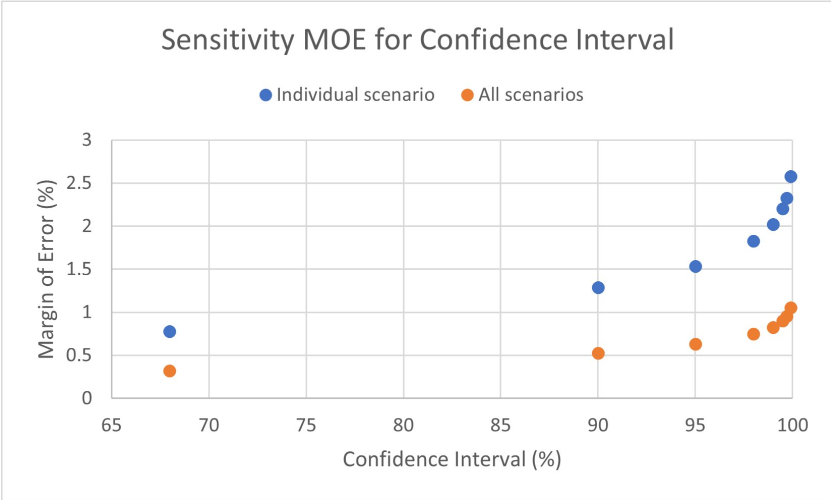


Figure 6.9: Sensitivity Analysis on Margin of Error for input Confidence Interval

6.6. Conclusion

This chapter has answered the sub-question: **What is the performance of the Naive Bayes Classifier compared to current RFID system?** Without Dock Door Discrimination method, the data points are plotted and it is possible to visually record what happens during the scenarios. If a tag is frequently read by a Gate and there is an increasing and decreasing trend in the number of reads, then the tag has been transitioned through that particular Gate. The current RFID transition system distinguishes the tags by looking at the number of reads on the front and back of the dock door. This is done only during a designated time window for the ”Front” and ”Back” antennas to determine the transition. A clear transition can be seen when the number of reads first increases and decreases at the front two antennas, then at the back two antennas.

The accuracy of the current RFID transition system is highest in scenarios where there is one adjacent Gate. Therefore, the system has a lot of difficulty with Dock Door Discrimination when the Gate in question is in the midst of two adjacent Gates. The accuracy for all scenarios combined is highest when the threshold value is set to $n = 9$ for this data set, being 82.1±0.8%.

The accuracy of Dock Door Discrimination using the Naive Bayes Classifier is approximated per background feature. The Naive Bayes Classifier is configured with either 1, 2 or 3 input features to determine the accuracy for Dock Door Discrimination. The accuracy per scenario is highest with the input feature Read Counts. The accuracy for all scenarios combined is highest with the combination of input features Read Counts + RSSI max, being 93.6±0.5%.

The comparison between the current RFID system for Dock Door Discrimination and the new system with the Naive Bayes Classifier, showed that it is an improvement. For individual scenarios, the Naive Bayes Classifier with one input feature (Read Counts) performs best. Also, it is concluded that the Naive Bayes Classifier can better deal with the deviations in the data set and thus is better suited for different situations in Dock Door Discrimination. Notable as well is that the Naive Bayes Classifier can perform Dock Door Discrimination better with two adjacent Gates than with one adjacent Gate. Accuracy for all scenarios combined is highest with the Naive Bayes Classifier with two input features (Read Counts + RSSI max). The final improvement of Dock Door Discrimination for all scenarios combined is from 82.1±0.8% to 93.6±0.5%.

A sensitivity analysis proved the robustness of the results. With the Margin of Error as the output factor whose influence of three input factors was investigated. The first is accuracy; for an accuracy of 50%, the Margin of Error is largest and becomes smaller as the accuracy moves toward 0% or 100%. Second is Sample Size; as the Sample Size becomes much larger, the influence on the Margin of Error becomes smaller and smaller. Third is the Confidence Level, as a higher Confidence Level is required, the Margin of Error increases to ensure that more values lie in the Confidence Interval.

Chapter 7

Conclusion and Future Work

7.1. Conclusion

The main research question answered during this study is: **What impact has implementing the Naive Bayes Classifier on RFID Dock Door Discrimination?** The purpose of RFID Dock Door Discrimination is to properly distinguish the RFID tags passing through the dock doors from location so that no errors are made in the delivery of products. One problem that arises here are cross-reads, when the tag is considered to have passed through multiple dock doors at the same time is considered to have passed through the wrong door or multiple doors at the same time. Another problem that arises are miss-reads, when the tag is not considered passed by any of the dock doors.

The current RFID transition system for Dock Door Discrimination uses the number of RFID reads to determine a transition at a dock door. To recognise a transition, a distinction is made between the two antennas at the front of the dock door and at the back of the dock door. A time window per antenna set can be used to determine whether the product (RFID tag) has actually transitioned through the dock door. The number of reads observed at the front and back of the dock door applies as a way to prevent the cross-reads, but undesirably creates more miss-reads. To compare the accuracy of the current RFID system with the new RFID system for Dock Door Discrimination, the KPI for accuracy is introduced. The current RFID system determines Dock Door Discrimination by setting a threshold value for the number of reads ”n”. When there are equal or more than ”n” readings at the front and back, the tag is considered to have passed through that dock door. The accuracy of the current RFID transition system for Dock Door Discrimination is determined by the number of clear transitions of a RFID tag out of the total number of RFID tags. The Margin of Error is used to determine the confidence interval of the results

In RFID technology, an RFID reader controls the RFID antennas. Through Ultra High Frequency radio waves, the RFID antennas communicate with the RFID tags to register the identity of the products. Low-level read data, recorded along with the identity of the RFID tag, is used to find out the direction of movement of the tag. As a Dock Door Discrimination method, the Bayes method has the most potential to determine at which location or through which door the RFID tags have passed. The probability that a tag is in a specific location is calculated by the observations of the multiple readers together. For this, the reader uses both the spatial and temporal observations of the tag.

The new design proposed for Dock Door Discrimination is divided into hardware and software design. For the hardware design, there is a detailed design for three consecutive dock doors according to the industry standard sizes. In this way, three different situations can be simulated; a dock door adjacent to the left, right and both sides of the respective dock door. The dock doors are fitted with 4 antennas at the dock door to register the moving tags and the 5th antenna, that is added compared to the current RFID transition system, registers the staging area, otherwise known as the static tags. For the software design the Naive Bayes Classifier is implemented. This classifier is entirely based on the Bayes theorem and compares the probability of a given tag being at a dock door with the probability of the tag being somewhere else and then assigns the value with the highest probability to it. With the RFID reader configuration and data transformation, data can be collected that is useful for the RFID Dock Door Discrimination with Naive Bayes Classifier. The accuracy of Dock Door Discrimination using the Naive Bayes Classifier is approximated per background feature. For this, the data set is first split into 75% train data and 25% test data. In order to determine the accuracy the good predictions are divided by the total number of prediction done by the Naive Bayes Classifier. Based on the Confidence level, sample proportion (accuracy) and sample size the Margin of Error is determined.

To compare the two RFID Dock Door Discrimination methods, data is collected with an Experimental Setup and Plan. The Experimental Setup is a 1:2 scaled version of the detailed design. The Experimental Plan is to transition through the three different dock doors, each time in a different configuration. Three scenarios without and with noise at other dock doors to gather a diverse data set. The sample size for performing the six different tests is set at 30 runs. With the RFID reader configuration and data transformation, a data set can be created per background feature with values per Gate, which the NBC then needs to start recognising patterns in for Dock Door Discrimination.

The accuracy of the current RFID system for Dock Door Discrimination is determined for the individual scenarios as well as all scenarios combined. Individually, the current RFID transition system struggles a lot with Dock Door Discrimination from the collected data set. Only at Gate 1 and Gate 3 without noise at other Gates gives reasonable accuracy at 95.3±1.3% and 95.0±1.3% respectively. The accuracy for all scenarios combined for the full data set of all six scenarios is 82.1±0.8%. The accuracy for the Naive Bayes Classifier is highest with Read Counts as the input feature for the individual scenarios. In each scenario, higher accuracy was obtained than with the current RFID transition system. Also, it is concluded that the Naive Bayes Classifier can better deal with the deviations in the data set and thus is better suited for different situations in Dock Door Discrimination. Notable as well is that the Naive Bayes Classifier can perform Dock Door Discrimination better with two adjacent Gates than with one adjacent Gate. The accuracy for all scenarios combined is remarkably highest for the Naive Bayes Classifier with Read Counts + RSSI max as input feature for the accumulated data set. This accuracy is 93.6±0.5% and thus an improvement of 11.5% with the current RFID transition system.

7.2. Future work

There are a few for future work to further improve RFID Dock Door Discrimination. The study found that the Naive Bayes Classifier achieves higher accuracy with two adjacent Gates than with one adjacent Gate. This means that higher accuracy can be achieved when there are several dock doors lined up in a row, because then there are more Gates with two adjacent Gates. A subsequent study could test whether this is actually true by testing the Dock Door Discrimination with the Naive Bayes Classifier on a larger scale with multiple adjacent dock doors.

Interesting for a next study could also be to actively test the Dock Door Discrimination method, i.e. not with an amassed data set but live during truck loading. This active Dock Door Discrimination will ensure live feedback whether products and pallets are loaded on the right trucks. This will make sure the errors are actively detected an can be handled accordingly.

References

Ahson, S. A., & Ilyas, M. (2017). *Rfid handbook: Applications, technology, security, and privacy*. CRC press.

Alfian, G., Syafrudin, M., Farooq, U., Ma'arif, M. R., Syaekhoni, M. A., Fitriyani, N. L., Lee, J., & Rhee, J. (2020). Improving efficiency of rfid-based traceability system for perishable food by utilizing iot sensors and machine learning model. *Food Control*, *110*, 107016.

Alfian, G., Syafrudin, M., Lee, J., & Rhee, J. (2019). Detecting movement and direction of tags for rfid gate. *2019 5th International Conference on Science and Technology (ICST)*, *1*, 1–5.

Alfian, G., Syafrudin, M., Yoon, B., & Rhee, J. (2019). False positive rfid detection using classification models. *Applied Sciences*, *9*(6), 1154.

Armo. (2018). Dock levellers: Standard or special dimensions? https://www.armoweb.it/en/insights/dock-levellers-standard-or-special-dimensions/

Buffi, A., Tellini, B., Motroni, A., & Nepa, P. (2019). A phase-based method for uhf rfid gate access control. *2019 IEEE International Conference on RFID Technology and Applications (RFID-TA)*, 131–135.

Chawla, V., & Ha, D. S. (2007). An overview of passive rfid. *IEEE Communications Magazine*, *45*(9), 11–17.

Chunli, L., & Donghui, L. (2012). Application and development of rfid technique. *2012 2nd International conference on consumer electronics, communications and networks (CECNet)*, 900–903.

Dekking, F. M., Kraaikamp, C., Lopuhaä, H. P., & Meester, L. E. (2005). *A modern introduction to probability and statistics: Understanding why and how* (Vol. 488). Springer.

Dobkin, D. (2012). *The rf in rfid: Uhf rfid in practice*. Newnes.

Duroc, Y., & Tedjini, S. (2018). Rfid: A key technology for humanity. *Comptes Rendus Physique*, *19*(1-2), 64–71.

Farid, D. M., Zhang, L., Rahman, C. M., Hossain, M., & Strachan, R. (2014). Hybrid decision tree and naïve bayes classifiers for multi-class classification tasks. *Expert Systems with Applications*, *41*(4, Part 2), 1937–1946. https://doi.org/https://doi.org/10.1016/j.eswa.2013.08.089

Fx series rfid reader integration guide. (2022). Zebra Technologies Corporation.

Gerstman, B. (2021). Two tails of z. https://www.sjsu.edu/faculty/gerstman/StatPrimer/z-two-tails.pdf

Goller, M. (2013). Probabilistic modeling in rfid systems.

Goller, M., & Brandner, M. (2011). Probabilistic modeling of rfid business processes. *2011 IEEE International Conference on RFID-Technologies and Applications*, 432–436.

Goller, M., & Brandner, M. (2012). Process-level localization of rfid tags using probabilistic models. *2012 IEEE International Conference on Wireless Information Technology and Systems (ICWITS)*, 1–4.

Hauser, M., Günther, S. A., Flath, C. M., & Thiesse, F. (2019). Towards digital transformation in fashion retailing: A design-oriented is research study of automated checkout systems. *Business & Information Systems Engineering*, *61*(1), 51–66.

How much does an rfid tag cost [Accessed: 2022-04-11]. (n.d.).

Jie, W., Minghua, Z., Bo, X., & Wei, H. (2018). Rfid based motion direction estimation in gate systems. *2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design ((CSCWD))*, 588–593.

Johnson, B. (2022). Warehouse management process flow. https://www.bluecart.com/blog/warehouse-process-flow

Joyce, J. (2003). Bayes' theorem.

Kaur, M., Sandhu, M., Mohan, N., & Sandhu, P. S. (2011). Rfid technology principles, advantages, limitations & its applications. *International Journal of Computer and Electrical Engineering*, *3*(1), 151.

Kelepouris, T., Harrison, M., & McFarlane, D. (2011). Bayesian supply chain tracking using serial-level information. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *41*(5), 733–742.

Keller, T., Thiesse, F., & Fleisch, E. (2014). Classification models for rfid-based real-time detection of process events in the supply chain: An empirical study. *ACM Transactions on Management Information Systems (TMIS)*, *5*(4), 1–30.

Keller, T., Thiesse, F., Ilic, A., & Fleisch, E. (2012). Decreasing false-positive rfid tag reads by improved portal antenna setups. *2012 3rd IEEE International Conference on the Internet of Things*, 99–106.

Keller, T., Thiesse, F., Kungl, J., & Fleisch, E. (2010). Using low-level reader data to detect false-positive rfid tag reads. *2010 Internet of Things (IOT)*, 1–8.

Khan, M. A., Sharma, M., & Prabhu, B. R. (2009). A survey of rfid tags. *International Journal of Recent Trends in Engineering*, *1*(4), 68.

Krishna, P., & Husak, D. (2007). Rfid infrastructure. *IEEE Communications Magazine*, *45*(9), 4–10.

Lyzs. (2022). Truck at a loading bay stock photo. https://www.westend61.de/en/imageView/LYF000406/truck-at-a-loading-bay

Ma, H., Wang, Y., & Wang, K. (2018). Automatic detection of false positive rfid readings using machine learning algorithms. *Expert Systems with Applications*, *91*, 442–451.

Majumder, P. (2020). Gaussian naive bayes. https://iq.opengenus.org/gaussian-naive-bayes/#:~:text=Gaussian20Naive20Bayes20supports20continuous,(independent20dimensions)20between20dimensions.

Mecalux, I. (2018). Designing warehouse layouts: What do you need to know? https://www.interlakemecalux.com/blog/designing-warehouse-layouts

Mojix. (2016). Dock Door Discrimination; Use Case 4.

Nova. (2022). Design the loading dock. https://www.novalocks.com/design-the-loading-dock/

Phrasant. (2020). Naive bayes classifier in python. https://www.kaggle.com/code/prashant111/naive-bayes-classifier-in-python#13.-Predict-the-results-

Prak, H. (2020). Confidence interval. https://www.geeksforgeeks.org/confidence-interval/

Qualtrics. (2022). Your guide to margin of error (with calculator). https://www.qualtrics.com/experience-management/research/margin-of-error/

Reva. (2006). *Solving the dock door cross read problem in dense reader environments* (tech. rep.). Reva Systems Corporation.

Scharwächter, V. (2022). Betrouwbaarheidsinterval berekenen: Formules amp; voorbeelden. https://www.scribbr.nl/statistiek/betrouwbaarheidsinterval/

Stertil. (2022). Het laadperron inrichten. https://stertil-dockproducts.com/nl/dock-producten/inrichten-laadperron/het-perron

Sunol, H. (2022). The principles of a warehouse layout design. https://articles.cyzerg.com/warehouse-layout-design-principles

Tran, T., Sutton, C., Cocci, R., Nie, Y., Diao, Y., & Shenoy, P. (2009). Probabilistic inference over rfid streams in mobile environments. *2009 IEEE 25th International Conference on Data Engineering*, 1096–1107.

Want, R. (2006). An introduction to rfid technology. *IEEE pervasive computing*, *5*(1), 25–33.

Weinstein, R. (2005). Rfid: A technical overview and its application to the enterprise. *IT professional*, *7*(3), 27–33.

Xu, H., Ding, Y., Li, P., Wang, R., & Li, Y. (2017). An rfid indoor positioning algorithm based on bayesian probability and k-nearest neighbor. *Sensors*, *17*(8), 1806.

Yuan, Y., & Yu, D. (2011). Reader planning in uhf rfid application. *2011 IEEE International Conference on RFID-Technologies and Applications*, 272–278.

Chapter A

Scientific Paper

In separate file

Increasing accuracy RFID Dock Door Discrimination with Naive Bayes Classifier

B.W. Berenschot

***Abstract*—Errors during truck loading at dock doors lead to unwanted wrong deliveries in logistics. Due to the falling price of RFID tags, RFID dock door discrimination is now being used for product registration. The problems that arise with the current system for RFID Dock Door Discrimination are cross-reads and miss-reads. The purpose of this research is to increase the accuracy of product registration by the proposed Dock Door Discrimination method with Naive Bayes Classifier (NBC). A hardware design, including 4 RFID antennas at three adjacent dock doors and 1 added antenna at the staging area, and software design, including the implementation of the NBC, are proposed to improve the RFID Dock Door Discrimination. The Experimental Setup and Plan were used to gather data to compare the current RFID transition system with the new proposed NBC system in six scenarios, three with and without noise at other gates. For each individual scenario, the accuracy improved most with NBC with one input feature. The accuracy for all scenarios combined for the collected data improved from 82.1±0.8% (current) to 93.6±0.5% (NBC).These results mean that there is a solid improvement in implementing the Naive Bayes Classifier over the current RFID transition system for Dock Door Discrimination.**

I. Introduction

A. Background

Logistics can be found in many places around the world. In particular, many logistics processes take place at large warehouses and distribution centres. One of these logistic processes is the loading of goods on trucks at the so-called dock doors. It is important to keep a good understanding of where products are and where the products should go, something called stock registration. Errors in the loading of trucks at the dock doors can lead to incorrect deliveries. As a result, customers may claim damages because the delivery is wrong or late, which is not desirable. It is therefore important to register the products in order to have proof of which product went into which truck. The hands-free automatic registration solution for identifying these products is Radio Frequency IDentification (RFID) technology. The operation of such an RFID technology relies on Ultra-High Frequency (UHF) radio waves to identify the RFID tags, which are attached to all products. Due to the falling price of RFID tags ([1]), RFID technology is increasingly being used for stock registration when loading different kinds of products on trucks. With the help of an RFID technology at the dock doors, it will be registered which products or crates have been identified and where they are stashed.

B. Problem

The problem that arises when installing an RFID system in a multiple dock door environment, is in the proximity of the dock doors. Figure 1 shows schematically how the installation of an RFID system is carried out at three contiguous dock doors. In this figure, the RFID antennas with their reading field are shown in red, the static products stored in the staging areas between the loading lines in yellow and the products in motion towards the truck by means of the arrow.

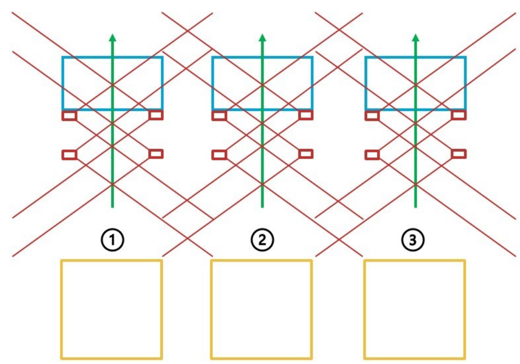


Fig. 1. Problem

As shown in the figure, the reading fields are not limited to just the dock door; RFID tags at other dock doors and staging areas are also read. This phenomenon is called **cross-reads** and makes dock door discrimination difficult. When the tag is not read, it is called **miss-reads**.

C. Goal

Other methods for RFID Dock Door Discrimination exist, such as the Satellite portal method [2], Zone Discrimination [3] and Metal shielding [4][5]. But because these are expensive alternatives [6], the logistics industry is looking for cheaper alternatives. A preliminary literature study revealed that a DDD method based on Bayes' probability theory has the most potential to make a positive impact as a solution to cross-readings and miss-readings in a multiple dock door environment. The objective of this study is therefor to increase the accuracy of RFID Dock Door Discrimination with the Naive Bayes Classifier. The main research question is: **What impact has implementing the Naive Bayes Classifier on RFID Dock Door Discrimination?** To answer the question, first a Process performance analysis of current RFID dock door system is carried out. Then a literature study on RFID technology and the relation to Bayes theorem is done. Subsequently, the proposed Design of RFID Dock Door Discrimination method with Naive Bayes Classifier is discussed. Data is then collected to be able to compare the Dock Door Discrimination methods with a Experimental Setup and Plan. Later, the results for the performance of the Naive Bayes Classifier is compared with the current RFID dock door system. At last, it is concluded what the impact of the Naive Bayes Classifier is on RFID Dock Door Discrimination, as well as recommendations for Future Works.

II. Analysis

A. Process

The outbound process flow in a warehouse, this is the process of storing the products that via picking end up in the staging areas and then loaded onto the truck. When zooming in on truck loading between staging and shipping, the current RFID system for Dock Door Discrimination is used. The current RFID system for one dock door is shown in figure 2. To determine a transition, a distinction is made between the first two antennas (light green) and the back two antennas (red). In this way, a distinction can be made between the number of reads "at the front" of the dock door and "at the back" of the dock door. Two factors play a major role here, the time window and the number of reads. The number of reads observed at the front and back of the dock door applies as a way to prevent the cross-readings, but undesirably creates more miss-readings.

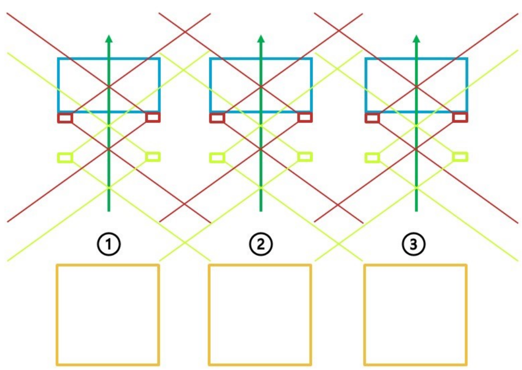


Fig. 2. Current RFID transition system for Dock Door Discrimination

The time window is the time used for the transition through the dock door from inside the warehouse to the truck. Here a time window of x seconds is used, this may differ according to the application of the RFID transition system. This time window determines how long the antennas on one side of the dock door are "on". This means that the antennas on the front side of the dock door are "on" for the first 2/3rds of the time window $[0\ s : \frac{2}{3}x\ s]$ and the back side for the second 2/3rds of the time window $[\frac{1}{3}x\ s : x\ s]$. This makes the middle 1/3rd of the time window the transition area, where the RFID tags are registered by both sets of antennas.

B. Performance

The Key Performance Indicator (KPI) Accuracy is used to make the methods quantifiable for comparison. The accuracy is determined using performance classifications. The results of the different tag movements are evaluated in four groups as follows:

- True Positives (tp_i)
- False Positives (fp_i)
- True Negatives (tn_i)
- False Negatives (fn_i)

After the results have been divided into the four groups, the test is evaluated in terms of accuracy. This key performance indicator for accuracy is determined as follows ([7]):

$$Average\ accuracy = \frac{\sum_{i=l}^{l} \frac{tp_i+tn_i}{tp_i+fn_i+fp_i+tn_i}}{l} \qquad (1)$$

The data used to compare the newly proposed method with the current method for Dock Door Discrimination is gathered using the Experimental Setup and Plan, which is mention in section V.

B. Accuracy and Margin of Error

The current RFID system determines Dock Door Discrimination by setting a threshold value for the number of reads "n". When there are equal or more than "n" readings at the front and back, the tag is considered to have passed through that dock door. The accuracy of the current RFID transition system for Dock Door Discrimination is determined by the number of clear transitions of a RFID tag out of the total number of RFID tags, as given in equation 2 with the proportion of static tags to moving tags is s to m tags.

$$Accuracy = \left(\frac{\frac{Number\ of\ "Yes"\ values}{Total\ number\ of\ values}}{m} + \frac{1}{s}\right) * Total tags * 100\% \qquad (2)$$

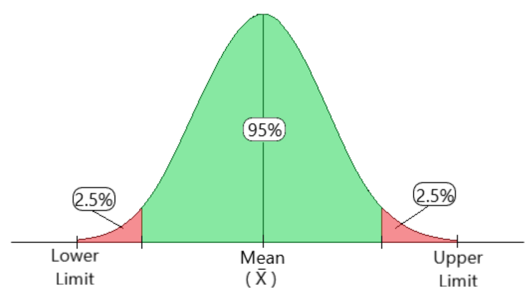


Fig. 3. 95 % Confidence Interval ([8])

To determine the inaccuracy of the results, a Margin of Error is used. In this study a confidence interval is assumed with a 95% confidence level, this indicates that 95 percent of the estimates are convinced to fall within the upper and lower bounds of the confidence interval, as shown in figure 3. With this information the Margin of Error can be determined using equation 3.

$$Margin\ of\ Error = z_\gamma * \sqrt{\frac{p(1-p)}{n}} \qquad (3)$$

III. Literature Review

A. Radio Frequency IDentification Technology

An RFID system consists of a number of components that are in contact with each other, being a reader, a tag, antennas and a host computer [9], shown in figure 4. The antennas are connected to the tags or the reader. In the case of the tag, the antenna is physically integrated. In addition, a tag also has an integrated circuit to provide the tag with its own identification and logic. In turn, the reader is either integrated with the antennas or is connected separately to the reader via cables. The antennas of the readers give energy to the tags via radio waves, this is called the downlink. The other way round, where the tag sends its energy together with its identification to the reader, is called uplink. The information that reaches the reader is often passed on to a connected computer. This computer is part of a communication network that makes it possible to process the data coming from the RFID system. A distinction is then made between three types of tag: active, passive and semi-passive. The active tag has its own power supply [10], the passive none [11] and the semi-passive only for the auxiliary electronics circuit [12]. The coupling method depends on the frequency of the operation. Low frequency and high frequency fall under near-field RFID, which is based on magnetic induction [13]. Ultra-high frequency and microwaves, on the other hand, fall under far-field RFID, where the working principle is based on electromagnetism [13].

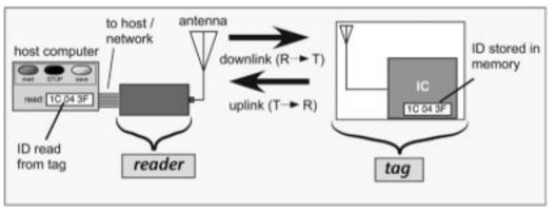


Fig. 4. RFID working principle

B. Low-level read data

In table I it is shown what low-level read data is retrieved during a tag read. This are subdivided into EPC, RSSI, Timestamp, Antenna and Phase angle per tag read. When the tags are detected, the direction of movement of the tags needs to be determined. Also, based on the low-level read data, extracted features are determined in order to be able to recognise the patterns and distinguish between different tag reads.

EPC	Timestamp	Antenna	RSSI	Phase angle
3032...7D	1,453,989,765.31	15	-59.0	3.50
3032...D1	1,453,989,765.31	15	-56.0	2.91
3032...7D	1,453,989,765.34	4	-69.0	2.72
3032...7D	1,453,989,765.34	17	-56.0	3.07

Tab. I
EXAMPLE OF LOWER-LEVEL FEATURES FROM TAG READ ([14])

C. Related works

In related works are discussed how methods based on Bayes theorem and other probabilistic methods are applied with RFID technology. The Process model localization method can estimate the location of a product in the supply chain at a given point in time [15]. he Tracking method can accurately determine future location of the tags based on past and present locations [16]. The indoor localization method determines the location of the target tags by the use of reference tags and by a Gaussian filter for abnormal RSSI values and Bayes probability in combination with a k-Nearest Neighbor algorithm [17]. The method of Inference combines multiple models in order to describe the local probability distribution for the RFID tags [18].

D. Bayes theorem for Dock Door Discrimination

The Bayes theorem can be used for Dock Door Discrimination, according to a white paper [19]. This white paper is the foundation of this study because it has the most potential for a Dock Door Discrimination method. Using Bayes' theorem, it is then determined at which location or through which door the RFID tags have passed. The probability that a tag is in a specific location is calculated by the observations of the multiple readers together using equation 4. For this, the reader uses both the spatial and temporal observations of the tag.

$$P(Loc < T, x, t > |Obs < T, R, t >) = \frac{P(Obs < T, R, t > |Loc < T, x, t >) x P(T \in x)}{P < T, R, t >} \quad (4)$$

IV. DESIGN

A. Requirements

The requirements to be met by the design are shown in Table II, both functional and non-functional.

Functional requirements	Non-functional requirements
Standard dock door dimensions	Zebra FX9600 readers
Standard antenna configuration	Mojix UHF RFID antennas
Dock Door Discrimination by software	Antenna cables
No extra hardware	RFID tags
Distinguishing moving to static	Internet connection
Better performance than current RFID system	Node-RED
	Python
	Pallets (boxes)
	Forklift (carriers)
	(ISB-profiles)

Tab. II
REQUIREMENTS (FUNCTIONAL AND NON-FUNCTIONAL)

B. System Design - Hardware

the overall hardware design is started in an overall entity such as a warehouse as shown in figure 5. 15 consecutive dock doors (crosses) are fitted with 4 antennas at the dock door to register the moving tags and the 5th antenna is added compared to the current RFID transition system, to register the products (RFID tags) in the staging area, otherwise known as the static tags.

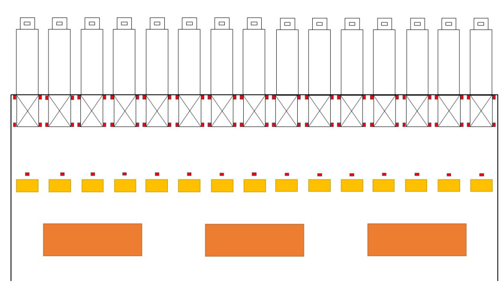


Fig. 5. Warehouse layout design

For the detailed system design, shown in figure 6, the standard sizes are considered. This means the dimensions for a dock door are 2.44 m in height and 2.74 m in width [20]. The distance between the centre lines of adjacent dock doors is at least equal to 3.70 m, and the space in between two adjacent dock doors is 1.26 m. The figure shows the front view of the three adjacent dock doors in the detailed system design, with the antennas indicated in red and the incoming and outgoing products with RFID tags indicated by the purple cross and dot symbol. By default, the antennas are at a 45-degree angle, as this is the conventional way that RFID systems are implemented in the real world, according to RFID experts at Mieloo & Alexander.

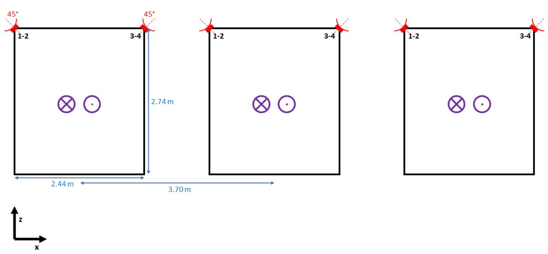


Fig. 6. Basic configuration and dimensions, front view

Figure 7 shows the top view of the detailed system design. The direction detection of the products (RFID tags), again shown in purple, is done by means of four antennas, to create a high accuracy. For this a margin of 2.5 m ([21]) between antennas 1-3 and 2-4 are given at the dock doors, which is the standard dimensions for the dock leveller in order to load the products into the truck. Then there is one antenna towards the staging area, where the static products (tags) are stored. The storage of the static products is shown by means of the yellow lining. These are positioned at 4.57 m from the dock door RFID system to provide enough space for vehicles such as forklifts to move through the passage way and straight into the dock doors to load the trucks ([22]), which is represented by the green arrows.

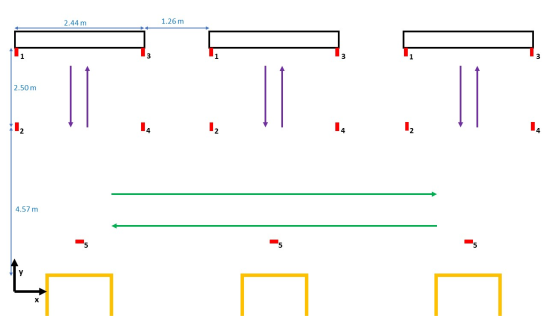


Fig. 7. Basic configuration and dimensions, top view

C. System Design - Software

The Naive Bayes Classifier is a form of machine learning that uses the Bayes theory in order to predict to which class the data points in a data set belong. The classes in this study entail the dock doors through which the RFID tags move. The advantages of Naive Bayes are that it is easy to use and the Classifier only needs one training data set to generate the class probability ([23]). The assumptions to be made with the Naive Bayes Classifier is that the Classes are independent and equal. This means that there is no dependency between the different features used in the NBC. This is quite possible in the case of Dock Door Discrimination since the features for Read Counts, RSSI and Phase of the reads are independent of each other. All features also contribute equally to the final outcome since no features are considered irrelevant. The Naive Bayes classifier predicts that the class label of instance X is the class C_i, if and only if

$$P(X|C_i)P(C_i) > P(X|C_j)P(C_j) \tag{5}$$

In Equation 5, $1 \leq j \leq m$ and $j \neq i$. That is the predicted class label is the class C_i for which $P(X|C_i)P(C_i)$ is the maximum probability. An example of classification using the Gaussian Naive Bayes Classifier is shown in figure 8. At each data point in this case, a z-score is given for each Class that is available. This z-score is based on the distance between that data point and the class-mean divided by the standard deviation of the class.

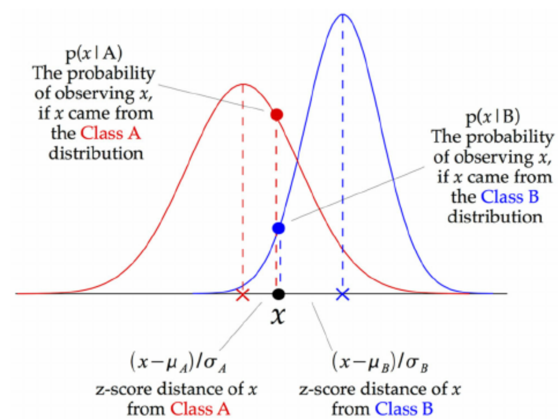


Fig. 8. Gaussian Naive Bayes Classifier ([24])

To implement the Naive Bayes Classifer in Python, the following steps of Phrasant are followed:

1) Import Libraries
2) Import Dataset
3) Exploratory Data Analysis
4) Declare feature vector and target variable
5) Split dat into separate training and data set
6) Encode categorical variables
7) Feature scaling
8) Model training
9) Predict the results
10) Check accuracy score

D. Configuration and Transformation

The reader configuration should ensure that the RFID antennas will transmit and capture the radio waves. The antennas are connected to the reader via antenna cabels and can thus transmit the information. To get the RFID reads required in a design, attention must be paid to setting up the RFID reader. To carefully set up a reader, the following points must be carefully determined:

- Endpoint configuration
- Mode configuration
- Node-RED to Data set configuration

Before the Dock Door Discrimination method with NBC can be applied, data transition has to be carried out. The data set coming from the reader configuration shows the Tag Data Events, shown in table I. The data of a test is divided into the number of runs the test has gone through. The observations are then divided into the gates where they are located, being Gate 1, Gate 2, Gate 3 or Static for static tags, shown in table III. The variable applicable at that time is examined. For this, the options are; read counts, RSSI and phase of the observations. For the RSSI and Phase values of the observations, different background features can be accessed to obtain inputs with the highest accuracy for the NBC.

Data	Description
EPC	Identification code of RFID Tag
Class	The Gate through which the RFID tag actually passed.
Gate 1	The value of the variable read by antennas from Gate 1.
Gate 2	The value of the variable read by antennas from Gate 2.
Gate 3	The value of the variable read by antennas from Gate 3.
Static	The value of the variable read by static antennas.

Tab. III
DATA SET INFORMATION

E. Performance Evaluation

After obtaining the wanted data set after the transformation, the set is divided into train data and test data. This division is equivalent to 75% training and 25% test data. The rows chosen for train and test data are completely random. To maximise the accuracy of the Naive Bayes Classifier, the different background features are used to determine the impact. These background features are shown below in table IV.

Variable	**Background feature**
Read Count	Sum
RSSI	Max, Mean, Median
Phase	Std dev., Variance, Mean Abs. Dev.

Tab. IV
DATA EVALUATION

The accuracy of the Naive Bayes Classifier depends on input variable, the background feature, which is used for Dock Door Discrimination. Through data analysis of the highest accuracies per input variable, the best configuration for the Naive Bayes Classifier in the case of this data set is determined. After this is figured out for a single input feature, combinations of two or three input features are also created. This is to see if the Naive Bayes Classifier for Dock Door Discrimination becomes even more accurate in that case. Based on the designated input feature, the Naive Bayes Classifier predicts a class at the corresponding tag, i.e. Gate 1, 2, 3 or Static. Afterwards, the prediction is compared with the real class, which is known beforehand. An example of this is shown in the confusion matrix set up per scenario. Obviously, when the tag is assigned to the right class, there is good Dock Door Discrimination. When a wrong class is assigned, there is wrong Dock Door Discrimination. The accuracy of the Dock Door Discrimination method with Naive Bayes Classifier is then determined by following equation 6. Margin of error is determined using equation 3 again.

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} * 100\% \quad (6)$$

V. EXPERIMENTAL SETUP AND PLAN

A. Experimental Setup

The experimental setup used during data collection is scaled at 1:2 of the hardware design. In figure 9, three adjacent dock doors are shown from a front view, the yellow, blue and green zone are gates 1,2 and 3 respectively. each dock door has four RFID antennas in the top corners, under a 45 degree angle. These are then spaced 1.22 m apart, and the dock doors are spaced 0.63 m. The height of antennas is 1.37 m in the scaled design. The blue and green square depict the RFID readers and can drive up to eight associated antennas. In this experimental setup; Gate 1 is blue antennas 1-2-4-5, Gate 2 is antennas 3-6 of both colours and Gate 3 is green antennas 1-2-4-5.

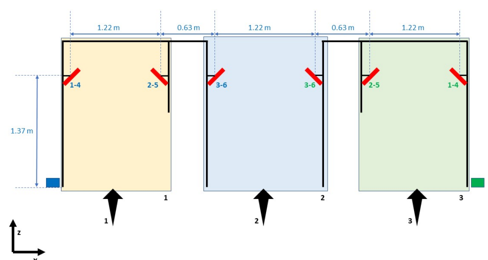


Fig. 9. Experimental Setup and Plan, Front view

In figure 10 the top view of the experimental setup is given. The distance between the front and back two RFID antennas of a dock door is 1.25 m (for instance for Gate 1 between antennas 1-2 and 4-5). At 2.30 m from the dock doors in the negative y direction are the RFID antennas which are meant to be reading the static tags. This fifth antenna is placed to distinguish between the moving and static RFID tags. The antenna with blue number 8 belongs to Gate 1, while the antennas with the green numbers 7 and 8 belong to Gate 2 and 3 respectively.

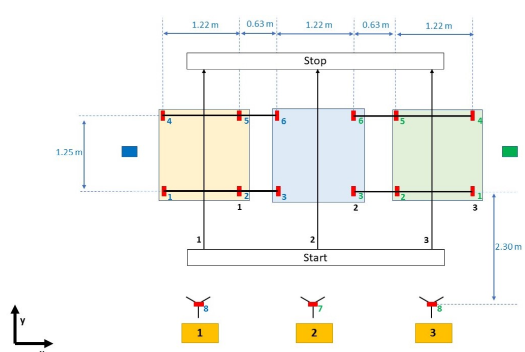


Fig. 10. Experimental Setup and Plan, Top view

B. Experimental Plan

To generate various data, several tests are carried out, all of which simulate different situations at the dock doors, going from the start to the stop area in figure 10 Boxes are used in conducting the tests. A number of RFID tags are stuck on these boxes to mimic a pallet of products. In total, there are 6 boxes, 3 of which are boxes for the moving RFID tags and 3 for the static tags. The moving boxes are provided with 12 RFID tags per box, 4 on the left, 4 in the middle and 4 on the right. These boxes are referred to as boxes 1, 2 and 3. The static boxes are provided with 4, 4 and 5 tags at boxes S1, S2 and S3 respectively. One tag left and right and 2 or 3 tags in the middle. Different scenarios are tested to obtain a diverse data set and to properly compare the Dock Door Discrimination methods. So for this purpose, different combinations of possible scenarios that could happen in reality are simulated. An overview of these can be found in table V.

Test	Dock Door 1	Dock Door 2	Dock Door 3	Staging Area
Gate 1 w/o	Box 1, 2, 3	-	-	13 RFID Tags
Gate 2 w/o	-	Box 1, 2, 3	-	13 RFID Tags
Gate 3 w/o	-	-	Box 1, 2, 3	13 RFID Tags
Gate 1 w/	Box 1 (moving)	Box 2 (static)	Box 3 (static)	Box S1, S2, S3
Gate 2 w/	Box 1 (static)	Box 2 (moving)	Box 3 (static)	Box S1, S2, S3
Gate 3 w/	Box 1 (static)	Box 2 (static)	Box 3 (moving)	Box S1, S2, S3

Tab. V
TEST SCENARIOS

In reality, the experimental setup and plan have been implemented. The result of the setup to simulate the three dock doors is shown in figure 11. The skeleton of the setup are the ISB profiles that guarantee the height and width of the antennas relative to each other. In total 15 UHF RFID antennas from Mojix are used to register the RFID tags at the dock doors. These antennas are connected via antenna cables to the RFID reader, Zebra's model FX9600. The configuration of these readers is discussed in the next section. One more antenna is placed per dock door at 2.30 m using a tripod in order to distinguish the static RFID tags. The boxes with the tags are walked through the dock doors using manual carriers.

Fig. 11. Real setup

C. Sample size

Before the different scenarios of testing can be carried out, the sample size has to be determined with equation 7. This sample size n is the minimum number of times one test must be run, otherwise the test is not sufficient.

$$n \geq (\frac{2z_{a/2}\sigma}{w})^2 \tag{7}$$

In this equation the minimal sample size is dependent on; the confidence interval α, σ the standard deviation, $z_{\alpha/2}$ is a the z-score determined by the confidence interval and w is the confidence level. A confidence interval is a percentage of the chance that a sample falls in the confidence level, being $w = 1 - \alpha$. The standard deviation σ is the square root of the variance, being the mean deviation of the data points. The z-score z has a constant value depending on the confidence interval.

The detectability of the RFID tags through the dock door is tested in 30 runs for the standard deviation to determine the sample size. In this process, a total of 49 RFID tags are tested for detectability. With the results, the formula is finally filled in. To determine the minimum number of tests, the formula is filled. For a confidence interval α of 95%, it means that the confidence level $w = 1 - 0.95 = 0.05$ and the z-score is 1.96. The standard deviation σ depends on the mean deviation, and then the root of that again, which comes out to 6.73% and thus σ = 0.0673. It follows that the sample size $n \geq 26.4$ must be adhered to. To be on the safe side, sample size 30 will be used for running the different scenarios of the testing plan.

D. Configuration and Transformation

To extract data from the tests, the reader must be configured. Those settings can be used to ensure that the RFID reads are converted to data points in a CSV file, making it usable for computer modeling ([25]). For the endpoint configuration, the reads are sent to an MQTT broker, which can transport data between different devices. For the mode configuration, the mode "Inventory" with a reporting interval 0 s ensures that each tag read is included in the data set. For the Node-RED to CSV file configuration, a structure was created that converts the data from the reads into a usable CSV data set, for programming with Python.

timestamp	antenna	epc	rssi	phase	reads	readername	reader
2022-11-10T15:33:48.403+0000	1.0	0012	-69.0	78.27997589111328	1.0	Reader77	/FX9600FCB1EA_ssl/tevents
2022-11-10T15:33:48.405+0000	1.0	0019	-63.0	-4.394665241241455	1.0	Reader46	/FX9600F445D8_ssl/tevents
2022-11-10T15:33:48.487+0000	4.0	005b	-74.0	-146.78732299804688	1.0	Reader77	/FX9600FCB1EA_ssl/tevents
2022-11-10T15:33:48.492+0000	4.0	001c	-73.0	-47.6656379699707	1.0	Reader77	/FX9600FCB1EA_ssl/tevents
2022-11-10T15:33:48.584+0000	7.0	0057	-48.0	76.1265869140625	1.0	Reader77	/FX9600FCB1EA_ssl/tevents
2022-11-10T15:33:48.570+0000	8.0	0010	-70.0	99.38536071777344	1.0	Reader46	/FX9600F445D8_ssl/tevents
2022-11-10T15:33:48.710+0000	3.0	001a	-72.0	-63.17331314086914	1.0	Reader46	/FX9600F445D8_ssl/tevents

Fig. 12. Tag Data Events

The Tag Data Events, example given in figure 12, are transformed to the specific value that can affect the Dock Door Discrimination method. An example of this is shown in table VI. Here, the first 5 observations are listed, where the sum of the read counts is the background feature. The Class indicates that the tag passed through Gate 1. Based on the values in the other columns, the Naive Bayes Classifier tries to find patterns to distinguish between the different Gates. Combinations of background features can also be made. In that case, all columns except Class are concatenated into a new input data set.

Class	gate1	gate2	gate3	Static
Gate1	18.0	8.0	13.0	
Gate1	17.0	13.0	9.0	7.0
Gate1	21.0	6.0	12.0	1.0
Gate1	19.0	5.0		
Gate1	25.0	2.0	2.0	

Tab. VI
MATRIX SUM READ COUNTS FOR SCENARIO DOCK DOOR 1 W/O NOISE

VI. RESULTS

A. Scatter Data Plots

Without a Dock Door Discrimination method, the scatter plot of the reads is shown in figure 13. A distinction is made between Gate 1,2 and 3 in the number of reads of an RFID tags over the period of one second. The figure clearly shows that the tag passed through Gate 1, as that is where the tag is read most often. What is also noticeable is that the sum of the number of reads first increases and then decreases again, clearly showing that Gate 1 has seen the tag first approach and then leave and thus has passed. The tag was also read at Gates 2 and 3, but not enough to question the actual direction.

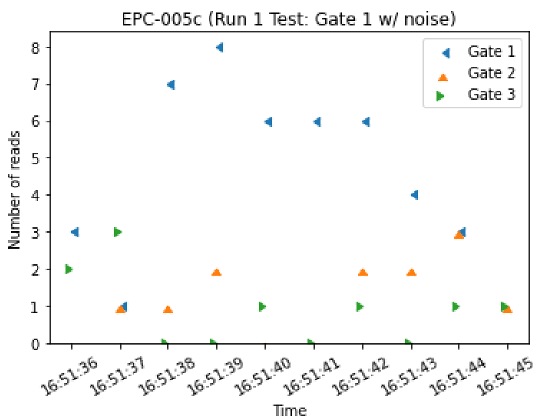


Fig. 13. Tag ending 005c during run 1 through Gate 1 with noise

The current system distinguishes the tags by looking at the number of reads on the front and back of the dock door. This is done only during a designated time window for the "Front" and "Back" antennas to determine the transition. A clear transition can be seen when the number of reads first increases and decreases at the front two antennas, then at the back two antennas, as shown in figure 14.

B. Accuracy Current RFID system

The accuracy of the current RFID transition system is highest in scenarios where there is one adjacent Gate. Therefore, the system has a lot of difficulty with Dock Door Discrimination when the Gate in question is in the midst of two adjacent Gates. The accuracy for all scenarios combined is highest when the threshold value is set to $n = 9$ for this data set, being 82.1±0.8%.

C. Accuracy Naive Bayes Classifier

The accuracy of Dock Door Discrimination using the Naive Bayes Classifier is approximated per background feature. The Naive Bayes Classifier is configured with either 1, 2 or 3 input features to determine the accuracy for Dock Door

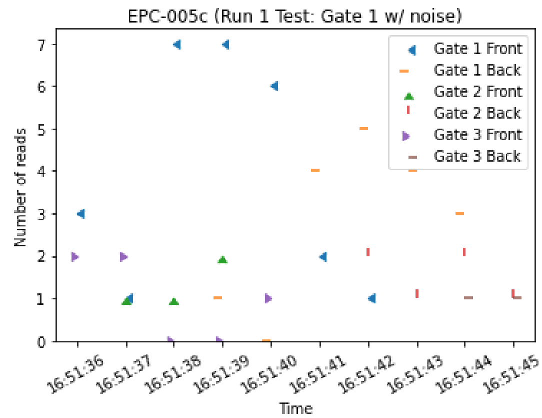


Fig. 14. Tag ending 005c during run 1 through Gate 1 with noise

Discrimination. The accuracy per scenario is highest with the input feature Read Counts. The accuracy for all scenarios combined is highest with the combination of input features Read Counts + RSSI max, being 93.6±0.5%.

D. Comparison

Table VII shows the comparison with each scenario shown in the left column and the method used in the top row. In this, Current is the current system and NBC1, NBC 2 and NBC3 are the DDD method using the Naive Bayes Classifier with 1,2 and 3 input features respectively. After the six different scenarios, the accuracy for all scenarios combined of the DDD method is also given.

Method	Current	NBC1	NBC2	NBC3
Gate 1 w/o	95.3±1.3	100	100	100
Gate 2 w/o	68.2±2.5	100	100	100
Gate 3 w/o	95.0±1.3	100	100	100
Gate 1 w/	83.2±2.2	89.0±1.6	86.9±1.6	82.9±1.9
Gate 2 w/	80.5±2.3	97.1±0.9	96.1±1.0	94.8±1.1
Gate 3 w/	74.1±2.4	91.1±1.5	88.2±1.6	87.4±1.7
All scenarios	82.1±0.8	92.6±0.5	93.6±0.5	93.4±0.5

Tab. VII
COMPARISON ACCURACY DOCK DOOR DISCRIMINATION METHODS

The comparison between the current RFID system for Dock Door Discrimination and the new system with the Naive Bayes Classifier, showed that it is an improvement. For individual scenarios, the Naive Bayes Classifier with one input feature (Read Counts) performs best. Also, it is concluded that the Naive Bayes Classifier can better deal with the deviations in the data set and thus is better suited for different situations in Dock Door Discrimination. Notable as well is that the Naive Bayes Classifier can perform Dock Door Discrimination better with two adjacent Gates than with one adjacent Gate. Accuracy for all scenarios combined is highest with the Naive Bayes Classifier with two input features (Read Counts + RSSI max). The final improvement of Dock Door Discrimination for all scenarios combined is from 82.1±0.8% to 93.6±0.5%.

E. Sensitivity Analysis

A sensitivity analysis proved the robustness of the results. With the Margin of Error as the output factor whose influence of three input factors was investigated, which equation is given in 3. The first is accuracy; for an accuracy of 50%, the Margin of Error is largest and becomes smaller as the accuracy moves toward 0% or 100%. Second is Sample Size; as the Sample Size becomes much larger, the influence on the Margin of Error becomes smaller and smaller. Third is the Confidence Level, as a higher Confidence Level is required, the Margin of Error increases to ensure that more values lie in the Confidence Interval.

VII. Conclusion and Future Work

A. Conclusion

The accuracy of the current RFID system for Dock Door Discrimination both per individual scenario and overall determined. Individually, the current RFID transition system struggles a lot with Dock Door Discrimination from the collected data set. Only at Gate 1 and Gate 3 without noise at other Gates gives reasonable accuracy at 95.3$\pm$1.3% and 95.0$\pm$1.3% respectively. The accuracy for all scenarios combined for the full data set of all six scenarios is 82.1$\pm$0.8%. The accuracy for the Naive Bayes Classifier is highest with Read Counts as the input feature for the individual scenarios. In each scenario, higher accuracy was obtained than with the current RFID transition system. Also, it is concluded that the Naive Bayes Classifier can better deal with the deviations in the data set and thus is better suited for different situations in Dock Door Discrimination. Notable as well is that the Naive Bayes Classifier can perform Dock Door Discrimination better with two adjacent Gates than with one adjacent Gate. The accuracy for all scenarios combined is remarkably highest for the Naive Bayes Classifier with Read Counts + RSSI max as input feature for the accumulated data set. This accuracy is 93.6$\pm$0.5% and thus an improvement of 11.5% with the current RFID transition system.

B. Future Work

There are a few for future work to further improve RFID Dock Door Discrimination. The study found that the Naive Bayes Classifier achieves higher accuracy with two adjacent Gates than with one adjacent Gate. This means that higher accuracy can be achieved when there are several dock doors lined up in a row, because then there are more Gates with two adjacent Gates. A subsequent study could test whether this is actually true by testing the Dock Door Discrimination with the Naive Bayes Classifier on a larger scale with multiple adjacent dock doors. Interesting for a next study could also be to actively test the Dock Door Discrimination method, i.e. not with an amassed data set but live during truck loading. This active Dock Door Discrimination will ensure live feedback whether products and pallets are loaded on the right trucks. This will make sure the errors are actively detected an can be handled accordingly.

References

[1] V. Chawla and D. S. Ha, "An overview of passive rfid," *IEEE Communications Magazine*, vol. 45, no. 9, pp. 11–17, 2007.

[2] T. Keller, F. Thiesse, A. Ilic, and E. Fleisch, "Decreasing false-positive rfid tag reads by improved portal antenna setups," in *2012 3rd IEEE International Conference on the Internet of Things*, IEEE, 2012, pp. 99–106.

[3] Mojix, *Zone Dock Door Discrimination*, 2016.

[4] S. A. Ahson and M. Ilyas, *RFID handbook: applications, technology, security, and privacy*. CRC press, 2017.

[5] Y. Yuan and D. Yu, "Reader planning in uhf rfid application," in *2011 IEEE International Conference on RFID-Technologies and Applications*, IEEE, 2011, pp. 272–278.

[6] P. Krishna and D. Husak, "Rfid infrastructure," *IEEE Communications Magazine*, vol. 45, no. 9, pp. 4–10, 2007.

[7] G. Alfian, M. Syafrudin, U. Farooq, *et al.*, "Improving efficiency of rfid-based traceability system for perishable food by utilizing iot sensors and machine learning model," *Food Control*, vol. 110, p. 107 016, 2020.

[8] H. Prak, *Confidence interval*, Nov. 2020. [Online]. Available: https://www.geeksforgeeks.org/confidence-interval/.

[9] R. Want, "An introduction to rfid technology," *IEEE pervasive computing*, vol. 5, no. 1, pp. 25–33, 2006.

[10] R. Weinstein, "Rfid: A technical overview and its application to the enterprise," *IT professional*, vol. 7, no. 3, pp. 27–33, 2005.

[11] D. Dobkin, *The rf in RFID: uhf RFID in practice*. Newnes, 2012.

[12] M. A. Khan, M. Sharma, and B. R. Prabhu, "A survey of rfid tags," *International Journal of Recent Trends in Engineering*, vol. 1, no. 4, p. 68, 2009.

[13] M. Kaur, M. Sandhu, N. Mohan, and P. S. Sandhu, "Rfid technology principles, advantages, limitations & its applications," *International Journal of Computer and Electrical Engineering*, vol. 3, no. 1, p. 151, 2011.

[14] M. Hauser, S. A. Günther, C. M. Flath, and F. Thiesse, "Towards digital transformation in fashion retailing: A design-oriented is research study of automated checkout systems," *Business & Information Systems Engineering*, vol. 61, no. 1, pp. 51–66, 2019.

[15] M. Goller, "Probabilistic modeling in rfid systems," 2013.

[16] T. Kelepouris, M. Harrison, and D. McFarlane, "Bayesian supply chain tracking using serial-level information," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 5, pp. 733–742, 2011.

[17] H. Xu, Y. Ding, P. Li, R. Wang, and Y. Li, "An rfid indoor positioning algorithm based on bayesian probability and k-nearest neighbor," *Sensors*, vol. 17, no. 8, p. 1806, 2017.

[18] T. Tran, C. Sutton, R. Cocci, Y. Nie, Y. Diao, and P. Shenoy, "Probabilistic inference over rfid streams in mobile environments," in *2009 IEEE 25th International Conference on Data Engineering*, IEEE, 2009, pp. 1096–1107.

[19] Reva, “Solving the dock door cross read problem in dense reader environments,” Reva Systems Corporation, Tech. Rep., 2006.

[20] Stertil, *Het laadperron inrichten*, 2022. [Online]. Available: https://stertil-dockproducts.com/nl/dock-producten/inrichten-laadperron/het-perron.

[21] Armo, *Dock levellers: Standard or special dimensions?* Jul. 2018. [Online]. Available: https://www.armoweb.it/en/insights/dock-levellers-standard-or-special-dimensions/.

[22] Nova, *Design the loading dock*, 2022. [Online]. Available: https://www.novalocks.com/design-the-loading-dock/.

[23] D. M. Farid, L. Zhang, C. M. Rahman, M. Hossain, and R. Strachan, “Hybrid decision tree and naïve bayes classifiers for multi-class classification tasks,” *Expert Systems with Applications*, vol. 41, no. 4, Part 2, pp. 1937–1946, 2014, ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2013.08.089. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417413007100.

[24] P. Majumder, *Gaussian naive bayes*, Feb. 2020. [Online]. Available: https://iq.opengenus.org/gaussian-naive-bayes/#:~:text=Gaussian20Naive20Bayes20supports20continuous , (independent20dimensions)20between20dimensions..

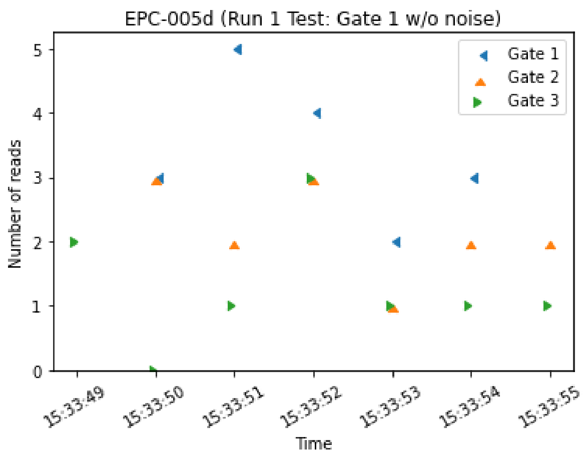
[25] *Fx series rfid reader integration guide*, Zebra Technologies Corporation, 2022.

Chapter B

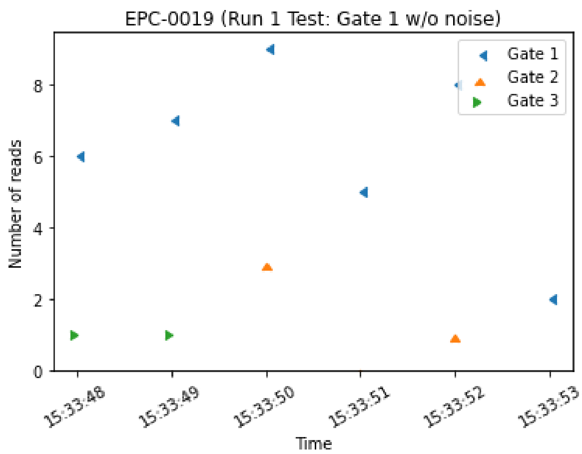
Plots

B.1. Without Dock Door Discrimination

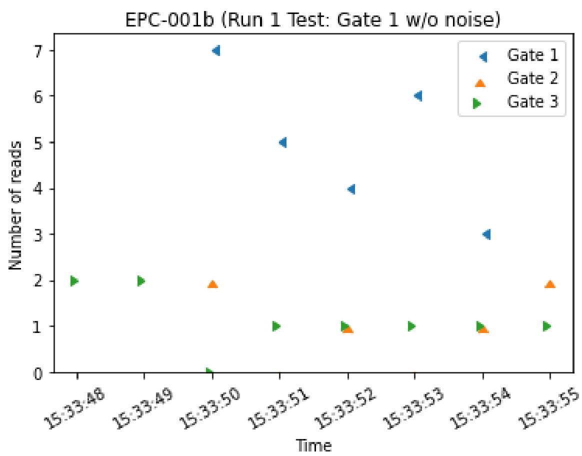
B.1.1. Gate 1 without noise



(a) Tag ending 005d during run 1 through Gate 1 without noise



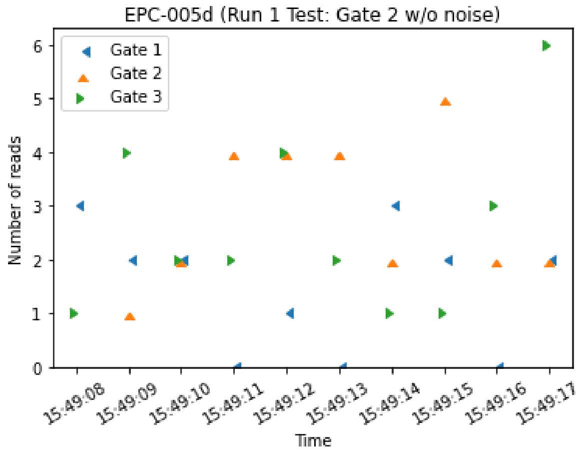
(b) Tag ending 0019 during run 1 through Gate 1 without noise



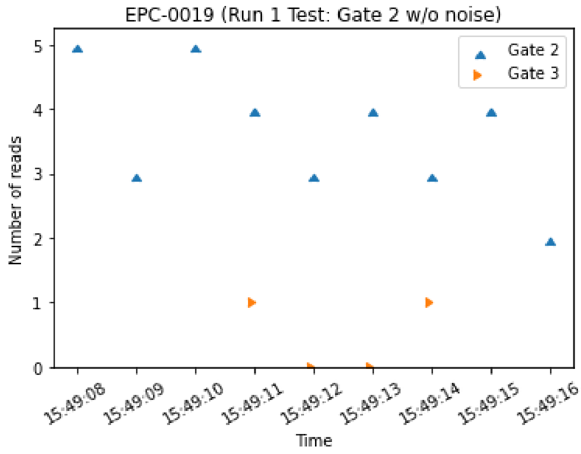
(c) Tag ending 001d during run 1 through Gate 1 without noise

Figure B.1: Plots for test: Gate 1 without noise

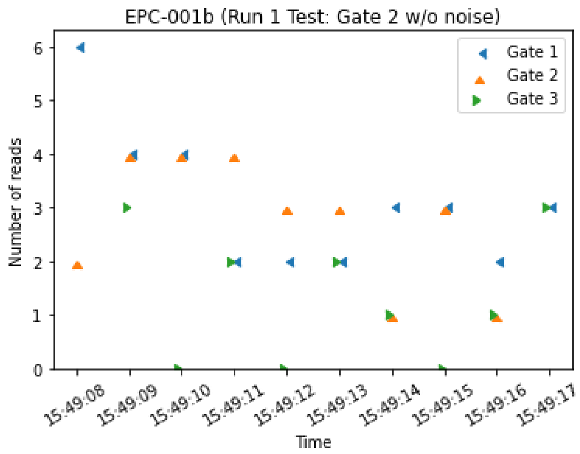
B.1.2. Gate 2 without noise



(a) Tag ending 005d during run 1 through Gate 2 without noise



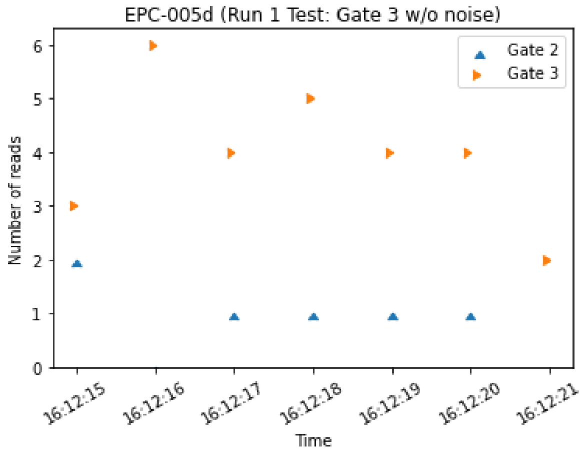
(b) Tag ending 0019 during run 1 through Gate 2 without noise



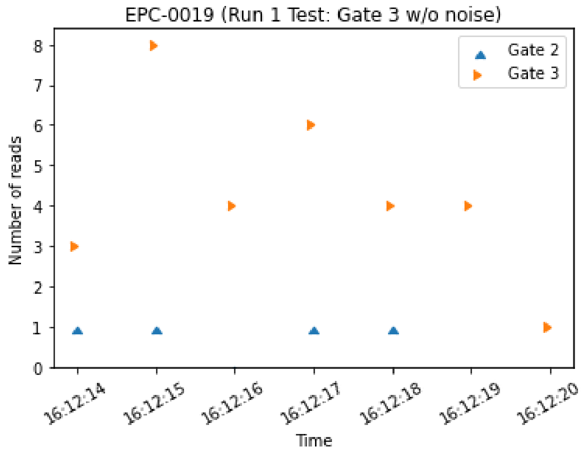
(c) Tag ending 001d during run 1 through Gate 2 without noise

Figure B.2: Plots for test: Gate 2 without noise

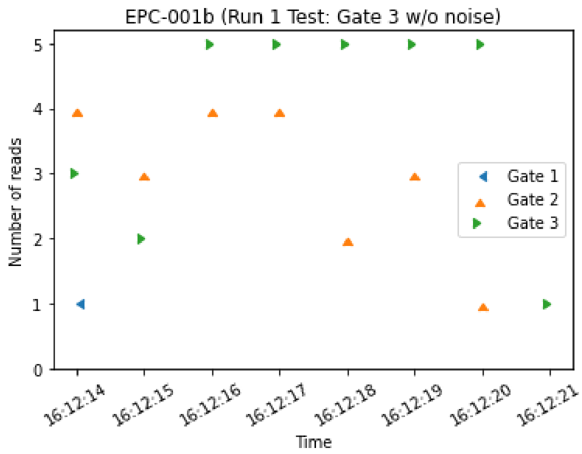
B.1.3. Gate 3 without noise



(a) Tag ending 005d during run 1 through Gate 3 without noise



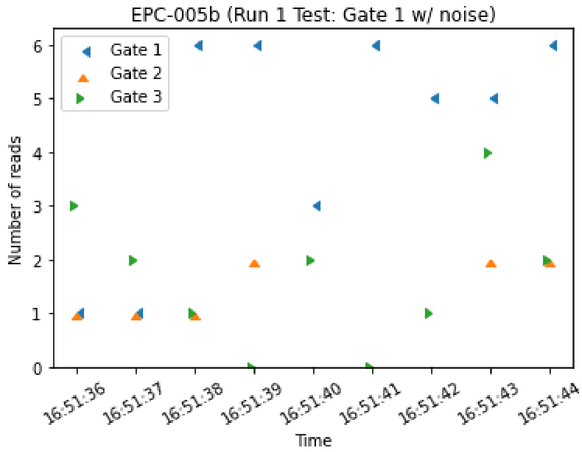
(b) Tag ending 0019 during run 1 through Gate 3 without noise



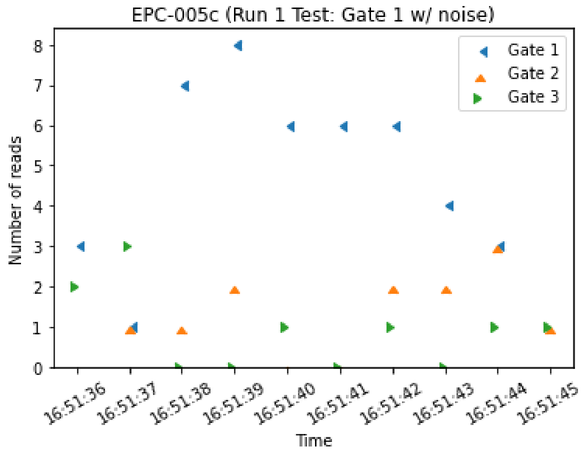
(c) Tag ending 001d during run 1 through Gate 3 without noise

Figure B.3: Plots for test: Gate 3 without noise

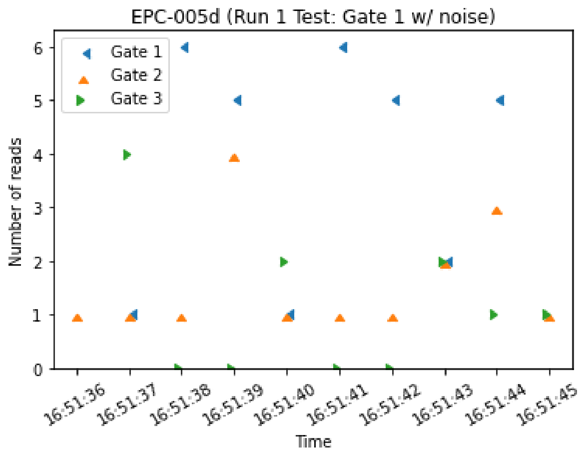
B.1.4. Gate 1 with noise



(a) Tag ending 005b during run 1 through Gate 1 with noise



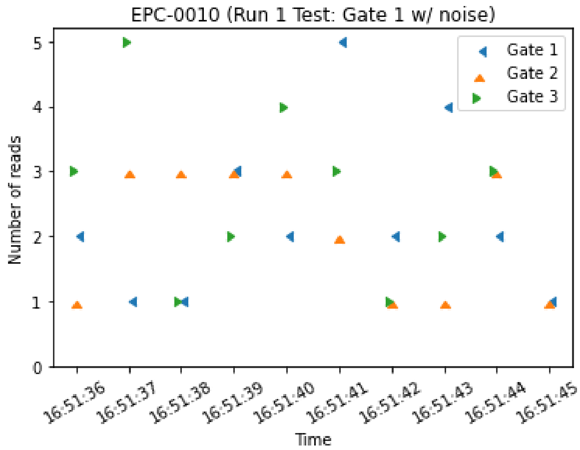
(b) Tag ending 005c during run 1 through Gate 1 with noise



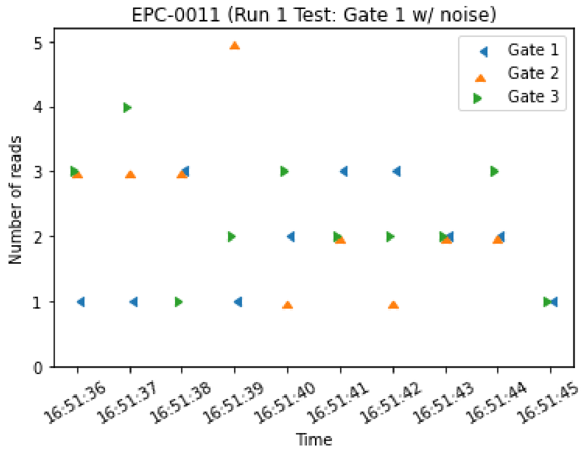
(c) Tag ending 005d during run 1 through Gate 1 with noise

Figure B.4: Plots for test: Gate 1 with noise

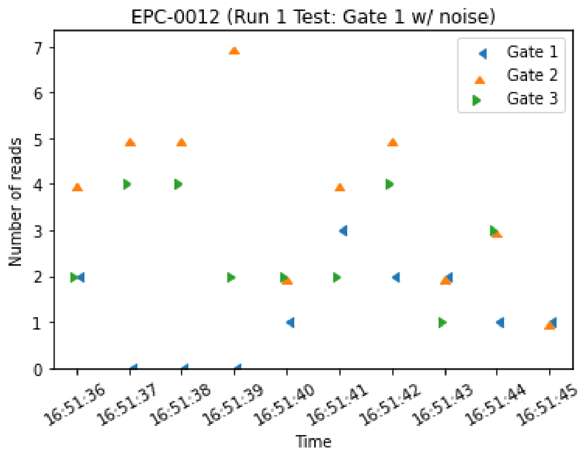
B.1.5. Gate 2 with noise



(a) Tag ending 0010 during run 1 through Gate 2 with noise



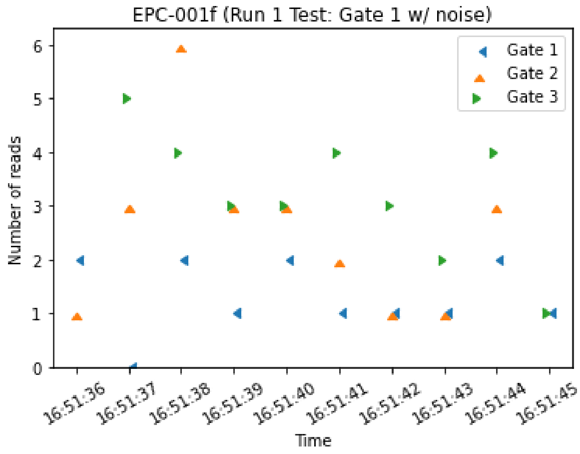
(b) Tag ending 0011 during run 1 through Gate 2 with noise



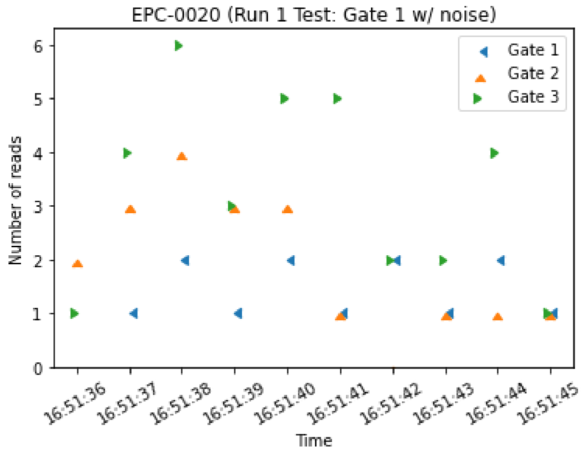
(c) Tag ending 0012 during run 1 through Gate 2 with noise

Figure B.5: Plots for test: Gate 2 with noise

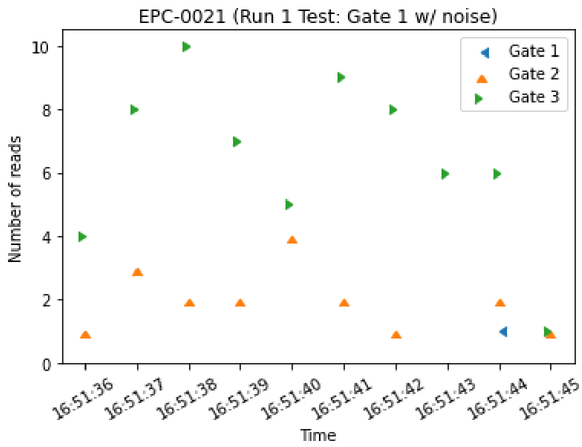
B.1.6. Gate 3 with noise



(a) Tag ending 001f during run 1 through Gate 3 with noise



(b) Tag ending 0020 during run 1 through Gate 3 with noise

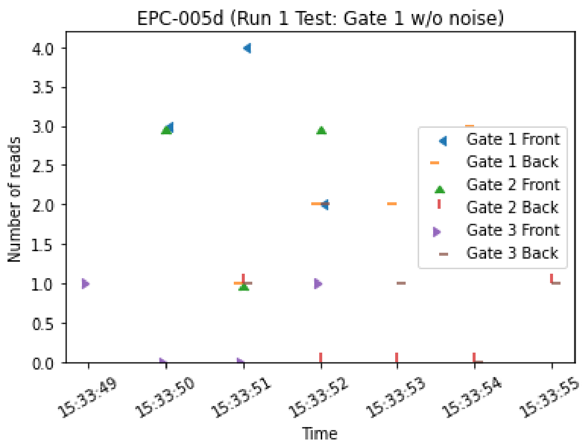


(c) Tag ending 0021 during run 1 through Gate 3 with noise

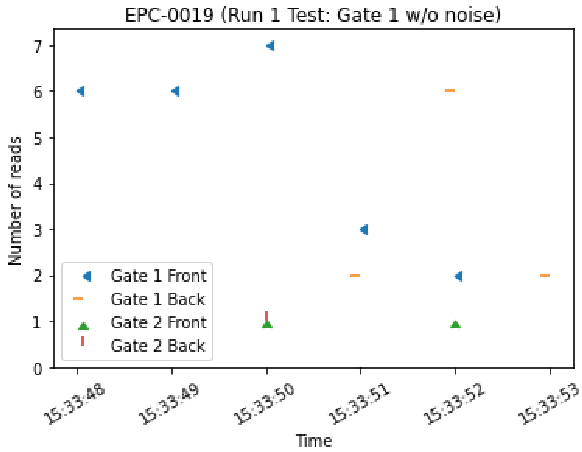
Figure B.6: Plots for test: Gate 3 with noise

B.2. Current RFID system for Dock Door Discrimination

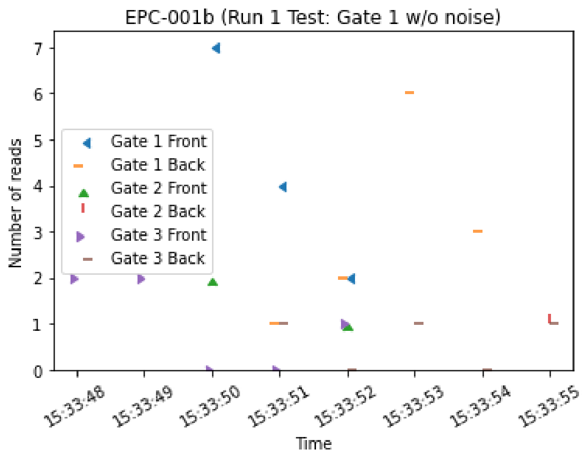
B.2.1. Gate 1 without noise



(a) Tag ending 005d during run 1 through Gate 1 without noise



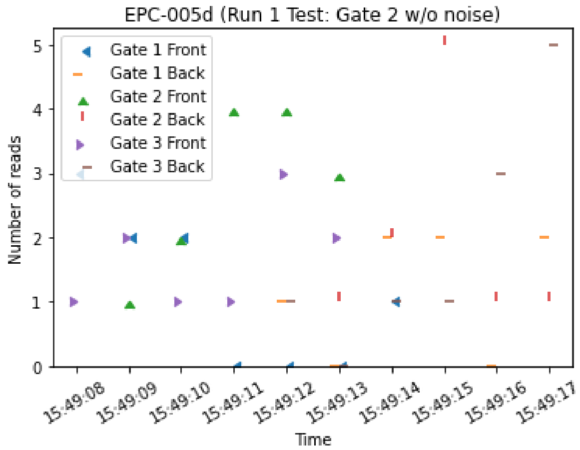
(b) Tag ending 0019 during run 1 through Gate 1 without noise



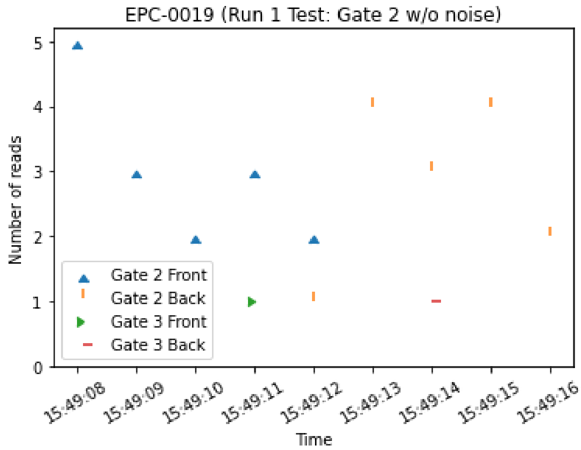
(c) Tag ending 001d during run 1 through Gate 1 without noise

Figure B.7: Plots for test: Gate 1 without noise

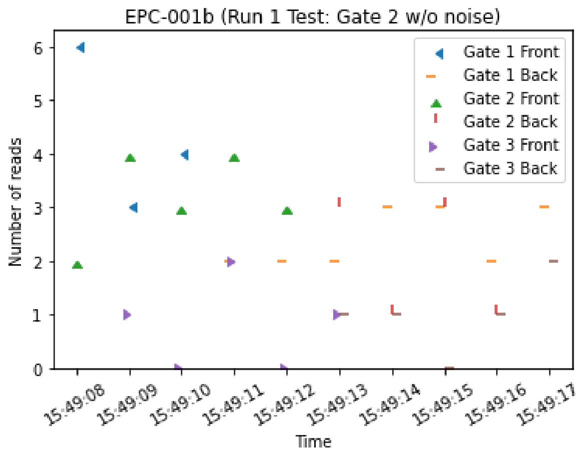
B.2.2. Gate 2 without noise



(a) Tag ending 005d during run 1 through Gate 2 without noise



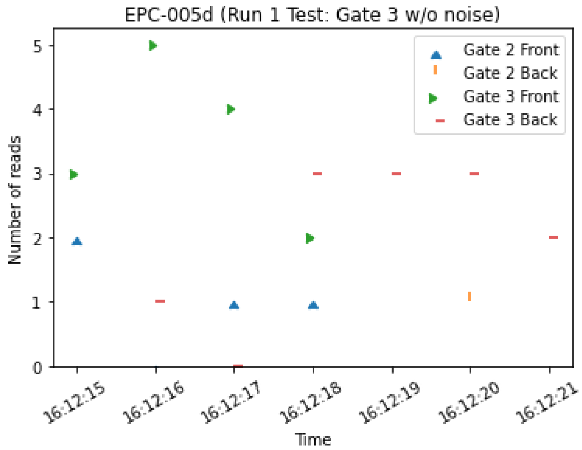
(b) Tag ending 0019 during run 1 through Gate 2 without noise



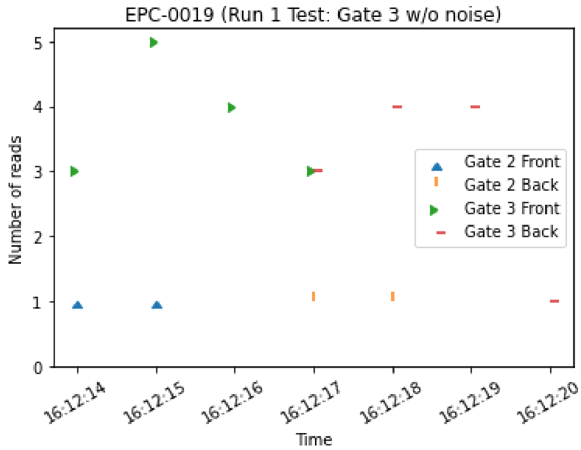
(c) Tag ending 001d during run 1 through Gate 2 without noise

Figure B.8: Plots for test: Gate 2 without noise

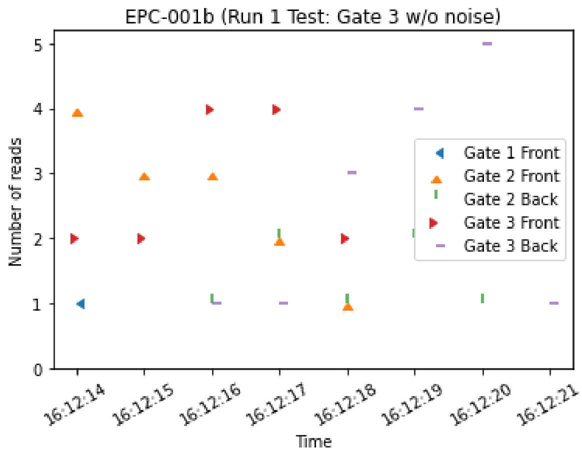
B.2.3. Gate 3 without noise



(a) Tag ending 005d during run 1 through Gate 3 without noise



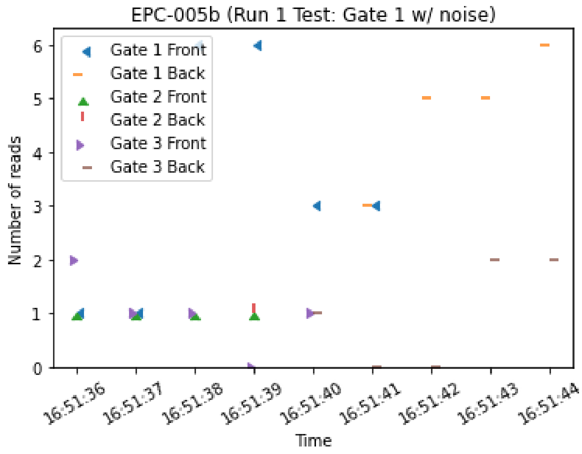
(b) Tag ending 0019 during run 1 through Gate 3 without noise



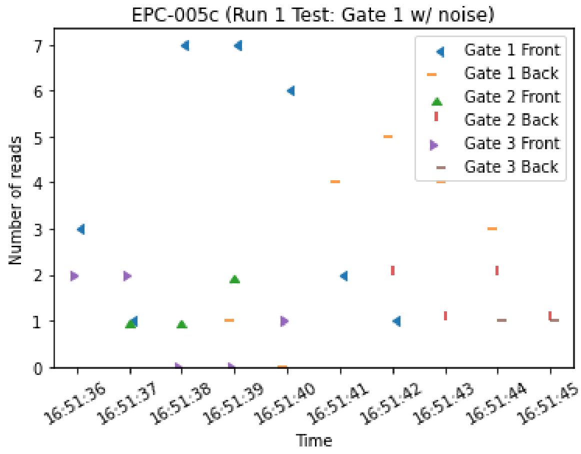
(c) Tag ending 001d during run 1 through Gate 3 without noise

Figure B.9: Plots for test: Gate 3 without noise

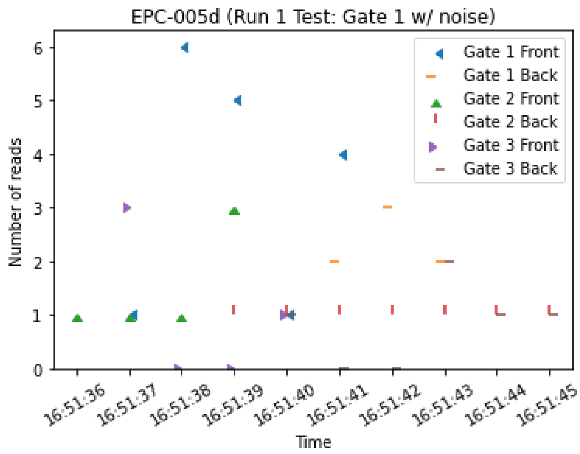
B.2.4. Gate 1 with noise



(a) Tag ending 005b during run 1 through Gate 1 with noise



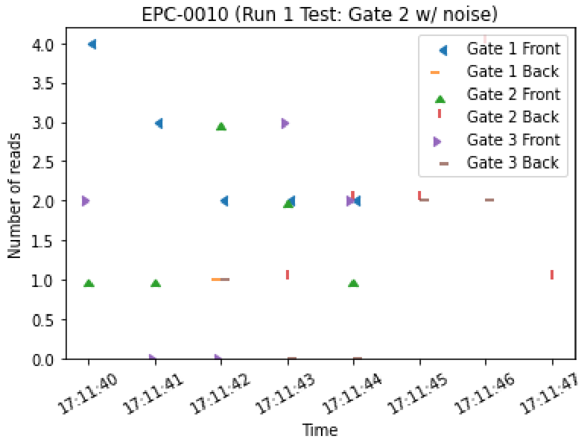
(b) Tag ending 005c during run 1 through Gate 1 with noise



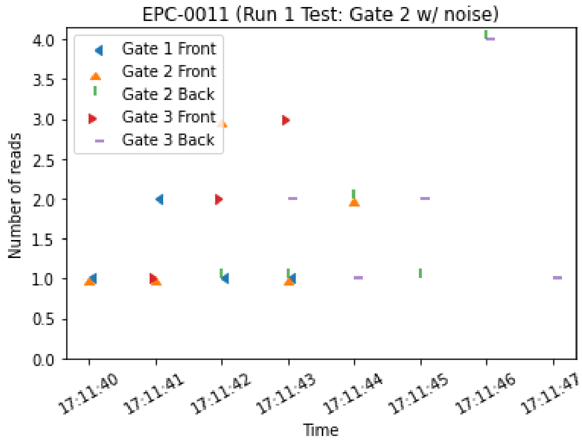
(c) Tag ending 005d during run 1 through Gate 1 with noise

Figure B.10: Plots for test: Gate 1 with noise

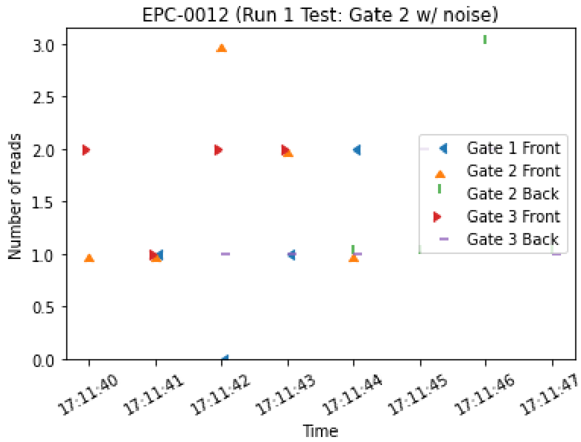
B.2.5. Gate 2 with noise



(a) Tag ending 0010 during run 1 through Gate 2 with noise



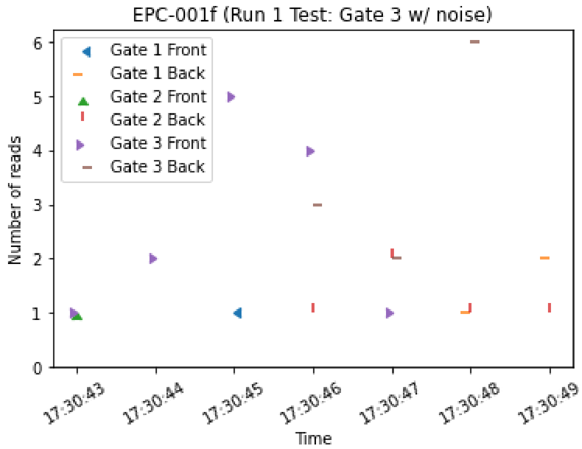
(b) Tag ending 0011 during run 1 through Gate 2 with noise



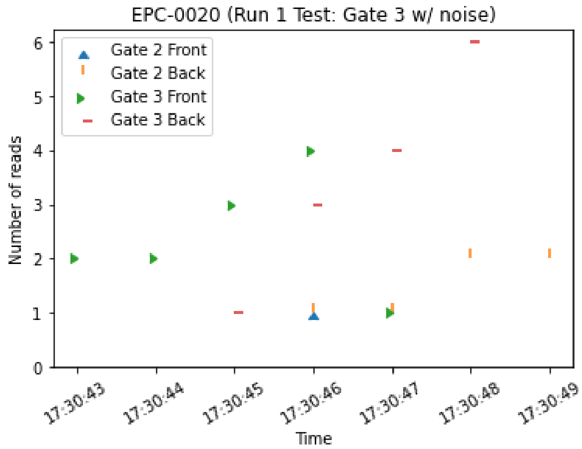
(c) Tag ending 0012 during run 1 through Gate 2 with noise

Figure B.11: Plots for test: Gate 2 with noise

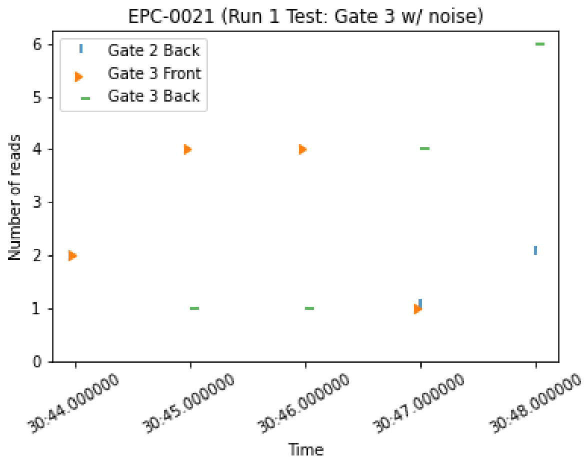
B.2.6. Gate 3 with noise



(a) Tag ending 001f during run 1 through Gate 3 with noise



(b) Tag ending 0020 during run 1 through Gate 3 with noise



(c) Tag ending 0021 during run 1 through Gate 3 with noise

Figure B.12: Plots for test: Gate 3 with noise