Localization and Tracking of UHF RFID Tags using Probabilistic Models

Michael Goller

September 2010

Graz University of Technology

Master Thesis

Localization and Tracking of UHF RFID Tags using Probabilistic Models

Master Thesis

at

Graz University of Technology

submitted by

Michael Goller

Institute for Electrical Measurement and Measurement Signal Processing,
Graz University of Technology
A-8010 Graz, Austria

30th September 2010

Advisor: Dipl. Ing. Dr. techn. Markus Brandner

TUG

Masterarbeit

Lokalisation und Verfolgung von UHF RFID Tags unter Verwendung Probabilistischer Modelle

Masterarbeit

an der

Technischen Universität Graz

vorgelegt von

Michael Goller

Institut für Messtechnik und Messsignalverarbeitung,
Technische Universität Graz
A-8010 Graz

30. September 2010

Begutachter: Dipl. Ing. Dr. techn. Markus Brandner

TUG

Abstract

During the last few years, Radio Frequency Identification (RFID) has gained increasing popularity in supply chains for tracking the state of goods. Recent developments and achievements together with a decrease of prices enable the tagging of goods on item level in various scenarios. With the increasing complexity of supply chains, also the requirements to RFID systems regarding localization and tracking have become more complex.

This work presents an approach for localization and tracking of RFID tagged items in practical applications based on probabilistic considerations. The novelty of this work is the fusion of different information sources of an RFID system in a probabilistic framework. Hidden Markov Models (HMMs) are used to model and classify RFID read events and sensor signals in an identification point model. In addition to that, a probabilistic model based on HMMs is used to consider information from the business layer. The appropriate fusion of the suggested models allows for a reliable localization and tracking of RFID tags in practical applications.

Keywords: RFID, localization, business process modeling, Hidden Markov Models

Kurzfassung

In den letzten Jahren verzeichnete die Radio Frequency Identification (RFID) Technologie einen enormen Aufschwung im Bereich der Logistik. Die rasant fortschreitenden Entwicklungen im Bereich von Transpondern und Lesegeräten ermöglichen es, einzelne Güter in Lieferketten von der Produktion bis zum Verkauf zu verfolgen. Die ständig wachsende Komplexität in Geschäftsprozessen und Lieferketten stellt dabei auch immer größere Herausforderungen an die eingesetzten RFID Systeme hinsichtlich der Lokalisierung und der Nachverfolgung einzelner Güter.

Diese Arbeit beschäftigt sich mit der Lokalisierung von RFID Transpondern in praktischen Anwendungsfällen auf Basis probabilistischer Modelle. Dazu werden die verschiedenen Informationsquellen eines RFID Systems in einem probabilistischen Framework kombiniert. Dieses Framework beinhaltet ein auf Hidden Markov Modellen basierendes Klassifikationsverfahren für RFID Lese-Events und Sensorsignale. Zusätzlich wird ein probabilistisches Prozess-Modell, ebenfalls basierend auf Hidden Markov Modellen, für die Miteinbeziehung von Informationen aus dem Business-Prozess Layer verwendet. Die geeignete Kombination dieser beiden Informationsquellen ermöglicht eine zuverlässige Lokalisierung von RFID Transpondern in praktischen Anwendungen.

Stichwörter: RFID, Lokalisierung, Business Prozess Modellierung, Hidden Markov Modell

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Place	Date	Signature

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Ort	Datum	Unterschrift

Contents

List of Figures

Acknowledgements

First of all, I would like to thank my advisor, Dr. Markus Brandner for his academic and personal guidance during the creation of my Master Thesis. In countless discussions, his critical reflections on various theoretical and practical issues provided me with important insights and new ideas.

I would also like to thank DI Alexander Gauby, CTO of RF-iT Solutions GmbH for creating the curiosity for the field of RFID and his support over the last two years. My special thanks go to my colleagues Michael Faschinger, Mario Katusic and Thomas Kempter for their help and support.

Especially, I want to thank my flatmates and friends Manuela Kathrein and Philipp Mattle for a great time over the last five years, their personal support and their friendship.

Finally, I want to express my sincerest gratitude to my family for all the support and assistance during my studies.

Michael Goller

Graz, Austria, September 2010

1 Introduction

The first chapter of this work gives an introductory overview of field of Automated Identification, with a special focus on Radio Frequency Identification (RFID). For this purpose, section 1.1 covers the principles of automated identification whereas section 1.2 deals with the basics of RFID and RFID systems. Moreover, also the use of probabilistic models in the field of localization of RFID tags is motivated by a discussion of the various effects that introduce a random behaviour in RFID systems.

1.1 Automated Identification - Auto-ID

Walking into a store, buying clothes or food, books or DVDs is a normal task in our everyday life. Every product we buy usually has been transported a long way from manufacturing over different storage halls and warehouses before it was finally placed in the shelf of a store. Along this way, products need to be identified for logistic and management purposes. This is the point where Automated Identification (Auto-ID) comes into play. Auto-ID summarizes different technologies that are concerned with the collection of information about certain objects, e.g. goods in a supply chain. The key feature of Auto-ID systems is that data processing is performed directly without human interaction. In order to be able to identify objects, they are provided with an identifier which is nothing else but a unique number that is associated with the considered object. In the context of Auto-ID, there are several commonly used terms that shall be explained here for clarification. An **item** is a certain object of interest, which is meant to be identified by the Auto-ID system. The terms item and object can be used synonymously. As stated above, items are provided with an identifier that is stored in a machine readable format using a **tag** or **label**. A **reader** is a device that is capable of performing an identification of items by reading the item identifier. A well known example for such a setup is the standard optical bar code: Optical labels contain the item identifier and bar code readers are used for the identification. In commercial applications, items are most often **tagged** (provided with the label that contains the identifier) already during manufacturing. An **identification point** consists of one or more readers that are used to identify objects at certain critical locations where information needs to be acquired. Auto-ID systems have in common that they are employed in some kind of **business process**. In general, this term covers the organization of activities and tasks that produce a certain product or offer a certain service [8]. In the context of Auto-ID, business processes describe the flow of goods over different stages, such as manufacturing, storage and transportation. These stages in the process can be mapped to identification points of an Auto-ID system, where all incoming goods need to be identified in order to update their regarding the level of detail at which items are tagged. The first family of systems uses a unique identifier for each and every single item. This is referred to as **item level tagging** and offers a higher transparency at the cost of additional processing efforts. The second family of systems aggregates several items (possibly of the same type) to larger units which are then provided with an identifier. Regardless of the level of detail at which items in a process are tagged, the task of an Auto-ID system is to provide information that allows to determine the current position of an item in the business process, i.e. to **localize** the item. Consequently, **tracking** of an item is the continuous determination of its position in the business process.

To make an automated object identification possible, different kinds of machine readable identifiers are widely used in the industry, such as bar codes, magnetic stripes and RFID (Radio Frequency Identification) transponders. Over the last decades, bar code labels have dominated in industrial and commercial applications. Although the technology is not new, RFID tags have gained increasing popularity during the last years which is mainly caused by the decrease in tag prices and the constantly improving performance. Currently, there is a competition between bar code and RFID systems, each having its advantages and disadvantages. Bar codes, especially the so called One-dimensional bar code are ubiquitous in commercial applications. Almost every product can be identified by means of its Universal Product Code (UPC) which is stored using a bar code that is printed onto the considered item. Advantages of this kind of identification are: Bar codes are very cheap, can be placed on objects of almost arbitrary size and shape and moreover reading devices are in the low price segment. However, bar codes also have their drawbacks: One significant disadvantage is that identification by means of an optical label requires a clear line of sight between reader and label. Occlusion, dirt or degeneration of the carrier material of the optical label can have severe negative impact on the read performance. Another drawback is that bar code readers can only identify one bar code at a time. Given a large number of items, the identification has to be performed sequentially which is a time consuming process. Moreover, bar code systems using the UPC only allow for an identification of the type of object, but not of the actual item itself. For example, there is a unique number for all blue shirts of a certain manufacturer, but one can not distinguish between shirts of the same type. For this reason, real item level tagging is not possible.

RFID tags do not suffer from these drawbacks. Since radio waves are used for the communication between tag and reader, a line of sight connection is not necessary as long as the tag is not shielded by surrounding metal objects or water. By using an anti-collision mechanism, up to several hundred tags can be identified within a second which is simply impossible when using standard optical labels. Today, tag and label manufacturers offer a huge variety of tags, especially designed for specific applications in the industry, transportation and the retail sector. Lately also so called “on metal tags” have been developed that are especially designed for tagging metal objects.

Identification of goods on item level offers high transparency of the underlying business process. It is possible to track the history of an item back to its manufacturing place and date. On contrast to bar code tagging, which only allows for an identification of the type of object, RFID tags offer the possibility to track a single, specific item. From the business process point of view, knowing where a certain item is at a certain time instant is valuable information. Most RFID systems today to a certain extend offer the possibility to localize a tag. Identification points at critical locations (outgoing goods, incoming goods etc.) provide information about when a specific item left a factory or warehouse or was delivered to a shop. Due to increasing requirements to the supply chain, this information may however not be sufficient. Just in time / just in place delivery and individual packing need more advanced approaches. Individual packing for example makes an assignment between specific items and packaging units necessary. RFID technology has the potential to provide solutions to these requirements, however there are still challenging issues to overcome.

Whereas RFID tags were predicted to replace optical bar codes some years ago, today's trends say that both technologies will coexist in the future. On the one hand, RFID tags are superior when there is a large number of items to be identified within a short time, however they are much more costly than optical labels. Bar codes on the other hand offer a high reliability and allow to keep the costs for labels and reader hardware low. For this reason, most Auto-ID systems use a combination of both technologies. Packaging units such as boxes are most often tagged with optical labels, whereas the contained items are equipped with RFID tags.

After this general introduction to the field of Auto-ID, the remainder of this chapter covers a brief overview over the basic mechanisms in RFID systems and outlines problems in terms of localization and tracking.

1.2 Radio Frequency Identification - RFID

The history of contactless identification using radio waves ranges back to the 30ies of the last century and has its origins in warfare [13]. During World War I, microwave radars were used to detect incoming aircrafts by means of backscattered radio waves. The major question was whether the detected plane belonged to allied or hostile forces. To solve this problem, German pilots started with roll maneuvers in order to change the backscattered signal and indicate that they are allied forces. This is a first, very primitive transmission of a single bit (friend or foe) using backscattered radio signals.

The achievements in semiconductor industry enabled the development of a variety of different RFID transponders during the last decades. The most promising RFID technology for the identification of a large number of items within a supply chain operates in the Ultra - High Frequency (UHF) band around $860\,\mathrm{MHz}$. This frequency band implies two consequences for modern RFID systems. First, the read range of up to several meters is quite considerable. Whereas this is desired in some applications to enable a distant identification, there are also drawbacks as will be discussed later on. Second, the short wavelength of signals in this frequency band implies a compact tag size. The EPCGlobal Class 1 Generation 2 [10] standard is the latest standard for UHF RFID systems. The advances in this standard enabled item level tagging in virtually any application. The standard is built on the use of passive RFID tags, where passive indicates that the tag itself does not have a power supply. Instead, it gets energized by the incoming radio wave and uses this energy to backscatter information, such as its identifier or data stored in the memory. In order to solve the multiple access problem when there are several tags in the RF-field, an anti-collision mechanism roughly based on the Slotted Aloha protocol [30] is defined in the EPCGlobal standard. During a so called inventory round, a reader energizes all tags in the field of its antennae and singles out the tag responses (consisting of the tag identifier) in a sequential manner. For this purpose, the tags choose a random number which is decremented at the start of every new round. As soon as the number approaches zero, the tag sends its identifier as response to the reader request.

Due to well known effects in UHF radio channels, such as multipath propagation and fading [22], different problems which result in unreliable readings do arise. Since the EPC Class 1 Generation 2 standard uses passive RFID tags, the first problem is to establish a stable energy supply for each tag as it passes an identification point. Power regulations in the used frequency band limit the reader output power and hence careful antenna positioning and orientation is critical to solve this issue. The second problem is to receive the backscattered signal which is several orders of magnitude smaller than the reader's emitted signal. This problem is a major challenge for reader manufacturers which provide higher and higher receiver sensitivities (up to $-80\,\mathrm{dBm}$ in state of the art readers). An additional challenge results from the problem of the shared medium: Several tag responses may overlap, causing collisions on the channel which a reader can not resolve in general. Since the standardized anti-collision algorithm is based on picking random numbers, this additionally introduces randomized behaviour in the tag - reader communication.

To summarize, the use of UHF signals offers advantages such as a considerable read range and compact tag sizes, but there are also drawbacks like unreliable read events due to unpredictable wave propagation. Additionally, randomized behaviour is introduced by means of the anti-collision mechanism.

1.2.1 RFID Systems

Modern RFID systems consist of a variety of components and share a common layered structure. The general layer model of an RFID system is shown in figure 1.1. The central layer of every system is the middleware as connection between sensing devices and a backend system. Sensing devices in an RFID system can be RFID readers, bar code readers, scales or motion sensors, but also more common devices such as light barriers or ultrasound sensors for distance measurements. These sensing devices have in common that they acquire some information about the items subject to a certain business process. Light barriers and ultrasound sensors can

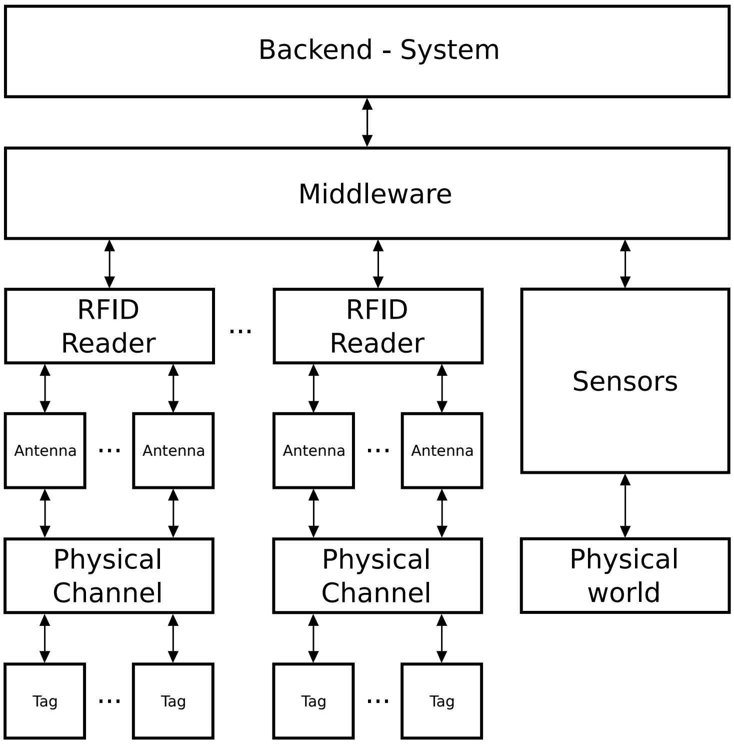


Figure 1.1: RFID system block diagram. The middleware can be considered as the central layer of an RFID system. It serves as an abstraction layer between sensing devices (RFID readers and sensors) that obtain information from the physical world and a backend system which is used to store and process information on a business process level. The task of the middleware is to manage the sensing devices, collect information from these devices and provide condensed information to the backend system.

be used to detect the presence of objects, measure their speed or for triggering RFID operations. Bar code readers are still widely in use to identify packaging units (such as boxes or trays) on conveyor belts. Finally, RFID readers acquire information about the RFID tagged items. On the one hand, this information consists of the item's identifier and possibly some memory content, on the other hand, an RFID reader can also provide information about the reading process itself. A read event for a specific tag consists of:

- The tag identifier I: As defined in the EPCGlobal standard, every tag carries an identifier of 96 or 240bit, which allows to assign a unique number to every single item in a process chain.
- Timestamp t: The time instant when the read event occurred. Usually, tag inventories are performed periodically, for this reason each tag will have more than one read event, separated by a unique timestamp.
- RSSI (Received Signal Strength Indicator) r: Information about the strength of the backscattered signal. Most state of the art RFID readers provide the RSSI information in a logarithmic scale (e.g. dBm), some even for the in-phase and quadrature phase component (I and Q) of the received signal.
- Antenna index i: Since most RFID readers have more than one antenna port which are used in a time multiplexed manner, also information about the antenna which inventoried a tag is provided.

In a more formal notation, a read event for a particular tag with identifier I, is a triple

$$\mathbf{e} = [t \quad r \quad i]. \tag{1.1}$$

A tag **inventory** is therefore defined as the creation of a triple $\mathbf{e}$ for a tag with identifier I as soon as the tag enters the read range of the antenna. Since the inventory is carried out periodically, each tag will have a series $\mathbf{E}$ of M consecutive read events, where M is also referred to as **read count**:

$$\mathbf{E} = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_M \end{bmatrix} = \begin{bmatrix} t_1 & r_1 & i_1 \\ t_2 & r_2 & i_2 \\ \vdots & \vdots & \vdots \\ t_M & r_M & i_M \end{bmatrix}. \tag{1.2}$$

The task of the middleware is to collect and evaluate the information acquired by the sensing devices and to report condensed information to the backend system. In general, the backend system is some kind of database that stores information about RFID tagged items, packaging units, orders and deliveries. In contrast to the inventory of a tag, the **identification** is defined as reporting a tag to the backend system. Based on a series of read events $\mathbf{E}$ and sensor signals $\mathbf{S}$, the middleware has to decide whether or not the tag will be reported to the backend system.

RFID systems are required to provide information about the current position of items subject to a certain business process, i.e. to localize the item in the process. In addition to that it is necessary to keep track of items as they move through the different stages of the process. Considering the limited reliability of RFID read events, a robust approach for localization and tracking is required.

1.2.2 Use Cases

In modern logistics and supply chains, individual requirements for deliveries also lead to increased requirements regarding the localization of items in the business process. In many applications it is desirable to know which packaging unit (e.g. a box or tray) contains which items, or if a box or tray is packed correctly (i.e. contains all the items that it should contain according to the order that was placed). Boxes and trays are usually transported by motorized conveyor belts inside storage halls, which offers the possibility to install identification points that provide information about the content of boxes as they are passing by. In order to allow for a reliable assignment between items and packaging units (i.e. tags and boxes), a large spacing between boxes would be ideal because antenna radiation patterns are not arbitrarily narrow and the read range of UHF RFID systems is up to several meters. Additionally, wave propagation in practical applications is difficult to predict due to reflections on metal items or the ground floor. However, a large spacing between subsequent boxes cannot be provided for economic reasons. Therefore, different approaches for the localization of tags with respect to packaging units need to be considered. Figure 1.2 depicts a typical conveyor belt application with an RFID identification point consisting of a single antenna. The conveyed boxes A and B contain a specific number of tagged items and all RFID tags are inventoried as the boxes pass the antenna. Due to the range of UHF radio signals and the antenna radiation pattern it is very likely that tags in both boxes will be inventoried at the same time, making it impossible to establish an assignment between tagged items and boxes. This implies that the large range of UHF signals, though desired in some applications can also introduce problems when certain requirements have to be fulfilled.

The effects described above also introduce so called *false positive reads*. Depending on the geometry and antenna orientation, it is possible that objects located in the vicinity of an iden-

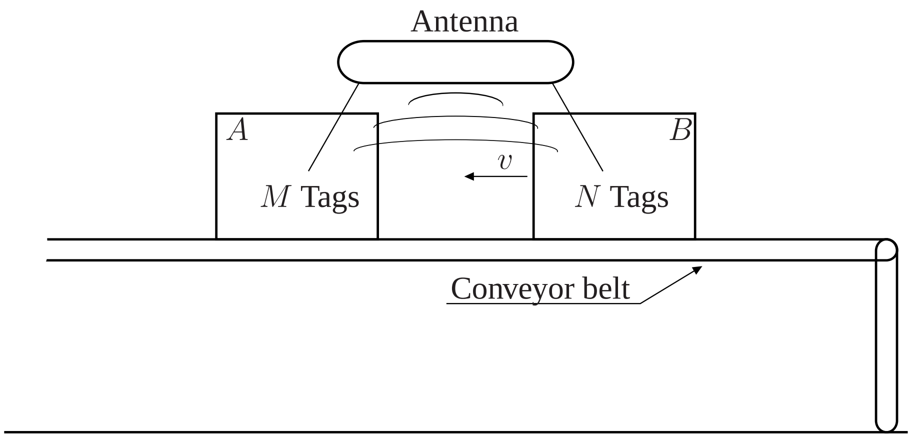


Figure 1.2: Conveyor belt application with RFID. In many practical applications, packaging units (boxes or trays) which contain a certain amount of tagged items are transported across storage halls using motorized conveyors. The boxes A and B are moving with a constant speed v and an antenna is used in order to identify the items in each box. An additional requirement here is to establish an assignment between tags and boxes as they are passing by.

tification point that are not meant to be identified will be inventoried. Imagine for example an RFID portal as shown in figure 1.3. In this scenario, boxes A and B are moving through the portal with a constant speed v and the items in these boxes should be identified by the RFID system. Since space is a scarce resource in storage halls, "reading free zones" around identification points can not be provided and for this reason, box C is located right next to the identification point. In contrast to the items in boxes A and B, the tagged items in box C should not be identified, and by no means be assigned to the content of the other two boxes. Due to the large read range, the RFID tags in box C will most probably be inventoried as well. If no classification among read events is performed, these tags will be reported to the backend system, causing a *false positive read*. Hence, the task of identification points is not just to inventory all RFID tags ever entering the reader field, but also a *classification* between objects that are actually and deliberately passing the identification point and objects which are located nearby. In this context, not the exact location of goods (with respect to a coordinate system) needs to be known - it suffices to have information about whether an item has passed the identification point along the predefined trajectory or not. In other words, the localization of goods is discretized according to the level of detail required by the business process. The discretization needs to be chosen such that identification points like the RFID portal in figure 1.3 can provide reliable information about the current location of goods with respect to the business process.

Usually, items in a business process have a predefined destination and follow a certain tra-

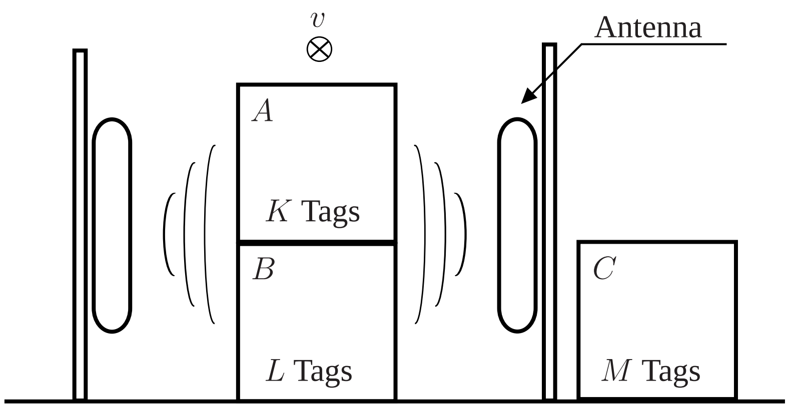


Figure 1.3: An RFID portal. In this application, the RFID systems is used to identify the items in boxes A and B which are moving at a constant speed v through the portal. Due to scarce space, box C is placed right next to the portal, but items in box C should not be identified. Due to the considerable read range of UHF RFID systems, items in box C will be inventoried as well, which leads to a false positive read if the middleware does not perform a classification of read events.

jectory through the process. Moving back to the example with the RFID portal in figure 1.3, this could mean that the tags in box C already have been moving through the portal (which might be an identification point for incoming goods at a warehouse) earlier and need not be identified anymore. Hence, the business process is able to provide information that can be used together with the information from RFID read events. On the downside, business processes are subject to errors, which result from human error or faults in hardware and software. For this reason, also the business process is not a fully reliable source of information, similar to the read events in the RFID system.

Based on these premises, this work presents an approach for robust localization and tracking of single items in an RFID business process. In this context, localization is always discretized with respect to the actual requirements of the process. The fact that RFID read events as well as business layer information are subject to random behaviour motivates a consideration of the different information sources in a probabilistic framework. Since standardization is a vital issue in the field of RFID, the presented system is based on standardized RFID hardware, compliant to the EPCGlobal Class 1 Generation 2 standard.

1.3 Organization of this work

After this introduction that discussed the basic components and tasks of an RFID system, chapter 2 deals with related work done in the field of localization in RFID. Since the idea in this work is to consider the information provided by the underlying business process, also publications about process mining and modeling will be reviewed. Chapter 3 then covers an introduction to the probabilistic model that will be used in this work for modeling RFID read events and business processes - the Hidden Markov Model. In chapter 4, the application of Hidden Markov Models to RFID systems is described and evaluated. Since the developed models for localization in RFID systems have found application in the industry, chapter 5 discusses two use cases from ongoing projects. Finally, chapter 6 concludes this work with a summary and an outlook to possible future fields of research.

2 Related Work

The topic on localization using RFID is subject to active research. Due to the decrease of hardware costs, RFID readers and tags are used in various fields as additional sources of information. The first section in this chapter discusses related work in the field of localization using RFID, where a distinction between reader and tag localization can be made. Since the focus of this work lies on the localization of RFID tags, the literature review also concentrates on this aspect. Another categorization in this context is whether the system uses probabilistic algorithms or directly relies on measurements.

The second section deals with publications in the field of business process modeling and data mining in processes. There is a large number of languages especially designed for modeling business processes, but there is a lack of structure and hierarchy across those languages. Recent approaches are dealing with more mathematical descriptions of business processes that also allow for an evaluation of the model quality.

After a short summary of the related work in the field of localization with RFID and business process modeling, the final section in this chapter discusses the contribution of this work and explains the advantages over previous approaches.

2.1 Localization using RFID

In the last few years, there have been several approaches for different applications in the context of localization using RFID. A main distinction of applications can be made regarding whether a system is concerned with the localization of RFID tags or of RFID readers. The first family of applications uses one or more readers to determine the location of several RFID tags, whereas the second family of applications deals with the localization of RFID readers using reference tags with known positions. An early system for localizing RFID tags is called SpotON [15] that uses several readers to perform a multilateration based on the measured RFID signal strength. The multilateration uses an empirically found relationship between received signal strength and distance, following a quadratic equation. The authors state that this approach suffers from significant limitations which are caused by the fact that the simple empirical model does not consider fluctuations in the received signal strength. The drawbacks regarding accuracy, sampling rate and evaluation time do not allow for a reliable system operation. The authors suggests an extension to the presented system that uses custom tags, which again can be considered as a significant disadvantage, since standardization in this field is vital for large companies that operate supply chains all over the world.

Another system, called LANDMARC (LocAtioN iDentification based on dynaMic Active RFID Calibration) [23] is based on the idea of a dense reader environment that is able to find a tag location by the proximity to a reader. This approach uses active RFID tags for calibration purposes and finding the k - nearest neighbours to a tag of interest. Since the used hardware is not capable of reporting received signal strength information, the transmit power of the RFID readers is swept in a certain range to get a measure that is similar to the RSSI. This measure is obtained for every unknown tag and then compared to the all reference tags by means of a Euclidean distance computation. The k - nearest neighbours are found as the reference tags that have the smallest Euclidean distance to the tag of interest. Since the position of the reference tags is known, this allows for the computation of the tag location up to a certain degree of accuracy, depending on the number and placement of reference tags as well as on the number of used RFID readers. However, this approach has several disadvantages. The first problem is that the system relies on the use of active tags as references, which are much more costly than passive RFID tags. The second issue results from the fact that the RFID hardware used in [23] is not capable of providing information about the received signal strength. The workaround of sweeping the transmit power implies a considerable processing time which is not desirable in practical applications. Last but not least, the use of several readers is another important cost factor in practical applications.

As an extension to the LANDMARC system in [23], the authors of [19] perform a localization in a three - dimensional space and utilize RFID readers capable of reporting the received signal strength. This eliminates the processing time introduced by sweeping the transmission power in the LANDMARC system. Instead of using active RFID tags, the extended LANDMARC system uses passive tags also as reference tags, which greatly reduces costs in practical applications. The achieved accuracy lies in the range of the LANDMARC system, but still requires several RFID readers for the computations.

Another approach that interprets the localization task as a Bayesian problem is presented in [2]. In this work, several RFID readers with rotating antennae are used and the task of localization is formulated as an inverse problem. The presented system uses RFID readers that are not capable of providing RSSI information, for this reason the transmitting power is also swept in a certain range. The posterior probability of detecting a tag at the given rotation angle can be computed as

$$P(\theta \,|\, d) = P(\theta)\frac{P(d \,|\, \theta)}{P(d)}, \tag{2.1}$$

where $\theta = (x, y)$ denotes the location of the tag and d are the acquired data. In equation (2.1), $P(\theta)$ provides a priori information about possible tag locations (e.g. limited by walls or objects) and $P(d \,|\, \theta)$ is the likelihood of receiving the data vector d at a certain tag position and antenna rotation. This likelihood can be acquired by means of an RFID reader model which directly depends on the antenna radiation pattern and transmission power. Despite this system also has the drawbacks of long processing time (due to the necessary power sweep) and the need for several RFID readers, it presents a novel interpretation of the localization problem within a statistical framework.

Following the idea of a probabilistic interpretation, several other papers have been published that formulate the localization task in a Bayesian context. In [14] and [17], approaches that are similar to the idea discussed above are presented and applied to mobile robots. The RFID system attached to a robot is used to localize either the robot or surrounding tags. In both cases, the RFID reader model is learned from obtained read events and is then applied to a Monte-Carlo localization. The RFID reader model learning is done in a recursive manner. Given the known tag positions and reader trajectory, the model stores detected and not detected tags together with the RSSI information. This set of read events is then used to compute the likelihood of an observation (with a given RSSI value) at a given position. To decrease the computational complexity, the considered space is discretized into several cells and a Gaussian distribution of RSSI values in a cell is assumed. The accurate results indicate that it is suitable to use a probabilistic approach for localization tasks with RFID. The big drawback with this approach is however that a vast number of read events is needed in order to be able to compute the reader or tag location using a Monte-Carlo algorithm. Considering a steady identification point in a practical system as discussed in chapter 1, there is no possibility to obtain thousands of read events for tags that are passing an identification point. In addition to that, the reader model in [17] only focuses on the relative orientation and pose between the reader antenna and tag, but does not include environmental considerations, such as object materials. In practical applications, such environmental considerations are necessary in order to cope with several types of tags and tagged items.

The approach for localization discussed in [32] uses so called RFID snapshots to estimate the position of a mobile robot. A snapshot consists of a list of reported tags within a certain time frame together with the number of read events for each tag. The event of tag detection is modeled as a random process, following a Binomial distribution. During a training phase, a large number of snapshots is recorded with a mobile robot that is equipped with two independent antennae. In the evaluation phase, the RFID snapshots are compared to the training data using a Monte-Carlo localization. Since the system is relying on densely tagged environments (with $\gg 100$ tags), the Particle Filter used for the Monte-Carlo localization suffers from problems with small particle weights [32]. This method has the advantage that no RFID reader model needs to be built or learned, however a large number of read events is necessary to achieve accurate results.

An algorithm that is similar to the LANDMARC approach but extends its idea in several ways is presented in [7]. In this work, two readers are used to determine the position of a tag with respect to N reference tags by means of a multilateration. The novelty in this approach is that an adaptive Kalman Filter is used in order to compensate for the noisy RSSI information. Moreover, a probabilistic RFID map is generated that represents the location error for each reader. In contrast to simple empirical models, a large scale path loss model [22] as frequently found in the literature is used to estimate the distance between the target tag and the detected reference tags. Since the goal is to estimate the location of the unknown tags with respect to the reference tags and not to the readers, the localization is performed among elements in the same environment that suffer from the same physical effects and therefore show a certain error correlation. Though the compensation of noisy RSSI information is a valuable and necessary step, the suggested approach also relies on a large number of read events to provide an acceptable accuracy. Still, the suggested method could be of interest when it comes to the localization of static objects when there is enough processing time. However, the results presented in the paper are based on simulations only and have not been reproduced with real measurement data.

Besides the systems for localization in RFID discussed above, there exist several other approaches which are based on Direction of Arrival (DoA) estimation [34]. The idea behind DoA estimation is to derive the direction of an incoming signal by measuring the phase difference between two antennae. Currently, no off-the-shelf RFID reader offers the possibility to acquire this information and hence this technique can not be applied to standard RFID products.

Some of the publications above provide information about the accuracy that can be achieved using the presented approach. These metrics however are difficult to interpret and do not allow for an objective comparison, since the used setups differ significantly regarding the number of used readers or reference tags and quantization intervals. Currently, there is no standardized way of comparing the performance of localization systems using RFID in terms of accuracy. For this reason, the review above focuses on criteria that are motivated from a practical point of view, such as number of readers needed, required processing time and the use of standardized RFID hardware.

2.2 Business Process Modeling and Workflow Mining

Modeling of phenomena of any kind is nothing else than an abstraction of ongoing processes that allows to explain certain observations. This statement can also be applied to business process models, where the goal is to describe a process in a formalized language that can be understood and interpreted by others. It is quite difficult to cover all facets of a business process and there exist a huge number of modeling languages [20]. These languages focus on different modeling aspects and lack a structural hierarchy. Nevertheless, efficient modeling languages are the key to describe large and complex processes and are needed to find optimization potential. Furthermore, modeling languages allow for a simulation of processes, which reduces costs in the design phase of a new process.

The Architecture of Integrated Information Systems (ARIS) framework [27] is a very popular and powerful way to describe processes within a company. Following a structured view, ARIS considers every process as part of a larger enterprise model. The aspect of process modeling within ARIS is covered by Event Driven Process Chains (EPCs) that describe a series of activities with clearly defined inputs and outputs. The term “event driven” arises from the fact that nearly all activities in a company are triggered by some event. Considering the example application in figure 1.3, an identification point for incoming goods is triggered by the delivery of goods to a certain warehouse. On the next layer, RFID read events are triggered by tagged items moving through an RFID portal etc. Since it is possible for a process to require more than one input and produce more than one output, ARIS is able to model decisions in a process and governs these decisions by simple probability measures. Using such measures, it is possible to make statements about how likely the outcome of a certain process chain is. However, the ARIS framework does not provide mathematical instruments to evaluate processes regarding their outcomes and does not allow for more complex mathematical investigations. Besides this drawback, ARIS is a very flexible and popular tool that is able to put every process chain in a larger context, making it suitable to model large and complex business processes.

Another common approach for modeling business processes, especially in the field of software oriented processes are Petri Nets [33] which are capable of describing concurrent processes. Petri Nets are the background for many commonly used modeling languages as the discussion in [20] implies. Petri Nets are good at modeling concurrent activities in processes, however they do not have the possibility to obtain any quantitative information about a process. For example, it is not possible to get any information about how often one specific path of two or more possibilities has been chosen or which steps in a process are the most likely ones. The discussion in chapter 1 indicates that knowledge about the stages and paths through a business process implicitly provide descriptive power that can be used as an additional source of information in RFID driven processes. For this reason, a process model that is capable of making quantitative statements rather than expressing structural dependencies is desirable.

Starting from a Bayesian viewpoint, there have been several approaches to find a mathematically descriptive and flexible way to model business processes. The authors of [24] describe an efficient mapping of Petri Nets onto Hidden Markov Models (HMMs) and shows how to use the HMM in order to evaluate the model quality. In contrast to the Petri Net representation, the HMM implicitly allows for the computation of a quantitative quality measure, as will be shown later in chapter 3. The work [24] is focused on general business processes rather than the special features of processes employing RFID. The novelty is however, that business processes can be described by means of a HMM, which provides great flexibility and descriptive power. On the downside, the idea is restricted to so called Simple Petri Nets which explicitly do not allow for parallelism. This drawback however does not negatively affect the modeling of supply chain business processes, since tagged items are usually not subject to two or more concurrent tasks, but are rather processed in a sequential manner. The authors moreover state that a generalization to concurrent tasks is possible at the expense of computational complexity and traceability.

An alternative approach, coming from the field of workflow mining (also called process mining) uses the capabilities of HMMs to cluster sequential data from event logs [1]. In general, workflow mining is concerned with the extraction of temporal patterns from event logs. Such an event log is a list of records that were made whenever an activity in a certain process was performed. In this case, HMMs can be used to find similar or equal state sequences in the event logs. Due to its flexibility, the approach of representing a business process as Hidden Markov Model seems a lot more suitable to mathematically describe business processes than other modeling languages which focus on the process structure.

2.3 Contribution of this work

Localization in RFID is a popular topic with a lot of recent publications. The two major problems regarding localization are concerned with reader localization on the one, and tag localization on the other hand. In general, there are several methods that directly rely on measurements of the received signal strength by means of a multilateration and methods that use statistical approaches to compensate for noisy measurements.

Due to the randomized nature of UHF signals and tag detection events, a probabilistic approach seems suitable also for localization purposes. The results in the publications reviewed above show that localization of RFID tags can be achieved based on probabilistic considerations. However, all presented methods rely on a large number of RFID read events which can not be provided in most practical applications like the ones demonstrated in chapter 1. Some of the methods discussed above provide an accuracy which may even not be needed for practical purposes. To decide whether a tag is moving through a portal or is located besides it, an exact localization with respect to a coordinate system is not necessary. Besides this fact, the presented methods do not consider the whole RFID system but rather focus on modeling the tag-reader communication. Since the business process in which RFID systems operate offer a great descriptive power about the current state of a tag, the combination of a model for tag - reader communication with a business process model is thought to provide the solutions for the problems outlined in the first chapter.

To be able to consider information from the business process layer, a formal description of ongoing processes is needed. Using this description, it should be possible to perform a quantitative evaluation of every step in the process. The discussion above indicates that a probabilistic view of business processes is suitable and that a mapping from commonly used process models (such as Petri Nets) onto Hidden Markov Models is possible. Therefore, the consideration of the whole RFID system in a probabilistic model is promising to solve localization and tracking tasks.

The novelty in the presented work is the fusion of different sources of information that belong to different layers in an RFID system. As discussed in chapter 1, RFID read events and the underlying business process layer both can not be considered as a fully reliable source of information. The idea is hence to take these sources of information and consider them together in a probabilistic model to perform localization and tracking on a business process level. Due to the fact that probabilistic models are able deal with random fluctuations, the suggested system is expected to be more reliable in terms of localization compared to systems that rely on RFID read events only. Since standardization is a critical issue for RFID processes, the developed system is solely based on EPCGlobal standard compliant RFID readers and standardized passive RFID tags.

3

Hidden Markov Models - HMMs

Hidden Markov Models (HMMs) are a flexible stochastic tool for modeling times series. The first publications about the theory of HMMs range back to the 1960s and since then, they have been applied in several fields such as speech recognition, pattern classification and bio informatics. The first section in this chapter presents a short introduction to the theory and application of HMMs, especially with a focus on modeling and classification of time series and is based on the tutorial by Lawrence R. Rabiner [21]. Whereas section 3.2 briefly covers general applications of Hidden Markov Models. Finally, section 3.3 describes the training of HMMs from a practical point of view and explains how trained HMMs can be used for classification of unknown time series.

3.1 An introduction to HMMs

Phenomena in the physical world in general can be observed by means of signals. Consider for example a vehicle moving at a certain speed, the air pressure as indicator for weather conditions or electrical currents. The state of the vehicle can be monitored by continuously measuring its speed. An important step that is necessary to allow for considerations in a mathematical way is to find a model of the considered physical phenomenon. For this purpose, systems with dedicated inputs and outputs are used. The system mathematically describes the relationship between the inputs and outputs. Such systems usually consider the physical process in a simplified way by neglecting certain effects. Ohm's law for example is a good model for the relationship between the electric current in a conductor and the applied voltage, provided that the voltage remains constant or changes slowly over time. However, it does not model the behaviour correctly when high frequencies are used.

Depending on the type of physical phenomenon, there are different families of signals and systems. One important family deals with phenomena that can not be described in a deterministic manner but require stochastic considerations. In this context, stochastic signals are interpreted as a realization of a random process. A more illustrative interpretation is that a random process is a "black box" that produces a random sequence of outputs. The simplest form of a random process is a discrete time random process, which means that the system produces an output at discrete time instances rather than in a continuous manner. Examples for random processes are Gaussian processes, Poisson processes and Markov processes. Since HMMs can be viewed as an extension to Markov processes, the first part of this introduction briefly covers discrete Markov processes.

3.1.1 Discrete Markov processes

A discrete Markov process is a system with N distinct states $S_1, S_2, \ldots, S_N$, where the state of the system changes at regular, discrete time instants, e.g. $t = 1, 2, \ldots, T$. Such a system is depicted in figure 3.1, with $N = 4$ states and some example state transitions. The current state of the system can be observed (sampled) at the present time instant t, and is denoted as q_t, following the notation in [21]. From each state, the system may change to any other state with a certain probability. These probabilities can be written in a compact form using the transition probability matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & \ldots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \ldots & a_{NN} \end{bmatrix} \tag{3.1}$$

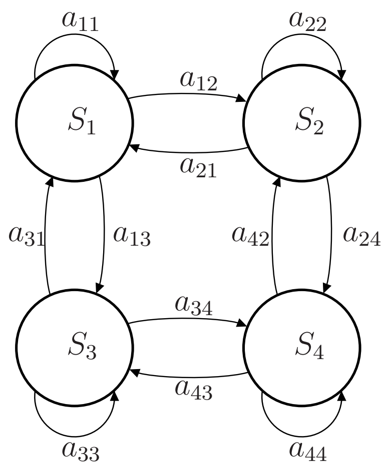


Figure 3.1: Discrete Markov chain with $N = 4$

where a_{ij} denotes the probability that the systems changes its state from S_i to S_j. The matrix $\mathbf{A}$ is a so called row stochastic matrix, which means that its rows sum up to 1

$$\sum_{j=1}^{N} a_{ij} = 1 \quad \forall i \in \{1 \ldots N\} \tag{3.2}$$

and all entries are non-negative:

$$a_{ij} \geq 0 \quad \forall i, j \in \{1 \ldots N\}. \tag{3.3}$$

Markov chains can be classified regarding their transition probabilities. Whereas the most general type of model allows arbitrary state transitions (i.e. $a_{ij} > 0 \forall i, j \in \{1 \ldots N\}$), some models only allow for certain state transitions. A detailed discussion of the different model types is provided in [21].

A special case of Markov processes are systems that obey the first order Markov assumption. This assumption states that the current state only depends on the direct predecessor state, i.e.

$$P(q_t = S_i \,|\, q_{t-1} = S_j, q_{t-2} = S_k, \ldots) = P(q_t = S_i \,|\, q_{t-1} = S_j). \tag{3.4}$$

The Markov assumption provides a simplification in terms of computational complexity, since the current state is conditionally independent of earlier predecessor states. Another simplification is achieved by the assumption that the transition probabilities are independent of time which means that the matrix $\mathbf{A}$ is constant. A commonly discussed example in the literature considers the weather as a three state Markov process with

- State S_1: rain or snow
- State S_2: cloudy
- State S_3: sunny.

Given proper transition probabilities, the model can be used to predict the weather, i.e. compute the probability that the weather is sunny tomorrow, given observations of previous days. In order to be able to compute the probability of a given sequence, the initial state probabilities need to be known, i.e. the probabilities that the system started in a certain state. This can be expressed in terms of

$$\boldsymbol{\pi} = [\pi_1, \ldots, \pi_N], \tag{3.5}$$

where $\pi_i = P(q_1 = S_i)$ denotes the probability that the system started in state S_i. To compute the probability of a certain state sequence $O = S_3 S_3 S_3 S_1 S_1 S_3$, consider the initial probability $P(q_1 = S_3)$ and the transition probabilities according to $P(q_2 = S_3 \,|\, q_1 = S_3)$, $P(q_3 = S_3 \,|\, q_2 = S_3)$ etc:

$$P(O \,|\, (\boldsymbol{\pi}, \mathbf{A})) = \pi_3 \cdot a_{33} \cdot a_{33} \cdot a_{31} \cdot a_{11} \cdot a_{13}. \tag{3.6}$$

Using this framework, it is possible to evaluate the probability of observed sequences and predict the probability of future states.

3.1.2 Hidden Markov Models

The type of model discussed above is appropriate whenever the state of a system can be observed directly, such as it is the case with the weather-example. Whereas the weather can be observed to decide if it is sunny, rainy or cloudy some systems can not be observed directly, but only e.g. through noisy measurements. In this case the observation that is made can be considered to have a probabilistic dependency on the true state of the system. This extension leads to the Hidden Markov Model. The term hidden results from the fact that the true state of the system can not be observed. To extend the weather-example, consider a prisoner that is locked in a cell, deep down in the tower of London. The cell does not have any windows and therefore the prisoner can not observe the weather (i.e. the state of the system) directly. Due to increasing boredom, the prisoner tries to guess about the weather outside. At noon, the warder serves lunch and depending on the weather, he carries an umbrella with him. The fact whether the warder carries an umbrella or not is the only piece of evidence the prisoner has about the weather outside. The observation O_t at time instant t therefore reduces to the event "umbrella" or "no umbrella" and depends on the true system state in a probabilistic manner. Since the weather forecast is not fully reliable, it may happen that the warder carries an umbrella despite it is not raining outside. On the opposite, it may also happen that he left his umbrella at home hoping for good weather, but it is raining despite that. For every state of the system, there is a probability that the warder takes the umbrella with him. Similar to the transition probability matrix $\mathbf{A}$, this can be summarized in the observation matrix $\mathbf{B}$ for discrete HMMs:

$$\mathbf{B} = \begin{bmatrix} b_{11} & \dots & b_{1M} \\ \vdots & \ddots & \vdots \\ b_{N1} & \dots & b_{NM} \end{bmatrix} \tag{3.7}$$

where b_{ij} denotes the probability that the system "emits" the symbol O_j, being in state S_i. Since it is possible that the number of symbols does not equal the number of states in the system (as with the weather example above), $\mathbf{B}$ is a $N \times M$ matrix, where N denotes the number of states in the system and M is the number of symbols. For the weather example, there is only one symbol, for this reason $M = 1$, i.e. $\mathbf{B}$ is a 3×1 matrix. An example of a Hidden Markov model with $N = 3$ states and $M = 3$ observation symbols is shown in figure 3.2.

In order to compute the probability for a system being in a certain state given the current

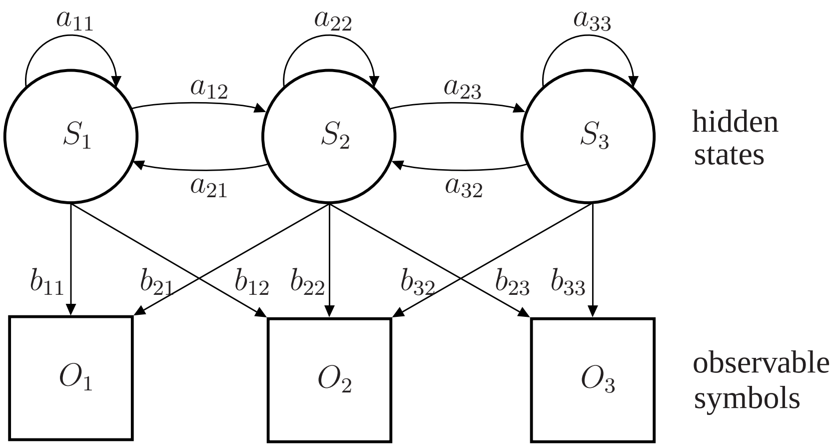


Figure 3.2: Hidden Markov Model with $N = 3$ and $M = 3$

observation, Bayes' rule can be used:

$$P(q_t = S_j \mid v_t = O_k) = \frac{P(v_t = O_k \mid q_t = S_j)P(q_t = S_j)}{P(v_t = O_k)}. \tag{3.8}$$

More general, the transition and emission probabilities allow for the computation of the probability for any given sequence of observations $O = O_1 O_2 \ldots O_T$.

To summarize, a Hidden Markov Model denoted by λ is characterized by the following quantities:

- The number of possible system states N. Despite the fact that the system states are not observable, they can still have a physical meaning, as with the weather example.
- The state transition probabilities, summarized in the transition matrix $\mathbf{A}$. In the most general form, every state can reach any other state within a single step, which means that $a_{ij} > 0 \quad \forall i, j \in \{1 \ldots N\}$.
- The number of distinct symbols M which can be emitted by the system. As discussed earlier, it is possible that the number of symbols is smaller (or even larger) than the number of distinct system states.
- The symbol observation probabilities, summarized in the observation matrix $\mathbf{B}$.
- The initial state distribution $\boldsymbol{\pi}$.

Since the two parameters M and N are implicitly given by the observation and transition probability matrices $\mathbf{A}$ and $\mathbf{B}$ respectively, the HMM can be represented as a triple

$$\lambda = (\boldsymbol{\pi}, \mathbf{A}, \mathbf{B}). \tag{3.9}$$

The literature discusses three basic problems with HMMs, for which solutions exist as well known and widely used algorithms:

Problem 1: Given an observation sequence $O = O_1 O_2 \ldots O_T$, what is the probability $P(O \mid \lambda)$, i.e. the probability that this sequence was generated by the model λ? The brute force approach for solving this problem is to evaluate the probabilities of all possible state sequences Q of length T

$$P(O \mid \lambda) = \sum_{\text{all } Q} P(O \mid Q, \lambda) P(Q \mid \lambda). \tag{3.10}$$

The expression above assumes statistical independence of the observations, and can be expanded to

$$P(O \mid \lambda) = \sum_{\text{all } Q} \pi_{q_1} \prod_{t=1}^{T} a_{q_{t-1} q_t} b_{q_t}(O_t). \tag{3.11}$$

Unfortunately, this approach requires a computationally unfeasible number of $2T \cdot N^{T+1}$ operations, which is why more efficient algorithms, as the the Forward-Backward procedure [3, 5] have been developed. Since this introduction is meant to give a brief overview over Hidden Markov Models from a practical point of view, the reader is referred to [21] for a more detailed discussion of the Forward-Backward procedure.

The question about the probability of a given observation sequence can be extended to a classification task. Consider a set of K Hidden Markov Models $\lambda_1, \lambda_2, \ldots \lambda_K$ and an arbitrary sequence O. In order to find out which class the sequence most likely belongs to, the probabilities $P(O \,|\, \lambda_i), \quad i = 1 \ldots K$ need to be computed. The model with the highest likelihood best describes the observed sequence and hence represents the class of the signal.

Problem 2: The second problem deals with finding the "best" state sequence $Q = q_1 q_2 \ldots q_T$ of a system given an observation sequence O. In this context, "best" means optimal according to some criterion, such as the "single best state" criterion. This criterion chooses the states that are individually most likely, i.e. maximize $P(Q, O \,|\, \lambda)$. The question about the actual system states given an observation sequence is encountered quite frequently in communications engineering, where the goal is to decode the actually sent symbols from the received observations. The algorithm used for decoding an observation sequence is called Viterbi algorithm [31], which is similar to the Forward-Backward procedure, except for a maximization step over previous states.

Problem 3: The last problem for HMMs is the estimation of the model parameters

$$\lambda = (\boldsymbol{\pi}, \mathbf{A}, \mathbf{B}) \tag{3.12}$$

from a set of given observation sequences. This procedure is referred to as training of a Hidden Markov Model using labeled data represented by means of observation sequences. In general, there is no closed solution to this problem. There are, however, several iterative algorithms like the Baum-Welch algorithm [4] which can be used to find local maxima of the probability $P(O \,|\, \lambda)$. An intuitive method, closely related to the Baum-Welch algorithm is presented in [21] and is based on counting the occurrence of events:

$$\hat{\pi}_i = \text{expected frequency of being in state } S_i \text{ at time } t = 1 \tag{3.13}$$

$$\hat{a}_{ij} = \frac{\text{expected number of transitions from state } S_i \text{ to } S_j}{\text{expected number of transitions from state } S_i} \tag{3.14}$$

$$\hat{b}_{ij} = \frac{\text{expected number of times in state } S_i \text{ observing symbol } v_j}{\text{expected number of times in state } S_i}. \tag{3.15}$$

A detailed description of the computations involved in the parameter estimation is given in [21]. All methods used for estimating the model parameters of a HMM have in common that the quality of the estimation relies heavily on the choice of initial parameters, since only local maxima of the probability can be found. For this reason, a good initial choice for $\boldsymbol{\pi}$, $\mathbf{A}$ and $\mathbf{B}$ is vital for the estimation of parameters and the descriptive power of the resulting HMM. Besides the choice of model type, number of system states N and emission symbols M, the parameter estimation is the key step in the design of HMMs. Whereas the former three can be managed most often in a straight forward manner, the latter needs special attention. One approach is to initialize the parameters with uniformly distributed values, another approach is to use knowledge about the underlying physical process if available. Another problem that comes into play when training a HMM is the fact that the training data will always be finite and therefore may be insufficient to cover all aspects of the underlying process. A possible solution to this problem of insufficient training data is to increase the training data set which is sometimes not possible or impractical.

So far, the discussion about Hidden Markov Models was about models with discrete observation symbols and sequences. For modeling certain processes and signals in the real world it is sometimes necessary to consider a more general type of HMM, where observation symbols are emitted according to continuous probability density functions (PDFs) rather than discrete probabilities. A common approach to deal with such models is to consider the PDFs as a mixture of Gaussian densities. For such mixtures, algorithms for the estimation of the mixture component weights and the parameters of the underlying Gaussian densities exist [12]. In contrast to discrete HMMs, the observation probabilities are no longer collected in a matrix $\mathbf{B}$ of discrete probabilities, but rather as a set of PDFs.

3.2 Application of HMMs

After giving a brief introduction about Markov chains and Hidden Markov Models, this section provides a general overview over possible applications of HMMs. Moreover, the application of HMMs to RFID systems and business process modeling is motivated.

HMMs can be used for classification and recognition tasks such as handwriting recognition, gesture or motion recognition from video signals or speech recognition. All recognition systems have in common that they perform a feature extraction from a given signal to derive an abstract representation by means of a time series. The extracted features are then processed by a set of HMMs and possibly some post processor which is concerned with certain characteristics of the considered time series, such as grammar for speech.

For example, HMMs are widely used in speech recognition in order to find and recognize basic units of speech, such as phonemes, syllables or even whole words. For this purpose, a feature extraction from the given speech signal is performed which most often includes a spectral representation. This representation is then analyzed by a set of previously trained HMMs. Basically, the HMMs perform a classification task and give answer to Problem 1 (i.e. how likely is it that the considered model produced the current set of observations) as discussed in the previous section.

The attractiveness of HMMs mainly results from the fact that they are not concerned with the actual signal representation (e.g. voltage over time), but rather rely on extracted features. Whenever it is possible to extract common features from a given signal, HMMs can be used for recognition and classification tasks if certain assumptions (such as the first order Markov assumption) are fulfilled up to a certain degree.

In an RFID system, read events are reported in a standardized way as shown in equation 1.2. Such a series of read events can be interpreted as a signal that is subject to random fluctuations as will be described later in chapter 4. Together with information from other sensing devices, an abstract representation of these signals can be found by means of a feature extraction. This representation can be mapped to a discrete time series of observation sequences which can be modeled by means of a HMM. Using the mathematical framework of HMMs, it is also possible to classify unknown feature sets. This makes HMMs suitable for the application to RFID systems, since identification points need to perform a classification of read events in order to report reliable information to the backend system.

Similarly, HMMs can also be used to model business processes. In a process, the identification of an item at a certain stage yields a discrete observation symbol. For example, if an item with identifier I was identified at a “goods-in” identification point, the middleware will report to that backend system that the item with identifier I was read at identification point k. An item moving through the different stages of the process hence implicitly provides an observation sequence of the last stages in the process

$$\mathcal{O} = k\ l\ m\ n, \tag{3.16}$$

where $k, \ldots, n$ denote the the corresponding identification points. This observation sequence can be modeled by a HMM in a straight forward way, since no feature extraction and mapping to observation symbols is necessary. Conversely, HMMs with appropriate parameters can be used to determine the probability of certain trajectories through the business process, which in turn can be used to support the localization of items. This discussion implies that HMMs are well suited for modeling RFID business processes. Moreover, they provide a flexible and elegant mathematical framework to numerically represent and evaluate information from the business process layer.

3.3 HMM training and classification in practical applications

This section describes from a general point of view how HMMs can be used for classification of signals obtained from sensing devices in an RFID system. Whereas the structure and characteristic of the HMM depends on the setup of the considered identification point, training and classification follow a general procedure.

In general, discrete HMMs require discrete observation sequences for training and classification. For this reason, a mapping from the signals obtained from sensing devices in the RFID system to discrete valued observation sequences needs to be found. This abstract representation is derived by extracting appropriate features from the signals. To account for the fact that discrete time series are needed, the available signals need to be sampled in an appropriate manner. Every sample yields a d dimensional feature vector

$$\mathbf{f} = [f_1 \quad f_2 \ldots f_d]^T \tag{3.17}$$

which means that the features span an d dimensional vector space $\mathbb{R}^d$. Consequently, sampling a signal K times, yields a feature set of K feature vectors

$$\mathbf{F} = [\mathbf{f}_1 \quad \mathbf{f}_2 \quad \dots \quad \mathbf{f}_K]. \tag{3.18}$$

In order to model the temporal evolution of the feature vector by means of a discrete HMM, it is necessary to assign every feature vector to a specific observation symbol $O_1, O_2, \dots O_M$. This is a common problem in statistical data analysis and there exist a variety of algorithms for this task. One standard approach is to use the K-Means algorithm as described in [28]. The K-Means algorithm tries to find M sets of clusters in a d dimensional feature space by minimizing the sum of squares within each cluster $\mathbf{O}_i$

$$\arg\min_{\mathbf{O}} \sum_{i=1}^{k} \sum_{\mathbf{f}_j \in \mathbf{O}_i} ||\mathbf{f_j} - \boldsymbol{\mu}_i||^2, \tag{3.19}$$

where $\boldsymbol{\mu}_i$ denotes the mean (or centroid) of cluster $\mathbf{O}_i$. Applying the K-Means algorithm to a set of training data

$$\tilde{\mathbf{F}} = \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \\ \vdots \\ \mathbf{F}_L \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{11} & \mathbf{f}_{12} & \dots & \mathbf{f}_{1K} \\ \mathbf{f}_{21} & \mathbf{f}_{22} & \dots & \mathbf{f}_{2K} \\ \vdots & & & \vdots \\ \mathbf{f}_{L1} & \mathbf{f}_{L2} & \dots & \mathbf{f}_{LK} \end{bmatrix} \tag{3.20}$$

of size L yields M clusters that can be interpreted as the observation symbols of the HMM as discussed in chapter 3. This means that the number of meaningful clusters equals the number of observation symbols M of the HMM. Provided these clusters, each feature set can directly be mapped to a series of observation symbols that can be used to train the considered HMM, for example with the Baum-Welch algorithm.

As mentioned earlier in this chapter, learning algorithms like the Baum-Welch algorithm rely heavily on a good initial guess of the parameters, since they are only able to find local maxima of the probability function $P(O \mid \boldsymbol{\pi}, \mathbf{A}, \mathbf{B})$ in a recursive manner. One approach is to establish an initial guess from a physical interpretation of the modeled process. This, however is a time consuming iterative process that requires a lot of expert knowledge and intuition.

Another possibility is to randomly initialize $\boldsymbol{\pi}$, $\mathbf{A}$ and $\mathbf{B}$, estimate the values from the given data set and evaluate the likelihood of the found model. This procedure is repeated until the model with the best parameters (yielding the highest likelihood) is found, or until the number of iter-

Algorithm 1 HMM Learning in pseudo-code

for $i = 1$ to MAXITERATIONS **do**
  initialize $\mathbf{A}_i$ with random numbers
  initialize $\mathbf{B}_i$ with random numbers
  initialize $\boldsymbol{\pi}_i$ with random numbers
  estimate $\hat{\boldsymbol{\pi}}$, $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ and evaluate $P(O \mid \lambda_i = (\hat{\boldsymbol{\pi}}, \hat{\mathbf{A}}, \hat{\mathbf{B}}))$ using the Baum-Welch algorithm
end for
$\lambda = \arg\max_{\lambda} P(O \mid \lambda_i = (\hat{\boldsymbol{\pi}}, \hat{\mathbf{A}}, \hat{\mathbf{B}}))$

ations exceeds a certain limit. The corresponding pseudo-code is shown in algorithm 1. Given the structure of a HMM (i.e. the number of states N and the number of observation symbols M), the HMM learned this way best describes the observation data set O. A third possibility to obtain the parameters of the HMM is to combine the two approaches described above. For example, it is sometimes possible to obtain an initial guess for the prior state probability vector $\boldsymbol{\pi}$ from a physical interpretation whereas the transition probabilities can be initialized randomly.

Applying the training procedure described above to different classes of feature sets leads to a set of HMMs which can be used to classify unknown feature sets. For this purpose, the same kind of sampling and feature extraction needs to be applied to every given unknown signal set. The next step is then to assign every sampled feature vector to an observation symbol. As stated above, the observation symbols are equivalent to the clusters derived by the K-Means algorithm. Each cluster $\mathbf{O}_i$ can be described by its mean vector $\boldsymbol{\mu}_i$ and covariance matrix $\boldsymbol{\Sigma}_i$ that can be estimated from the training feature vectors belonging to a particular cluster. In order to assign an unknown feature vector to one of the M clusters, the Mahalanobis distance

$$d_i = \sqrt{(\mathbf{f}_j - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{f}_j - \boldsymbol{\mu}_i)} \quad i = 1 \ldots M,\ j = 1 \ldots K \tag{3.21}$$

is computed for each feature vector and cluster respectively. The feature vector is then assigned to the cluster that yields the smallest Mahalanobis distance. This way, a series of K feature vectors is mapped to an observation sequence O which can be evaluated by means of the trained HMMs. For this purpose, the N previously trained HMMs are used to evaluate the observation sequence O by calculating the probabilities $P_i = P(O \mid \lambda_i)$. The classifier chooses the HMM that yields the highest probability and outputs the corresponding class i. A general classifier for N different signal classes using HMMs is shown in figure 3.3.

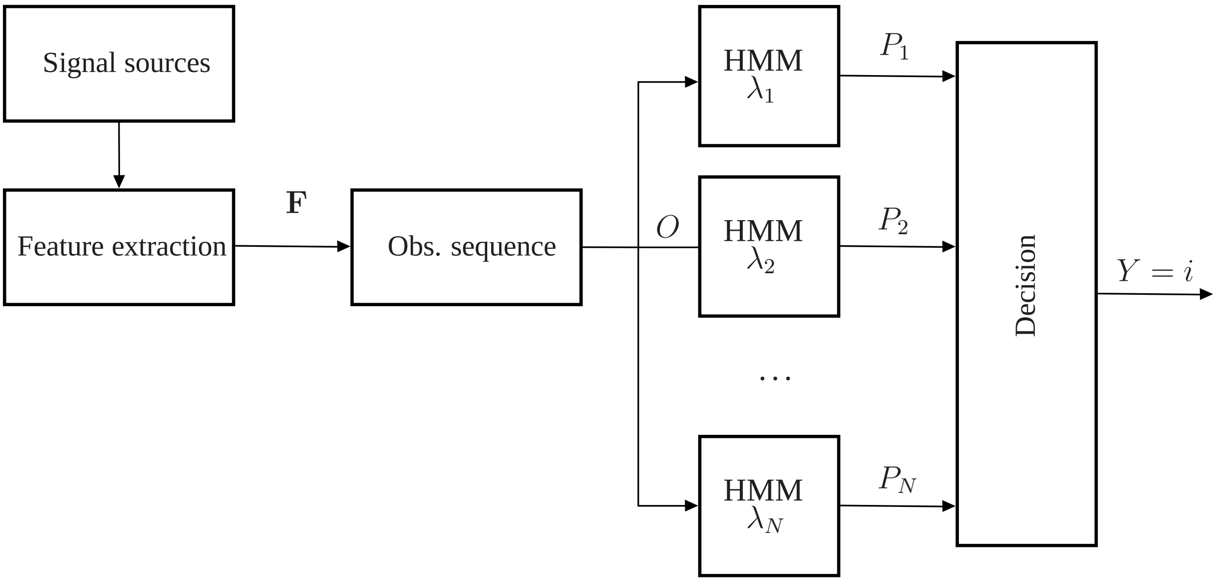


Figure 3.3: A general classifier for a set of N different classes. Given a set of signal sources, a feature extraction is performed that yields a discrete time feature set **F**. This feature set is mapped to an observation sequence O that can be classified by evaluating the probabilities $P_1, \ldots, P_N$. The classifier outputs the label of the class with the highest probability of producing the particular sequence.

Using the mechanisms for training and classification described above allows to consider RFID read events and sensor signals in a probabilistic manner that can account for random fluctuations. Due to the mathematical framework provided by HMMs, also an efficient classification of signals is possible.

4 Probabilistic System Model

After the general introduction to Hidden Markov Models and their applications to RFID system and business process modeling, this chapter presents details about the suggested system for localization.

Section 4.1 discusses the different types of information sources in an RFID system and shows how these sources can be considered in an appropriate way for localization and tracking tasks. Based on this discussion, the general localization system architecture is presented.

Section 4.2 deals with the details of modeling RFID read events and sensor signals by means of HMMs based on an exemplary identification point. For this purpose, the extraction of features from the available signals and the mapping to discrete observation sequences is described. In the next step, details about the training of HMMs and a model evaluation are presented. This includes an evaluation of the developed classification system based on real RFID read events.

Finally, section 4.3 presents an approach to consider business processes employing RFID in a probabilistic framework by means of a HMM. For this reason, a formal set of rules that map a given business process to a HMM is discussed. The obtained process model is then used together with the classification of feature sets for localization tasks. The section concludes with an evaluation of the overall system based on simulation results.

4.1 System components for localization and tracking

Considering the general layer model of an RFID system as shown in figure 1.1, there are different components in an RFID system that can be considered as sources of information. On the lower layers, there are RFID readers and sensors that provide information about RFID tagged items and objects around an identification point. As stated in chapter 1, this information is subject to random fluctuations for various reasons. Besides the sensing devices, also the backend system can be interpreted as an information source in terms of business process information. The backend system is able to store information about every tagged item, including the recent stages in the process or the physical condition of items and tags. Similar to RFID read events, this business process information is subject to fluctuations and errors.

The idea in this work is hence to combine the two different types of information sources – both considered as not fully reliable – and use them together for localization and tracking of items in the business process. For this purpose, a likelihood Λ_I for an item passing an identification point is computed based on the information obtained by the sensing devices in the RFID system. In addition to that, the localization system considers information from the business process, denoted by $\mathcal{H}$. The result of the localization is an estimate of the current position $\hat{\mathbf{x}}$ of an item in the process and can be used to update the business layer information. The structure of the suggested system is shown in figure 4.1.

The mechanisms for localization and tracking are well suited for the implementation in the middleware of an RFID system. As shown in the general system overview in figure 1.1, the middleware is the central layer of every RFID system with interfaces to sensing devices and the backend system. For this reason, the necessary information flow for localization as outlined in figure 4.1 does not require extensive changes to the architecture of an RFID system.

Starting from this general system architecture, the following two sections describe the probabilistic modeling of identification points and business processes using HMMs. Besides an in-depth discussion of HMM training and the classification of feature sets, this also includes an evaluation of the derived models using real RFID read events and an exemplary business process.

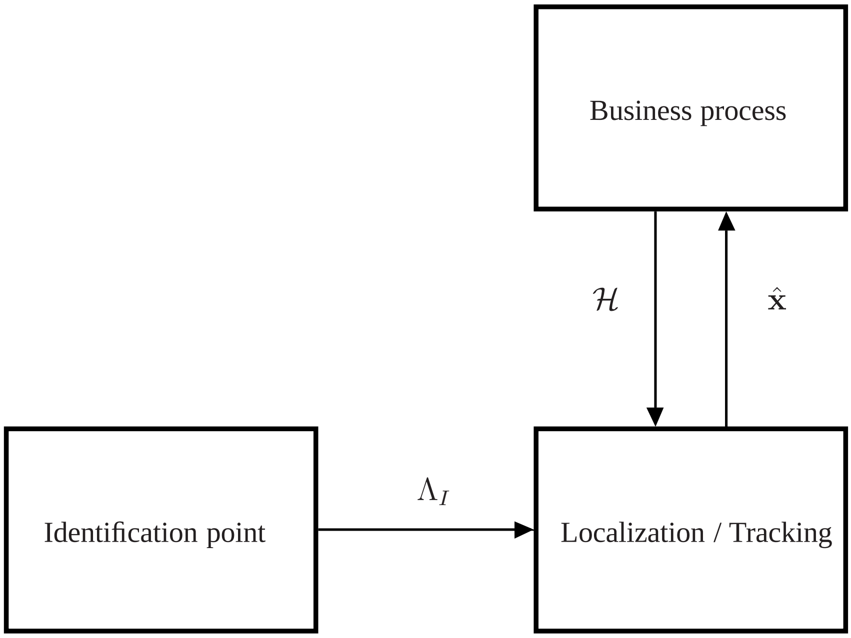


Figure 4.1: System components for localization and tracking. An identification point in an RFID systems provides a certain likelihood Λ_I of items passing the identification point, based on RFID read events and information from other sensing devices. Together with the information $\mathcal{H}$ provided by the business process, this likelihood can be used to estimate the current location $\hat{\mathbf{x}}$ of an item in the process.

4.2 Identification point modeling

The problem statement in chapter 1 briefly outlines the requirements to identification points using RFID. This section presents a probabilistic model for identification points based on an exemplary conveyor belt application.

4.2.1 Physical setup

The physical setup of the example identification point is shown in figure 4.2. Boxes with a

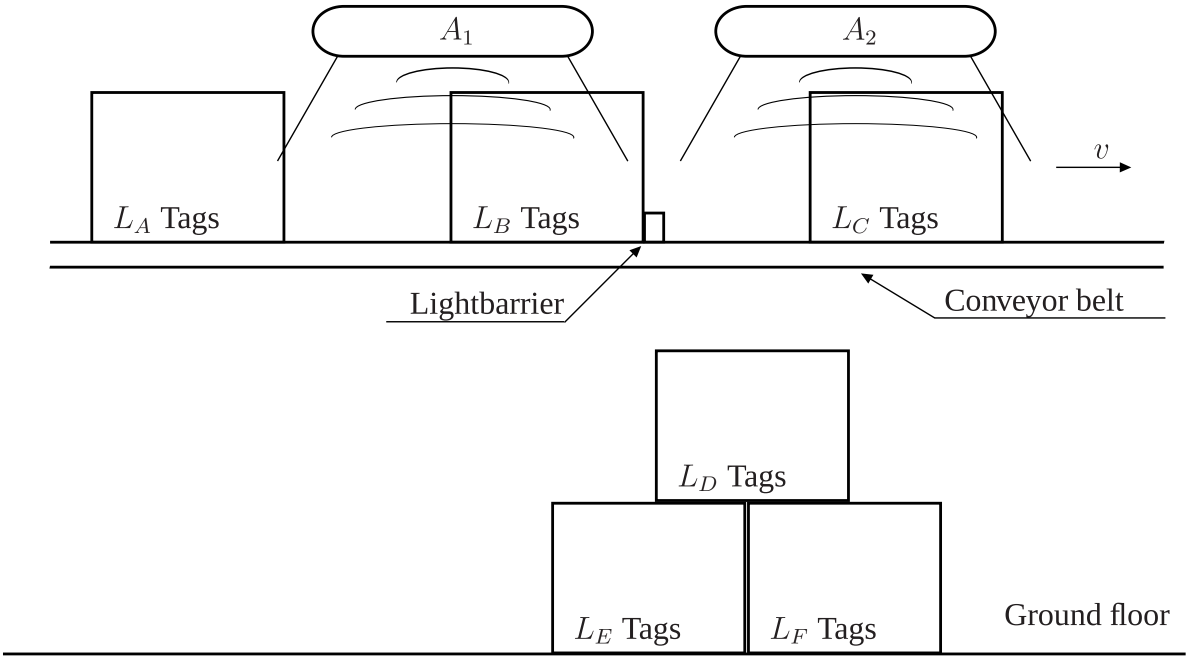


Figure 4.2: RFID conveyor belt application with two antennae. Boxes that contain a specific number L_i of tagged items are transported by a conveyor with constant speed v. In addition to that, there are boxes located besides the conveyor, for example on the ground floor. The task is to identify the items on the conveyor and assign them to the correct box. For this purpose, two antennae and a light barrier as sensing devices are used.

specific number of items are transported on a conveyor at constant speed v. The suggested approach uses an RFID reader with two antennae and a light barrier as sensing devices. The light barrier is meant to give information about when a box is passing the identification point by means of a digital signal. On the one hand, this signal is used to trigger the processing, on the other hand it also provides information about when a box enters the area between the two antennae. The two antennae are connected to an RFID reader which in turn reports read events to a host PC over an Ethernet connection. The conveyor belt is constantly moving boxes with tagged items into the RF - field of the two antennae. Moreover, there are boxes located near the identification point which contain tags that will also be inventoried.

4.2.2 Signals from sensing devices

After a description of the physical setup of the identification point, it is helpful to have a look the signals provided by the RFID reader and the light barrier. Given that a box with a tagged item is passing the identification point at constant speed, it is to expect that the tag will be first seen on antenna A_1 as soon as it enters the field that is determined by the antenna radiation pattern. As the conveyor belt moves on, there will be an overlapping section where the tag is inventoried by both antennae in an alternating manner. During this period, the light barrier will indicate that there is currently a box present. Finally, there will be a stage where the tag is only seen on antenna A_2. Figure 4.3 shows the resulting signals obtained from the read events on the two antennae and the light barrier. Although the read events obtained by the RFID reader

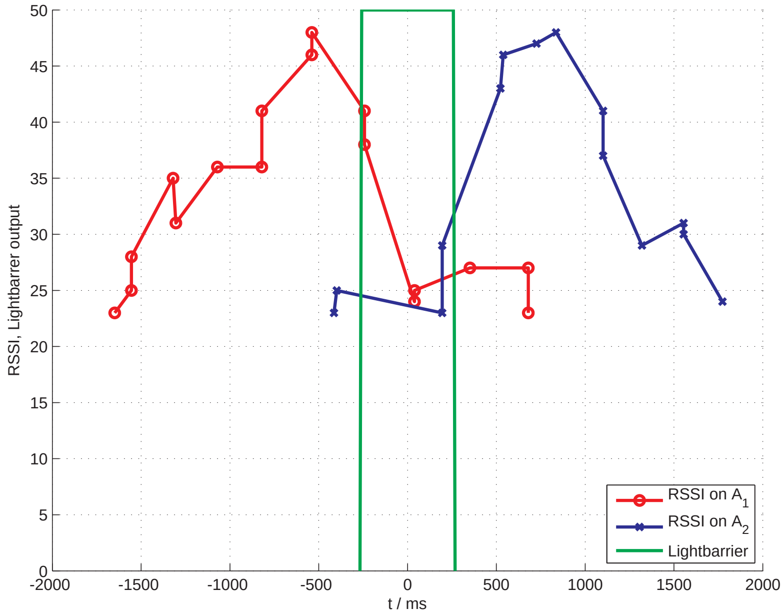


Figure 4.3: RSSI pattern and light barrier signal for a single tag passing two antennae at constant speed.

provide a pure discrete time signal, the plot in figure 4.3 connects the sampled points for a more convenient view. In addition to that, the the output of the light barrier signal is scaled to the magnitude of the RSSI values. The x-axis in figure 4.3 displays a time axis that is shifted to the time instant at which the box was exactly at the center of the identification point. Hence, read events that occurred prior to this time instant show a negative timestamp.

There are three interesting facts about the RSSI patterns from the two antennae in figure 4.3. The first fact is that the sampling intervals between two subsequent read events are not constant, i.e. read events are not subject to uniform sampling with respect to time. The reasons for the non uniform sampling are twofold: On the one hand, there are physical effects on the channel which cause that the intervals between two inventory rounds are varying, on the other hand there are also reasons located at the protocol layer as read events are reported to the host PC. The second interesting issue is that the amplitude of the signals shows fluctuations and discontinuities. The orientation and angle between tag and reader antenna, tag manufacturing issues and multipath propagation across the physical channel are possible causes for these fluctuations. The third noteworthy fact about the RSSI pattern is the limited number of read events for each tag passing the identification point. Depending on the speed of the tagged item, the antenna radiation pattern, the transmitted power and the total number of tags in the field, it is possible to obtain RSSI patterns with several ten read events in most practical applications. The number of read events is an implicit measure of “how well” a tag was seen at a specific identification point. The more read events, the more information can be obtained from the corresponding RSSI pattern.

4.2.3 HMM Training

Provided these signals, a feature extraction that enables a mapping to discrete observation sequences as discussed in section 3.3 can be applied. To overcome the issue of non uniform sampling intervals, a windowing technique is suggested which splits the whole RSSI pattern up into K windows of equal length T_W. Considering the symmetry of the physical setup and the RSSI pattern, it is suitable to perform a symmetric windowing across the RSSI pattern, indicating that K is an odd number. Figure 4.4 shows the RSSI pattern discussed above with $K = 5$ windows of length $T_W = 900\,\mathrm{ms}$. The total frame length is then

$$T_F = K \cdot T_W = 4500\,\mathrm{ms}. \tag{4.1}$$

The values for K and T_W were chosen according to the following considerations: The first parameter to be chosen is the total frame length which was adapted to the length of the RSSI pattern. It is chosen such that the frame contains all read events, i.e. the RSSI pattern is not

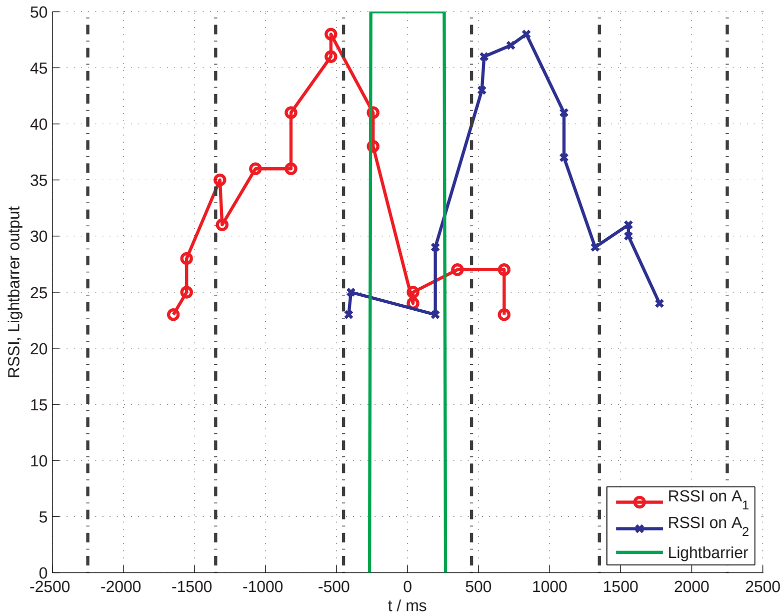


Figure 4.4: Windowed feature set for a single tag across two antennae.

truncated. Once the frame length is chosen, the number of windows K is the next parameter to be determined. Given a fixed frame length, a compromise needs to be found between the number of read events per window and the frequency at the signals are sampled. If the window size is small, this yields a higher sampling frequency but also implies that the number of read events within each window will be rather low. This is a disadvantage for the feature extraction, since the variance of the chosen features increases.

For the representation in a discrete state sequence, robust features need to be extracted from the available signals within each window. These features include information about the current light barrier state (i.e. whether a box is present or not), but also statistics regarding the RSSI pattern. Possible features regarding the RSSI pattern are

- Number of read events in a window
- RSSI statistics (median, mean value, variance . . .)
- RSSI signal energy in a window

$$s = \langle r[n], r[n] \rangle = \sum_{n=0}^{R} |r[n]|^2 \tag{4.2}$$

for real valued signals in a linear scale. The drawback of this measure is that the unit of RSSI values differs from manufacturer to manufacturer and is sometimes provided in logarithmic scale, for which the signal energy can not be computed this way.

- Slope of the RSSI pattern or higher order derivatives

Considering the random fluctuations of RSSI patterns, mathematical functions like first or higher order derivatives do not seem to be an appropriate measure, though the slope of the RSSI could provide information about whether a tag is moving towards an antenna or not. Statistical moments like mean or variance would compensate for the fluctuations of the RSSI pattern, however these moments do not cope well with statistical outliers. Hence, the median value seems to be an appropriate feature. The same argument applies to the signal energy within a window: Due to its integrating nature, outliers do not considerably affect the signal energy and hence could be used as a robust feature of the RSSI pattern. The number of read events inside a window is another possibility to obtain a statistic of the reported read events, especially in cases where no RSSI information is available. Considering that state of the art RFID readers do provide an RSSI value, other features such the median value are expected to be more descriptive. Provided these facts, the median RSSI value within each window is chosen as feature for the RSSI pattern. The resulting feature vector is hence

$$\mathbf{f} = [\tilde{r}_{A_1} \quad \tilde{r}_{A_2} \quad u]^T \tag{4.3}$$

where the tilde operator stands for the median and u denotes the output signal of the light barrier in an appropriate scale. In this case, the chosen features span vector space with $d = 3$ dimensions which also allows for a convenient graphical interpretation. For the mapping of feature sets to discrete observation symbols, signals from 165 tags in several boxes passing the identification point have been recorded. After extracting the features above, the K-Means algorithm can be applied to cluster the resulting feature vectors in a set of M groups. Figure 4.5 shows the result of applying the K-Means algorithm to the feature sets, with a number of clusters respectively observation symbols of $M = 4$. The four symbols correspond to four physically meaningful states of a tag within a window:

- Tag is seen on antenna A_1, corresponding to the observation symbol O_1
- Tag is seen on antenna A_1 and on antenna A_2 and the light barrier signal indicates that a box is between the antennae, corresponding to the observation symbol O_2
- Tag is seen on antenna A_2, corresponding to the observation symbol O_3
- Tag is not inventoried at all, corresponding to observation symbol O_4.

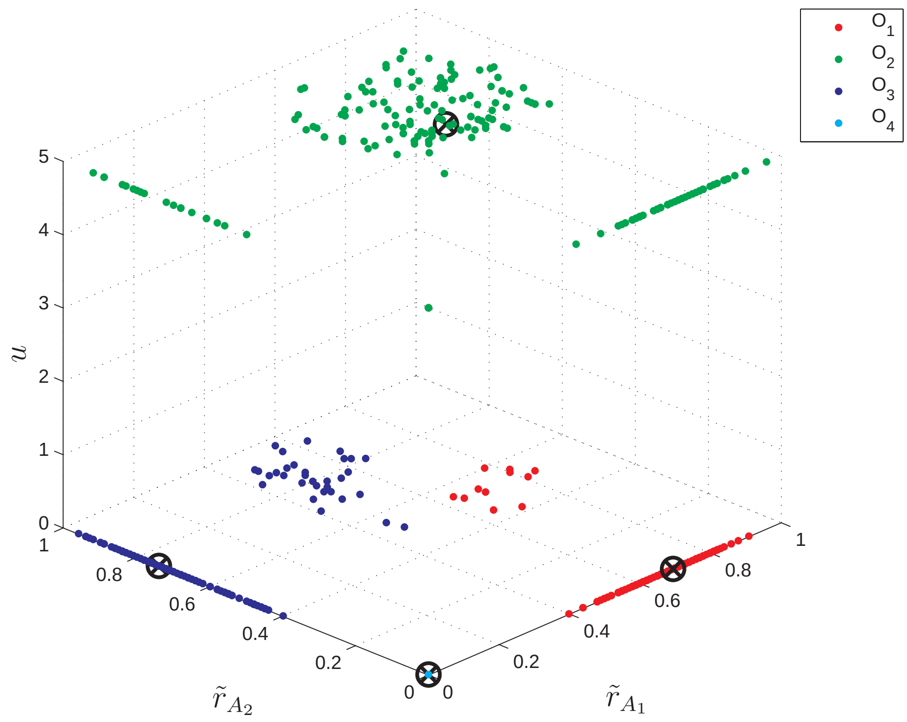


Figure 4.5: Feature space with clustered data points and cluster centroids. The plot shows feature vectors from 165 tags passing the identification point. The data points in the $d = 3$ dimensional space are grouped into $M = 4$ clusters, corresponding to the observation symbols of a HMM.

In figure 4.5, the median values for the RSSI pattern on antenna A_1 respectively A_2 have been normalized to the interval $[0 \dots 1]$, whereas the light barrier signal is weighted with a factor of $k_S = 5$. Since the light barrier provides a deterministic signal, this information is weighted stronger than the fluctuating information obtained from the RSSI pattern. Figure 4.5 does not provide information about the temporal evolution of feature sets, since it shows the whole set of obtained feature vectors, regardless of the actual sequence. Nevertheless it provides an intuitive interpretation of the extracted features that correspond to different states of a tag passing the identification point. In addition to the cluster centroids, the K-Means algorithm also provides a mapping between every feature vector $\mathbf{f}_i$ and the containing cluster, which can be used to determine the discrete time observation sequence $O_1, O_2, \dots O_K$. For the first four tags from the data set in figure 4.5, the corresponding observation sequences are

$$O^{(I_1)} = O_4\, O_4\, O_2\, O_3\, O_4, \tag{4.4}$$

$$O^{(I_2)} = O_1\, O_1\, O_2\, O_3\, O_4, \tag{4.5}$$

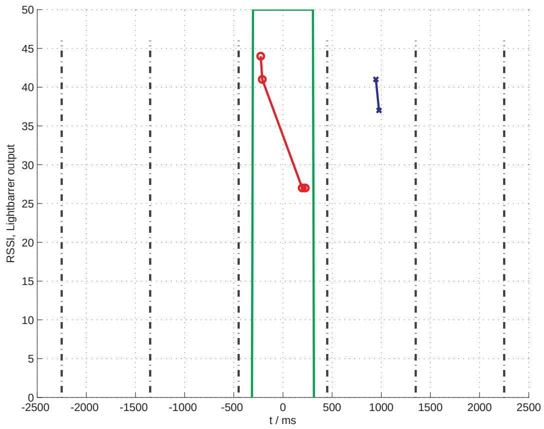
$$O^{(I_3)} = O_1\, O_1\, O_2\, O_3\, O_4 \tag{4.6}$$

and

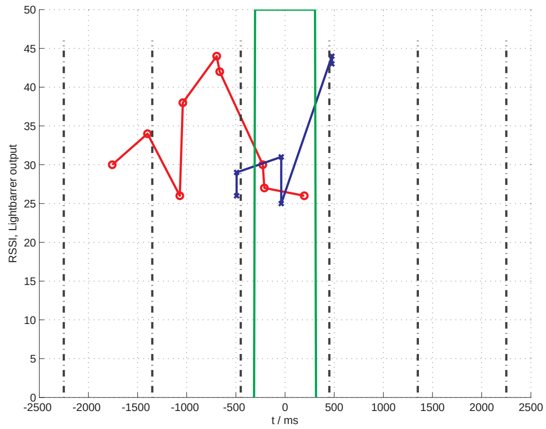
$$O^{(I_4)} = O_1\, O_1\, O_2\, O_3\, O_3. \tag{4.7}$$

This shows that four tags, though located in the same box as they pass the identification point take different trajectories through the feature space. This directly results from the random nature of the feature sets, which are illustrated in figure 4.6 for the four tags mentioned above. Whereas tags $I_2 \dots I_4$ provide a sufficient number of read events, the tag with identifier I_1 suffers from a relatively low read count, resulting in the observation sequence $O^{(I_1)}$. This illustrates the random fluctuations of RSSI patterns, especially when there is a large number of tags in the field of the reader. Despite the tags $I_1 \dots I_4$ are moving in the same box across the conveyor belt, the feature sets and resulting observation sequences are quite different. This is an issue for which a HMM as a versatile stochastic model for time series can account for.

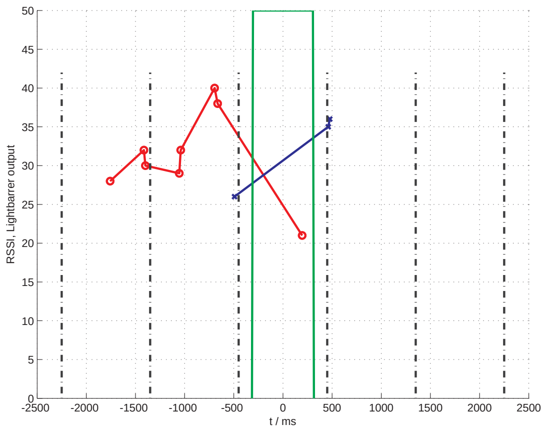
Considering that the number of observation symbols for the resulting HMM is derived from the number of clusters in the feature space, the next step is to define the number of hidden states. This can be derived in a straight forward manner from the previous discussion about the physical states of a tag and hence the number of states N is equal to the number of observation



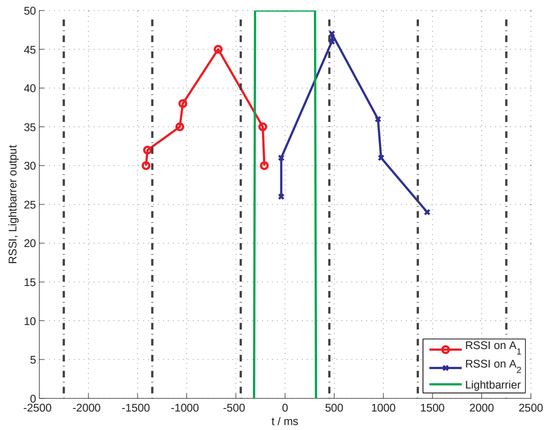
(a) Tag I_1



(b) Tag I_2



(c) Tag I_3



(d) Tag I_4

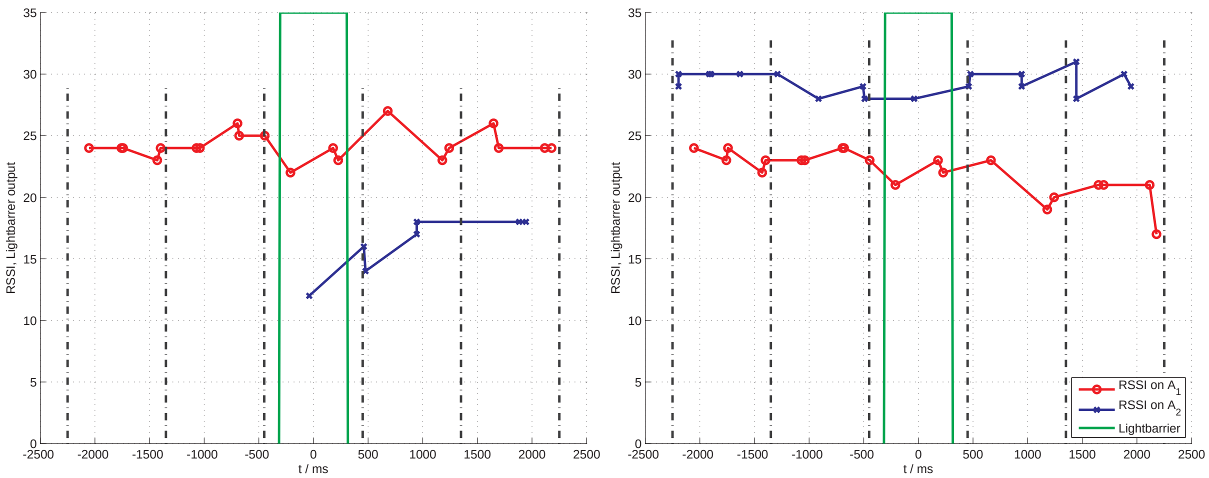
Figure 4.6: Feature sets of four tags inside the same box passing the identification point. Whereas the light barrier provides a deterministic signal, the RSSI patterns show strong fluctuations in amplitude, sampling intervals and number of read events.

symbols:

$$M = N = 4 \tag{4.8}$$

for the shown example. With the structure of the HMM defined, the obtained feature sets $\tilde{\mathbf{F}}$, mapped to a set of observation sequences O can be used to train the HMM, i.e. find appropriate values for $\boldsymbol{\pi}$, $\mathbf{A}$ and $\mathbf{B}$ according to the discussion in 3.3.

The Hidden Markov Model derived in the discussion above represents the class of tags passing the identification point at a constant speed. Since the task of the identification point is to distinguish between moving and stationary tags (perform a classification of the considered feature sets), the same derivations can be applied to the class of stationary tags. The idea is hence to use a second HMM representing the class of stationary tags. It follows from theoretical considerations regarding the received signal strength that tags passing an identification point show a different behaviour than tags which are stationary in the RF-field of an antenna. The latter class of tags will have a more or less constant feature set that is superimposed by random fluctuations. The feature sets for two stationary tags are shown in figure 4.7, as another box was passing the identification point. Due to the fact that the light barrier signal does not provide information about the state of stationary tags, the feature space for stationary tags only consists of the median RSSI value on each antenna. The obtained signals can be mapped to discrete observation



(a) Tag I_5 **(b)** Tag I_6

Figure 4.7: RSSI pattern and light barrier signal for two stationary tags next to the conveyor belt as another box is passing the identification point. In contrast to moving tags, the RSSI patterns are relatively constant, but still superimposed by random fluctuations.

symbols using the procedure described above. The physical setup indicates that the structure of the HMM representing stationary tags is equivalent to the structure of the HMM modeling moving tags, regarding number of hidden states N and number of observation symbols M. The behaviour of the HMM, resulting from the transition and observation probabilities however will be different, since stationary tags are likely to stay in a certain state in the feature space, emitting the same observation symbols again and again. The training process of the second HMM is straightforward with an appropriate data set of stationary tags.

The explanations and discussions so far showed how to derive a probabilistic model of an identification point using Hidden Markov Models. The first step in processing the signals obtained from sensing devices is to apply a windowing technique that splits the signal up into K windows of equal length. The next step is to extract robust features from the windowed signals. These steps are necessary for abstracting a discrete time series model from the reported read events. The extracted features span an d - dimensional vector space, where every sample point origins from the feature extraction within a single window. In order to obtain a training set of observations O from the RSSI patterns, the data points in the feature space need to be partitioned into M labeled clusters, corresponding to the discrete observation symbols $O_1 \ldots O_M$ of the HMM. Using this training data set, the parameters of the HMM can be estimated by employing an iterative algorithm.

The HMMs obtained this way describe the behaviour of moving and stationary tags at an identification point and can account for random fluctuations in RSSI patterns. Moreover, it is possible to use these HMMs for the classification of feature sets. The next section deals with the classification of unknown feature set and presents a framework for the evaluation of classifiers. Using this framework, the performance of the presented model for identification points will be evaluated.

4.2.4 Classification using HMMs

With two appropriately trained HMMs, a classifier with a structure similar to the one shown in figure 3.3 can be used to distinguish between feature sets from moving and stationary tags. The classifier for the particular example with two classes is shown in figure 4.8. The classification is based on the probabilities

$$P(O\,|\lambda_i) \tag{4.9}$$

that the an observation sequence derived from the feature set was generated by one of the two HMMs. This is done most easily by computing the likelihood ratio (LL ratio)

$$\Lambda_I(O) = \frac{P(O\,|\lambda_1)}{P(O\,|\lambda_2)} \lesseqgtr \gamma \tag{4.10}$$

as frequently used in statistical hypothesis testing. The subscript I denotes that this likelihood ratio corresponds to an identification point. In the equation above, γ denotes an appropriate threshold value. The feature extraction also provides two scalar values, w_1 and w_2 which can be interpreted as a measure of the reliability of the current feature set, for example the number of obtained read events. In general, a binary classifier assigns an unknown object to one of two

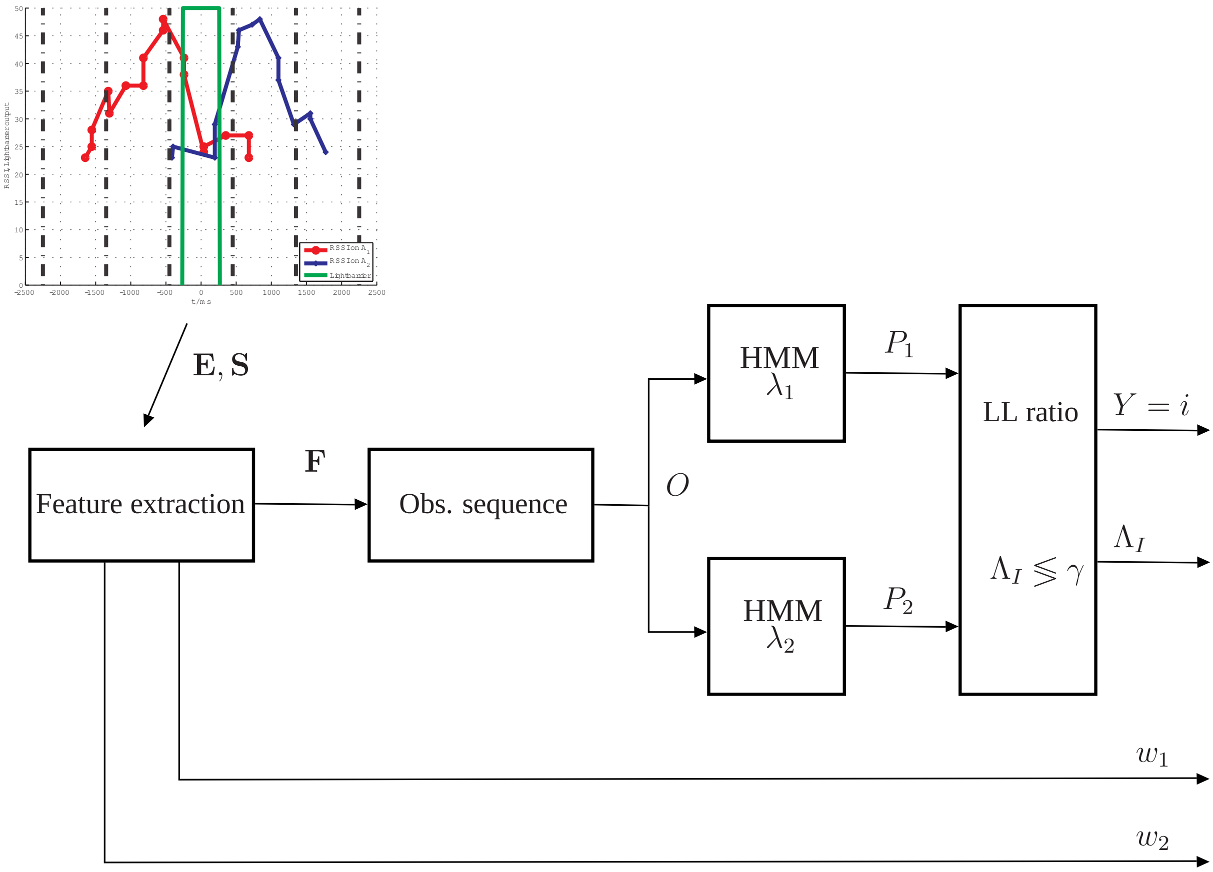


Figure 4.8: Classification of feature sets at an identification point. Based on the signals from sensing devices and the corresponding mapping to discrete observation symbols, two HMMs are used to evaluate the probability of the observed sequence. The classification is performed by evaluation of the resulting likelihood ratio $\Lambda_I = \frac{P_1}{P_2}$. The feature extraction additionally provides two scalar factors w_1 and w_2 which can be interpreted as a reliability measure of the used features.

assigned \ actual	positive: $X = 1$	negative: X = 0
positive: $Y = 1$	true positive	false positive
negative: $Y = 0$	false negative	true negative

Table 4.1: Confusion Matrix of a binary classifier. The actual class of an object is denoted by X whereas the classification result is denoted by Y.

classes. In this case, the output Y of the classifier can take the values 0 or 1, therefore $i \in \{0, 1\}$. An output of $Y = 1$ means that the considered feature set belongs to a tag which is passing the identification point and $Y = 0$ stands for stationary tags. Hence, there are four possibilities for the classification result, summarized in the so called confusion matrix [11], presented in table 4.1, where X denotes the actual class of the unknown pattern. A true positive describes an outcome where the classifier assigns the unknown object correctly to the specific class, i.e. $Y = X = 1$. In the case of feature sets originating from tags passing an identification point, this means that the classifier correctly decides that a feature set belongs to a tag which is passing the identification point. A true negative is also straight forward: The classifier correctly decides that the unknown object does not belong to the considered class and $Y = X = 0$. For a feature set, this means that the identification point correctly assigns the corresponding tag to the class of stationary tags. A false positive as outcome of the classification means that the classifier assigns the object to the considered class, despite it belongs to some other class: $Y = 1, X = 0$. Mapped to an identification point, this means that the classifier erroneously assigns a feature set to the class of tags passing the identification point. A false negative finally indicates that a tag passing the identification point is considered as stationary, hence $Y = 0, X = 1$. Both types of errors introduce severe problems in RFID systems, because incorrect data is reported to the backend system.

Two common metrics, directly following from the confusion matrix are the false positive rate (also called “false alarm rate”) P_{FA} and the true positive rate (also called “hit rate” or “recall”) P_{D}. Given a set of labeled test - data, it is possible to estimate these two metrics according to

$$\hat{P}_{\text{FA}} = \frac{N_{\text{FP}}}{N_{\text{N}}} \tag{4.11}$$

and

$$\hat{P}_{\text{D}} = \frac{N_{\text{TP}}}{N_{\text{P}}}, \tag{4.12}$$

where

$$
\begin{aligned}
N_{\mathrm{FP}} &\ldots \text{number of false positives} \\
N_{\mathrm{N}} &\ldots \text{total number of negatives} \\
N_{\mathrm{TP}} &\ldots \text{number of true positives,} \\
N_{\mathrm{P}} &\ldots \text{total number of positives.}
\end{aligned}
$$

These two performance metrics can be visualized in the so called ROC (Receiver Operating Characteristic) space, which is suitable for the comparison of different classifiers regarding their performance [11]. The general ROC space in figure 4.9 shows three interesting points. First, the point $(0, 0)$ represents classifiers that always reject objects from the assignment to a considered class. On the one hand, this means that no false positives can occur, on the other hand there are also no true positives, since the classifier always rejects the unknown pattern. The point $(1, 1)$ represents the opposite strategy, where all objects are assigned to the considered class, regardless of the observed data. In this case, all true positives are met, but the false positive rate also equals $P_{\mathrm{FA}} = 1$. The point $(0, 1)$ in the ROC space represents perfect classification: There are no false positives, since $P_{\mathrm{FA}} = 0$, and the classifier is able to assign all true positives to the correct class, hence $P_{\mathrm{D}} = 1$. The line $P_{\mathrm{D}} = P_{\mathrm{FA}}$ in figure 4.9 represents the performance of a randomly guessing classifier. If the true positives are guessed right at $50\,\%$ of the time, also the false positives will be $50\,\%$. The ROC space allows for an intuitive interpretation of the

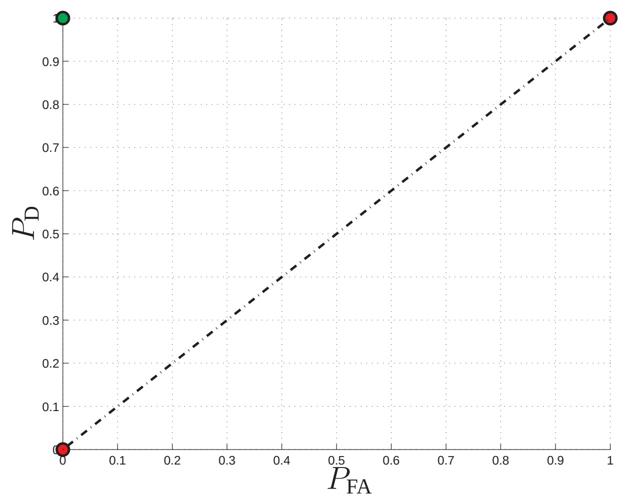


Figure 4.9: General ROC space. The ROC space can be used to assess the quality of binary classifiers by means of the true positive rate P_D and the false alarm rate P_{FA}. The point (0,1) in the ROC space represents a perfect classification result.

classifier performance. The closer the resulting point is to the upper left corner $(0, 1)$, the better is the performance of the classifier. Conversely, as soon as the resulting point lies in the lower right section of the ROC space (below the line $P_{\mathrm{D}} = P_{\mathrm{FA}}$), the performance of the classifier is worse than guessing. Whereas discrete classifiers only output the class of an unknown object, scoring classifiers also provide a probability measure that the object belongs to this class. In the case of two competing HMMs, this score is provided by the likelihood ratio

$$\Lambda_I(O) = \frac{P(O\,|\lambda_1)}{P(O\,|\lambda_2)} \lessgtr \gamma. \tag{4.13}$$

By sweeping the threshold γ for the likelihood ratio in the interval $\gamma \in [-\infty, +\infty]$, a ROC curve can be derived by evaluating the false positive and true positive rate for every value of γ. This way, it is also possible to derive an optimum value for γ as will be demonstrated later.

This framework for the performance analysis of classifiers is applied to the classification of RSSI patterns using a test data set. The data set consists of 15 boxes, each containing 15 tagged items (trousers and shirts), making up a total of 225 moving tags and 10 stationary tags in the vicinity of the reader antennae. The stationary tags have to be considered for every box passing the identification point, hence the total number of stationary tags is 150. This test data set is representative for many practical applications in the retail industry, where it is common that boxes contain around 15-20 tags on average. The resulting ROC curve is shown in figure 4.10. Analyzing the ROC curve now also allows for the determination of the threshold parameter γ, by picking the point that is closest to the point of perfect classification. For this reason, the Euclidean distance from every point along the ROC curve to the point of perfect classification is computed, and the point with the smallest distance is chosen, yielding $\gamma = 1$. This is quite intuitive, since two HMMs of the same structure are used, and the computed probabilities will be within the same range. The likelihood ratio test hence needs to evaluate

$$\Lambda_I(O) = \frac{P(O\,|\lambda_1)}{P(O\,|\lambda_2)} \lessgtr 1, \tag{4.14}$$

where $\Lambda_I(0) > 1$ indicates the assignment to the class of moving tags, and $\Lambda_I(0) < 1$ stands for the assignment to the class of stationary tags. The marginal case of $\Lambda_I(O) = 1$ will per definition also assign a feature set to the class of tags passing the identification point. This gives the classifier a rather liberal behaviour in the case of equally strong evidence.

For the optimal value of $\gamma = 1$, the results regarding false positive and true positive rate are

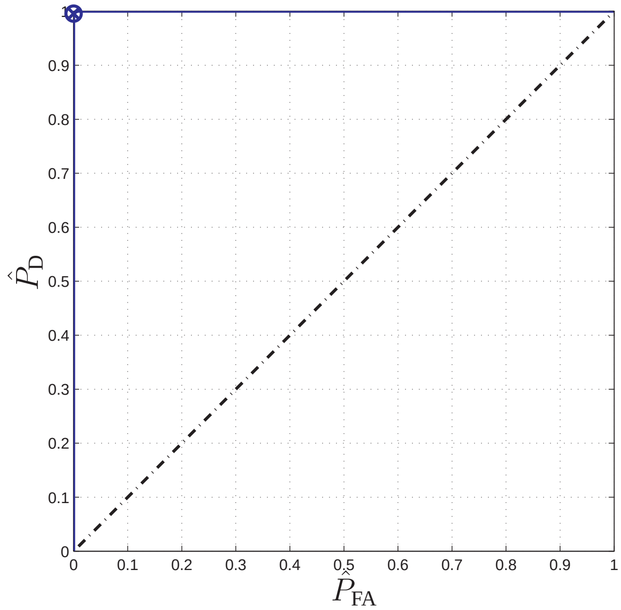


Figure 4.10: Classification result in the ROC space: $\hat{P}_\mathrm{FA} = 0.0067$ and $\hat{P}_\mathrm{D} = 1$ for $\gamma = 1$. These results imply that a single tag was erroneously assigned to the class of tags passing the identification point, despite it is stationary in the vicinity of the reader antenna.

$$N_\mathrm{FP} = 1 \quad \text{and} \quad N_\mathrm{TP} = 225 \tag{4.15}$$

yielding

$$\hat{P}_\mathrm{FA} = 0.0067 \quad \text{and} \quad \hat{P}_\mathrm{D} = 1. \tag{4.16}$$

This shows that the descriptive power of the trained Hidden Markov Models is suitable to perform a classification between stationary tags and tags that are passing the identification point.

Since the task of an identification point is not only to distinguish between moving and stationary tags, the approach above needs to be extended in order to allow for an assignment between tags and packaging units. The physical setup of the considered identification point in figure 4.2 suggests that due to the small spacing between subsequent boxes, it is highly probable that the time-frame of the current box will contain also read events from tags in the previous and subsequent boxes. This issue is demonstrated in figure 4.11. The plot shows RSSI patterns for three tags located in three consecutive boxes where the time-frame is centered to the timestamps of the box in the middle. The distance between the three boxes was chosen as $d_x = 0.5\,\mathrm{m}$ to meet the requirements in practical applications. There is a significant overlap of the consecutive feature sets: When the second box enters the read range of antenna A_1, the first box is still located under antenna A_2. To solve this issue and provide an assignment between tags and packaging

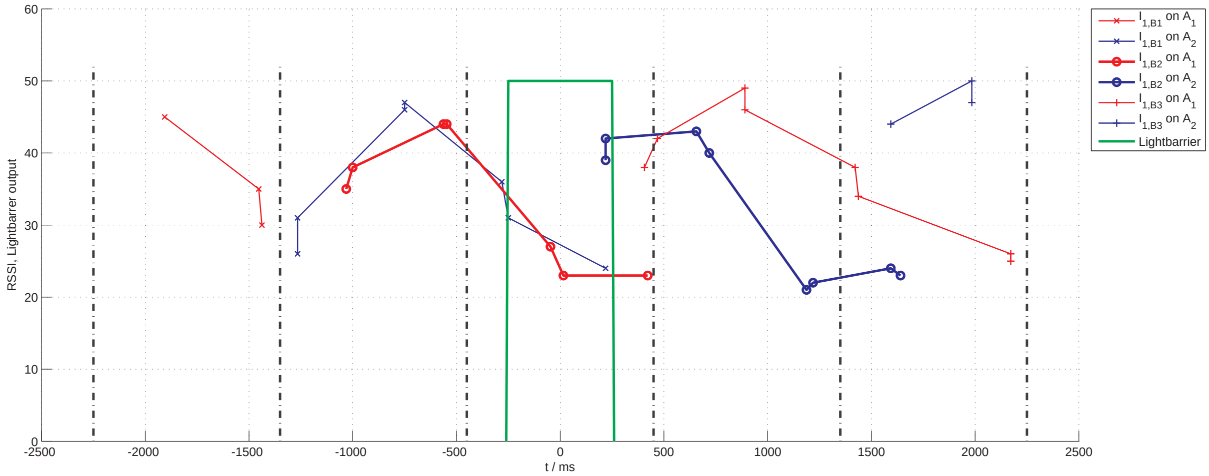


Figure 4.11: Feature sets for tags in three consecutive boxes. The small spacing between two subsequent boxes cause a significant overlap of the feature sets.

units, the classification approach discussed above needs to be extended. In general, every tag will be considered in T consecutive time frames and classification processes. Consequently, this means that there will be T likelihood ratios

$$\mathbf{\Lambda}_I = [\Lambda_{I_1} \quad \Lambda_{I_2} \ldots \Lambda_{I_T}] \tag{4.17}$$

computed for each tag. To assign a tag to a box, simply the time-frame with the highest likelihood ratio needs to be found. This means that the assignment between tag and box can be established $\lceil \frac{T}{2} \rceil$ time-frames after the box was passing the identification point, where $\lceil \cdot \rceil$ denotes the ceiling round operator. The same considerations apply to stationary tags, which will also be considered in the time-frames of subsequent boxes. In this case, the set of likelihood ratios will always indicate that the considered tag is stationary and hence the tag will never be assigned to any box. Figure 4.12 shows the result of the classification process when an assignment between tags and packaging units is performed. The test setup is equal to the tests that were performed for the classification among stationary and moving tags. The optimal point in the ROC space provides the following results regarding false positive and true positive rate:

$$N_{\mathrm{FP}} = 0 \quad \text{and} \quad N_{\mathrm{TP}} = 224 \tag{4.18}$$

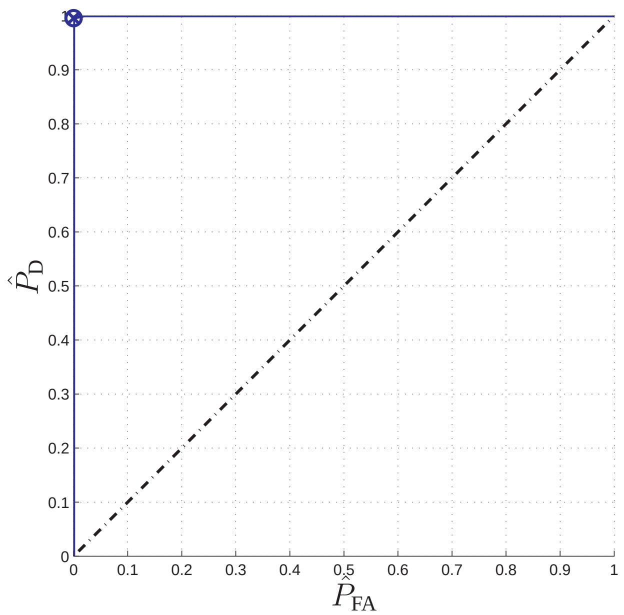


Figure 4.12: Classification result in the ROC space for the assignment between tags and boxes: $\hat{P}_{\mathrm{FA}} = 0.0$ and $\hat{P}_{\mathrm{D}} = 0.9956$ for $\gamma = 1$. This result implies an error free assignment between tags and boxes. However, one moving tag was considered to be stationary.

yielding

$$\hat{P}_{\mathrm{FA}} = 0.00 \quad \text{and} \quad \hat{P}_{\mathrm{D}} = 0.9956. \tag{4.19}$$

A closer look at the result shows that a single tag was considered as stationary, despite it was located in a box passing the identification point. The assignment between tags and boxes however was error free, which points out that the suggested approach is suitable for the localization of RFID tags in practical applications.

4.2.5 Summary

The previous discussion presents a probabilistic model for identification points in RFID systems. Using this model it is possible to perform a classification of feature sets obtained from RFID read events and sensor signals. Moreover, this approach can be extended to perform an assignment between tags and packaging units, which is a common requirement to RFID systems.

Since the read events reported by RFID readers suffer from a non uniform sampling, a windowing technique is used to obtain a discrete time series of feature vectors. Every feature vector represents a robust statistic of the RSSI pattern in each window combined with deterministic information from other sensing devices. The set of feature vectors is mapped to a discrete time observation sequence which can be modeled and evaluated using HMMs.

The presented experimental results proof that the suggested concept is suitable for the localization of RFID tags in practical applications. Since it suffices to know the location of tags with respect to packaging units, the localization can be performed by means of a classification of feature sets. Whereas demonstrated on a specific application, this concept is extendable to other applications, such as RFID portals or dockdoors. Provided the fact that the physical setup of the identification point is optimized to report a sufficient number of RFID read events, HMMs can be used to build a classifier that performs a classification between moving and stationary tags. In addition to that, it is possible to establish an assignment between tags and packaging units as the pass an identification point. Since the problem of false positive reads is encountered quite frequently in UHF RFID systems, this methods is promising to enhance the performance of identification points.

4.3 Business process modeling

The previous section was dealing with a probabilistic model for identification points that can be used in order to perform localization by means of a classification. This section describes how business process information can be used in order to support localization and tracking tasks. For this purpose, every stage of the process is considered as one state of a Hidden Markov Model. The representation of a business process as HMM makes a mathematical treatment possible and allows for the evaluation of the flow of goods in the process.

4.3.1 Basic approach

In general, the task of an RFID system in business processes is to report the identification of items at certain identification points to the backend system. Similar to many other information processing systems, the layer model of an RFID system as shown in figure 1.1 suggests that there are in general three types of packets that can be exchanged between the layers of a system:

- Control packets: These include certain commands, for example between the middleware and RFID readers, such as power settings, poll interval, protocol settings etc. This type of information is used to define the behaviour of an identification point or RFID reader.

- Status packets: Especially large RFID systems need to be monitored regarding the status of devices, connections etc. This is usually done by sending status packets in predefined intervals. For an RFID reader, such a packet can include information about its temperature, antenna connection status, current output power etc. This information is collected by the middleware and also reported to the backend system for management purposes.

- Data packets: Contain the actual information acquired by the RFID systems, such as read events and sensor data. On the higher layers, this information consists of abstract process data, e.g. that an item has changed its location from A to B.

The information flow in a RFID system is shown in figure 4.13. Usually, the middleware collects information about read events and reports directly to the backend system. Consequently, there is usually no data information flow concerning process data from the backend system to the middleware. The basic idea of this section is to establish an information flow between the backend system and the middleware as illustrated by the red arrow in figure 4.13 in order to support localization tasks. This data flow provides the business process information $\mathcal{H}$ as depicted in the system architecture in figure 4.1. As briefly outlined in chapter 1, the underlying business process is able to provide descriptive information about the tracked items subject to the process. Provided that process information is available to the middleware, certain issues regarding the localization of items can be investigated from the business process viewpoint. Since business processes also suffer from uncertainties resulting from hardware defects or human error, also a probabilistic model for business processes appears suitable as discussed in chapter 2. To account for information on the business level, an attempt to model processes using Hidden Markov Models is presented. This approach provides a framework that can be used to investigate on the following three questions in an RFID system. First, it is possible to obtain information about the history of every tagged item in a process. This means that the localization system has the possibility to "ask" where an item came from and how likely it is according to the business process that this particular item is inventoried at the current location. In a mathematical formulation, this can be described as

$$P(\mathrm{ID}^{(k)} \,|\, \mathcal{H}). \tag{4.20}$$

In the equation above, the term $\mathrm{ID}^{(k)}$ stands for the identification event of a tagged item at an identification point k. This probability measure can be used in order to support the classification regarding false positive and false negative reads. Consider again the conveyor belt application presented in the previous section. If an item was already identified at this stage earlier, the likelihood according to the process that it will again be passing this identification point is rather

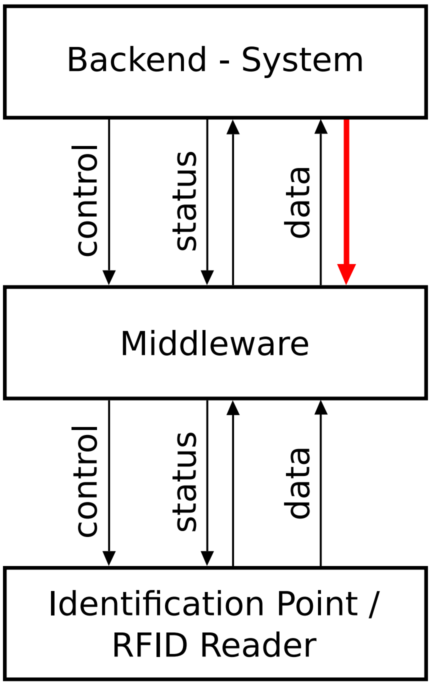


Figure 4.13: Information flow in an RFID system. Similar to many other information processing systems, there are control, status and data packets exchanged between the different layer of the system. According to the suggested system architecture in figure 4.1, establishing a data flow between the backend system and the middleware allows to support localization and tracking in RFID systems.

low. Second, the use of business information enables partly to compensate for weak RFID read events, which were also discussed already in the previous section. The approach of modeling and classifying feature sets with HMMs relies on the information that is contained in the provided RSSI pattern. Data from the business level can partly compensate for a potential lack of information in case of weak read events. If an item provides a low number of read events as it passes an identification point, the classification process might decide that this item was not passing the identification point due to the lack of evidence. If, however the item is following its ideal trajectory through the states of the business process, the likelihood according to the business process that this item is inventoried at the particular identification point will be rather high. Third, the use of business layer information can also be helpful for problems beyond localization. For example, it possible to derive certain measures and statements about hardware components in an RFID system, like RFID tags and readers. Consider for example a simple business process as shown in figure 4.14 where goods are inventoried by identification points at four stages. Imagine that one box leaving the factory provided a sufficient amount of read

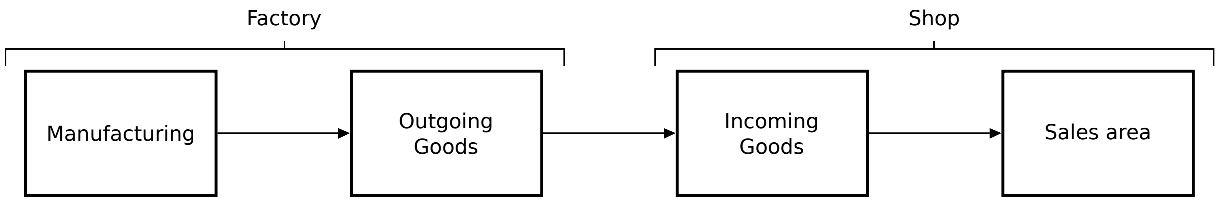


Figure 4.14: Example of a simple, linear business process with four stages.

events at the first two identification points, but was hardly read by the third RFID reader for incoming goods. Then again, the tagged items were identified on the sales floor with a high number of read events. This indicates that the RFID tags in the considered box perform well, but there is a certain probability for hardware problems at the identification point for incoming goods. Conversely, tags that continuously provide weak read events might suffer from degraded performance. This is in particular interesting for tagged items that are subject to periodic processes, in which the RFID tags can be replaced at a certain stage.

These are issues that can not be considered by identification points that operate in an isolated manner. Since the backend system has the possibility to store information about tagged items across the different stages of a process, it becomes apparent that providing this information to the middleware and offers considerable advantages.

The remainder of this section provides a formal set of rules that specify how to model business processes using HMMs and shows how the information provided by the business process can be used in order to support the localization of items. At the end of this section, simulation results from an illustrative business process example will be presented.

4.3.2 Mapping business processes to HMMs

In order to be able to consider the information provided by the business process, it is necessary to find a mathematical representation by means of a model. The mapping of common process models such as Petri Nets to Hidden Markov Models which offer more mathematical flexibility has been covered in recent publications [1, 24]. These publications present a general framework for creating HMMs from Petri Nets and use them for data mining and the evaluation of the quality of derived process models. This work focuses on the application of HMMs to model business processes for RFID systems. As discussed earlier in this chapter, an RFID business process consists of a set of distinct identification points, corresponding to the stages of the process. These stages can be interpreted as the states of a Hidden Markov Model. The transition probabilities between the different states are defined by the rules governing the process. The possible structures of HMMs offer a great flexibility for modeling processes: Whereas most processes are strictly linear, HMMs are also able to account for loops and periodic processes, simply by choosing appropriate transition probabilities between the corresponding states. The choice of transition probabilities also makes it possible to model uncertainties. Considering the example process in figure 4.14, it is possible that items skip certain stages in linear process models, either due to hardware defects or the lack of RFID read events. This behaviour can also be modeled by HMMs, by allowing transitions not only to the direct successor state, but also to the next state after that.

To map a given process to a Hidden Markov Model denoted by $\lambda_{\mathcal{B}}$, the following considerations need to be applied:

- Every identification point that acquires information about the items subject to the process corresponds to one state in the HMM. Since there are also other types of readers (bar code readers, scales etc), the term identification point needs to be considered in an extended way. In addition to that, one state of the HMM is required to model defect tags, i.e. tags that can not be identified by means of an RFID reader.

- The prior state distribution probabilities $\boldsymbol{\pi}$ are derived in a straight forward manner from the rules describing the business process. For example, some processes employing RFID have one or more “tagging stations” which are used to attach RFID tags to the considered items. These tagging stations can be considered as the start state in most cases.

- The transition probabilities a_{ij} follow directly from the ideal trajectory of an object through the different stages of the process. Whereas this is quite simple and intuitive for linear process chains, more complex structures such as loops additionally require considerations about possible successors to each state and the corresponding transition frequencies.

- The observation probabilities b_{ij} will equal unity for most cases, but can also be adjusted to consider crosstalk between closely spaced RFID identification points.

- To allow for arbitrary state transitions, the transition probability matrix $\mathbf{A}$ needs to be modified such that there are no impossible state transitions. This is done by setting all zero probabilities to an arbitrary small value ϵ. This can be thought of adding some noise to the state transitions, which in turn provides the considered HMM with more flexibility. To ensure that $\mathbf{A}$ is a row stochastic matrix (i.e. the rows sum up to 1), the following modifications need to be made:

$$A(m,n) = \begin{cases} \epsilon \cdot \frac{|\{(m,n') \mid A(m,n')>0\}|}{|\{(m,n') \mid A(m,n')=0\}|} & A(m,n) = 0 \\ A(m,n) - \epsilon & A(m,n) > 0 \end{cases} \tag{4.21}$$

  where $|\{(m,n') \mid A(m,n') = 0\}|$ denotes the number of elements in row m that are equal to zero and $|\{(m,n') \mid A(m,n') > 0\}|$ denotes the number of elements in row m which are greater than zero. Depending on whether the process is a strictly linear left to right process, backward transitions may or may not be allowed. The same modifications need to be done for the prior state probabilities vector $\boldsymbol{\pi}$ in order to allow for arbitrary start states.

The rules above indicate that some expert knowledge is necessary in order to find good parameters that define the behaviour of the Hidden Markov Model. Since this is an issue that may not be fulfilled, especially when considering large and complex processes, the parameters of the business process HMM $\lambda_{\mathcal{B}} = (\boldsymbol{\pi}, \mathbf{A}, \mathbf{B})$ can also be learned by means of logged observation sequences. This training procedure can be performed in the same iterative way as described in section 3.3.

To demonstrate the application of these rules to a given business process, a slightly more complex example compared to the process in figure 4.14 is investigated. Consider a process in the retail sector, consisting of the following five identification points:

- 1 - Packing table: After the manufacturing of clothes, they are put in packaging units for transportation. Packing tables are used quite frequently as tagging stations (in order to provide the items with RFID tags) and for quality assurance. This identification point consists of a single RFID reader per table that reports the identifier of each item to the backend system.
- 2 - Outgoing goods: As soon as the assembled packaging units are meant to leave the factory, their content is checked by means of an RFID portal, which allows for the identification of a large number of items, e.g. on pallets. At this stage, it is only possible to check if the current delivery is complete, since no assignment between items and packaging units can be established.
- 3 - Incoming goods I: The ordered items are shipped from the factory to the storage hall of a shop, where the completeness of the order is again checked by means of an RFID portal.
- 4 - Incoming goods II: Assuming a complete order, every packaging unit is checked a second time while it is moved on a conveyor belt to its desired storage space. This identification is used in order to guarantee that every packaging unit contains the desired items.
- 5 - Unpacking table: As soon as the items are meant to be placed on a sales floor in the shop, they need to be unpacked from the boxes in the storage hall. During this process, another identification is performed, similar to the packing table at the beginning of the process.

Applying the mapping rules to the process above yields the HMM shown in figure 4.15, where the transition probabilities have not yet been provided with additional noise. Since there are five identification points in this process, the total number of states in the HMM is six, because the defect state needs to be added. The transition probabilities across the different states follow straight forward from the process description above. In order to account for the possibility that items are not inventoried at the identification points 3, 4 and 5, also transitions to the states after the direct successors are allowed. Identification points 2 and 6 are however expected to provide reliable information, since human interaction is involved during packing and unpacking. The transition probability for every state to enter the defect state is according to the defect

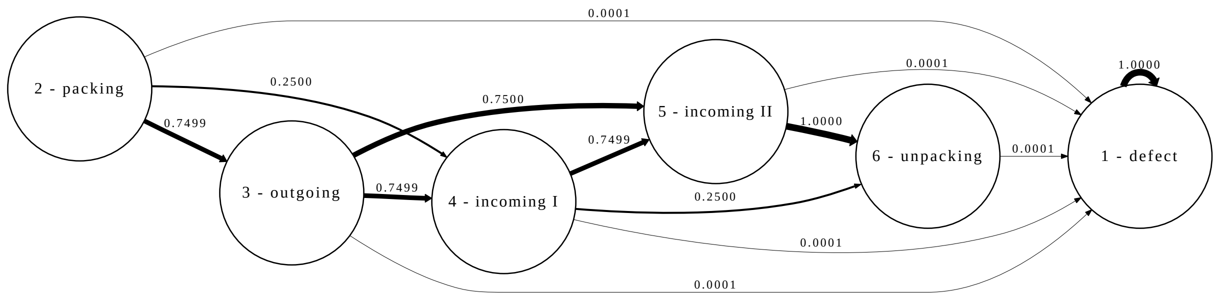


Figure 4.15: HMM structure for the exemplary business process with five stages. In addition to the stages of the process, a defect state needs to be modeled in order to account for the possibility that a tag can not be identified by an RFID reader. The transition probabilities between the states follow from the rules governing the process. For the sake of simplicity, the observation symbols are not shown. The numerical values of the HMM are provided in appendix A.1.

probability P_D of the used RFID tags, which is specified by the manufacturer. For the sake of simplicity, the observation symbols with the according probabilities are not shown in figure 4.15, since $b_{ij} = 1 \,\forall\, i = j$.

Using this HMM, it is possible to compute the probability $P(\mathcal{O} \mid \lambda_{\mathcal{B}})$ of trajectories (i.e. observation sequences $\mathcal{O} = \mathcal{O}_1\mathcal{O}_2 \ldots \mathcal{O}_N$) for any given item in the process. The calligraphic notation is used here in order to distinguish the quantities corresponding to the business process from the quantities used in the previous section to model signals at an identification point. The likelihood of an observed sequence can be compared to an ideal observation sequence $\mathcal{O}_i$, by means of the likelihood ratio

$$\Lambda_{\mathcal{B}}(\mathcal{O}) = \frac{P_{\mathcal{B}}(\mathcal{O})}{P_{\mathcal{B}}(\mathcal{O}_i)} = \frac{P(\mathcal{O} \mid \lambda_{\mathcal{B}})}{P(\mathcal{O}_i \mid \lambda_{\mathcal{B}})}. \tag{4.22}$$

The ideal observation sequence is the sequence that maximizes the likelihood at the current identification point and can be determined in a straight forward manner from the transition probability matrix $\mathbf{A}$. The likelihood ratio computed this way indicates to which extend an observation sequence of states represents a valid sequence in the business process. The resulting structure is depicted in figure 4.16. The derived process model can now be used in addition to the RFID read events and sensor signals to perform a classification at an identification point. For this reason, the classification system is extended by the business process classifier. There exist different methods of combining the result of several classifiers, especially for classifiers offering a "soft output" by means of a probability measure rather than a class label only [25].

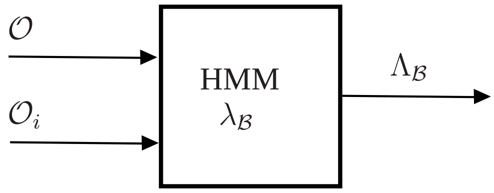


Figure 4.16: Classification according to the business layer information. The observation sequence $\mathcal{O}$ is compared to the ideal sequence $\mathcal{O}_i$ of an item at the current stage of the process. This provides the likelihood $\Lambda_{\mathcal{B}}$ of the item's trajectory through the process.

One approach, called *Bayes Average* is to compute the average posterior probability from all classifiers. Since the classifiers in this work provide likelihood measures rather than posterior probabilities, this idea needs to be adapted accordingly. To assign a tag in one of two classes, a weighted sum of the classification result based on the obtained feature sets and the business layer classification is computed:

$$\Lambda = w_1 \Lambda_I + w_2 \Lambda_{\mathcal{B}} \lesseqgtr \tilde{\gamma} \tag{4.23}$$

where $\tilde{\gamma}$ denotes an appropriate threshold value in the domain of the likelihood Λ. The scalar factors w_1 and w_2 can be used to adjust the weights of either classification result. The two weights are chosen such that

$$w_1 + w_2 = 1. \tag{4.24}$$

The structure of the suggested classifier, based on the fusion of the identification point classifier and the business layer information is shown in figure 4.17. In general, the feature extraction is able to provide reliability measures of the extracted feature sets. The idea for the fusion of the two classifiers is to use these reliability measures as weights for the individual classification results. Since the number of read events is a measure for how well a tag was seen at an identification point, this is one possibility to derive a reliability measure. In order to fulfil the requirement of equation 4.24, the total number of read counts needs to be normalized. Furthermore, it is suitable to apply a deterministic function to the number of read events that adjusts the weights accordingly. The final decision about the particular sequence of read events is performed by evaluating the weighted sum of the individual likelihood ratios.

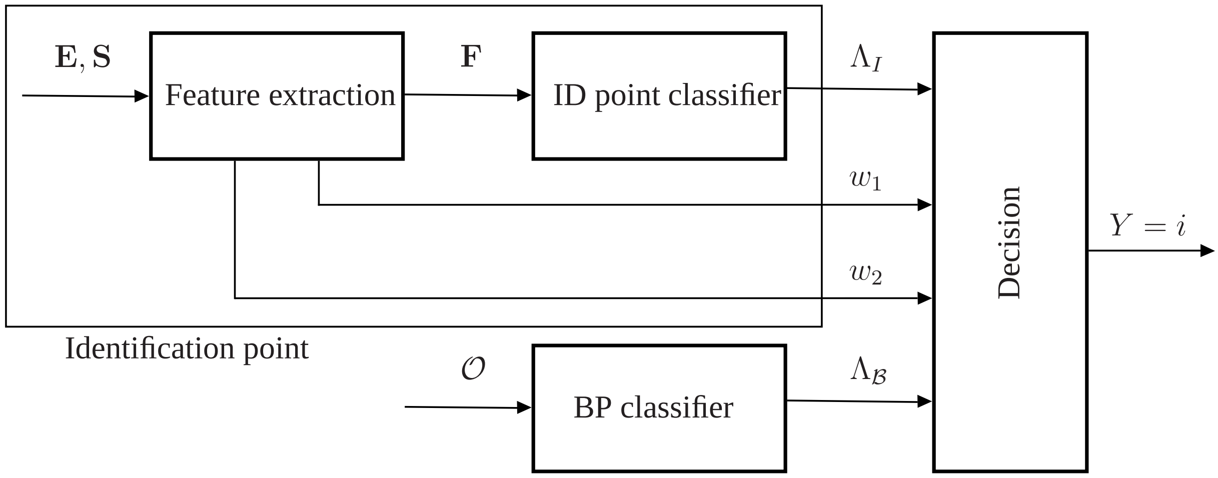


Figure 4.17: Classifier fusion: The classification result Λ_I based on the information obtained from the feature set $\mathbf{F}$ and the classification result $\Lambda_\mathcal{B}$ according to the business process information are used together for a decision whether an item was identified at the particular identification point.

4.3.3 Model Evaluation

The performance of the system depicted in figure 4.17 is analyzed by means of a simulation. To demonstrate the effects when business layer information is considered, the identification point "incoming goods I" from the exemplary business processed is simulated. For this purpose, an arbitrary identification point classifier, characterized by its false positive and true positive rate is combined with the business process classifier as described in the previous paragraph. To show the descriptive power of the business layer information, a deliberately weak classifier is chosen for the evaluation of feature sets. The chosen classifier is based on two overlapping Gaussian distributions for tags passing the identification point and stationary tags. The corresponding ROC curve is shown in figure 4.18. The chosen classifier has a considerable false positive rate and a moderate true positive rate which would not be acceptable in practice.

In general, a process as shown in figure 4.15 is subject to two kinds of errors. First, there is the possibility that an item is not identified by an RFID reader at a particular stage of the process. This is an issue for which the process model can account for by choosing appropriate transition probabilities. Second, there is the possibility that items are not moving along the ideal trajectory through the process. Applied to the example in figure 4.15, this could mean that an item that has already been shipped through the RFID portal for incoming goods is by accident passing the same identification point again. Or even worse, an item from a later stage of the

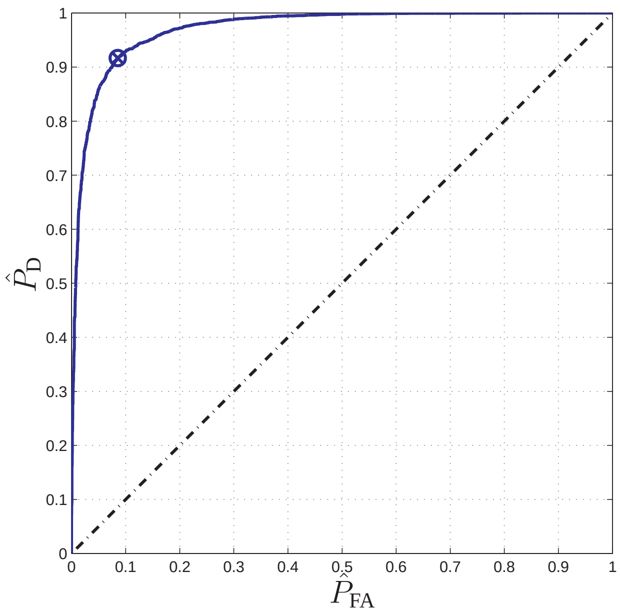


Figure 4.18: ROC curve of the simulated identification point classifier: $\hat{P}_{\mathrm{FA}} = 0.085$, $\hat{P}_{\mathrm{D}} = 0.917$.

process is again shipped through the RFID portal for incoming goods. From a process point of view, there are hence two possibilities:

- Items following a valid trajectory through the business process are passing an identification point. These items have a high probability of being identified according to the process model.
- Items following an invalid trajectory through the process. Consequently, these items should not be identified at the considered identification point since the probability according to the business process model is low.

Considering the possibility that items can also be placed in the vicinity of an identification point without shipping them through deliberately, this gives four possible combinations:

- Items following the ideal trajectory through the business process are passing an identification point.
- Items following an invalid trajectory through the business process are located in the vicinity of an identification point.
- Items following an invalid trajectory through the business process are passing an identification point.

- Items following the ideal trajectory through the business process are located in the vicinity of an identification point.

To account for these possibilities resulting from process imperfections, a noise level is introduced in the simulation which specifies the percentage of items following an invalid trajectory through the business process:

$$n = \frac{\text{\# of erroneous identification events}}{\text{\# of total identification events}}. \tag{4.25}$$

For example, $n = 0.1$ means that 5% of all identification events result from tags that are passing a particular identification point though they should be in another stage of the process and 5% result from stationary tags that should actually be identified, yielding a total percentage of 10%.

As described in section 4.2, the feature extraction provides a reliability measure w_1, w_2 of the considered features. As an exemplary reliability measure for feature sets obtained from RFID read events and sensor signals, the simulation also includes values for the number of inventories for each tag. Based on empirical data, the read count is assumed to have a Gaussian distribution

$$M \sim \mathcal{N}(\mu_M, \sigma_M^2) \tag{4.26}$$

where $\mu_M = 20$ and $\sigma_M^2 = 100$ represent the mean value and variance of the Gaussian distribution. Using this measure, the weighting factor w_1 is computed by means of an empirically found function

$$w_1 = f(m) = \frac{1}{k_r}\sqrt[4]{m} \tag{4.27}$$

where k_r is a normalizing factor that can be interpreted as the maximally expected number of read events, $k_r = \frac{1}{m_{\max}}$. The function applied to the number of read events ensures that the classification result obtained by unreliable read events get a low weight, as shown in figure 4.19. For the extreme case of a single read event $r = 1$, the classification result of this read events is almost neglected in favour of the business process information, since $w_1 \ll$. Conversely, if the number of read events is high, the classification result from the identification point classifier is considered as more reliable and the business information gets a lower weight. Whereas the general relationship between the weighting factor w_1 and the number of read events as reliability measure is based on the considerations above, the function in equation 4.27 was found empirically.

Using the described setup, several simulations are performed. The first simulation considers

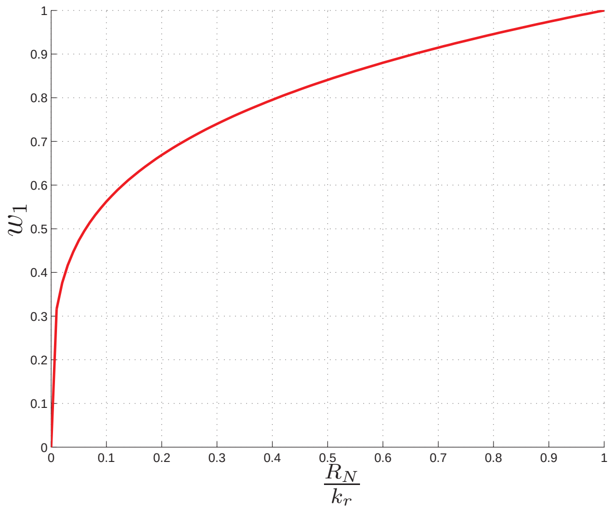


Figure 4.19: Empirically found relationship between weighting factor and relative read count. The weighting factor w_1 is computed according to the relative number of read events. The lower the number of read events, the lower is also the weight on the classification result of the identification point.

a reliable business process with $n = 0$. The ROC curve for $K = 10000$ identification events is shown in figure 4.20. In comparison to the identification point classifier, the result is improved considerably. The false alarm rate decreased by simultaneously increasing the true positive rate. Mapping this result to the RSSI pattern classifier presented in the last section suggests that a perfect classification is possible if the identification point model is used in combination with reliable business layer information. Moreover, the results show that this approach is suitable to compensate for weak RFID read events and weak classifier performance.

To show the impact of imperfect business processes, the next simulation shows the classification results when the noise level is varied in the interval $[0, 0.5]$. A noise level of $n = 0.5$ means that half of all read events result from tags that are not in the appropriate stage of the process. This is a value that exceeds the amount of every practical process by far, nevertheless it demonstrates the system behaviour if no reliable business process information is available.

For every noise level n, the performance characteristics of the system are evaluated. Figure 4.21(a) shows the resulting true positive rate as a function of the process noise level, whereas the corresponding false positive rate is visualized in figure 4.21(b). An increasing amount of process noise leads to a degradation of the classifier performance: The true positive rate decreases whereas the false positive rate constantly increases. The effects of unreliable business processes can be compensated to a certain extend by the probabilistic nature of the HMM. If,

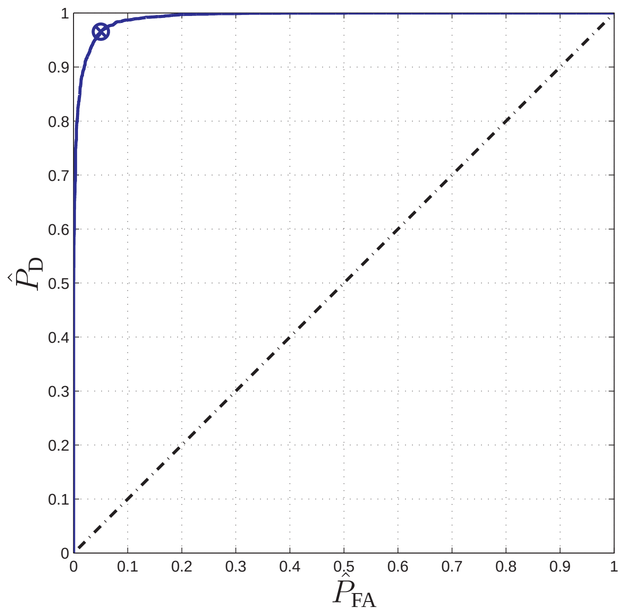
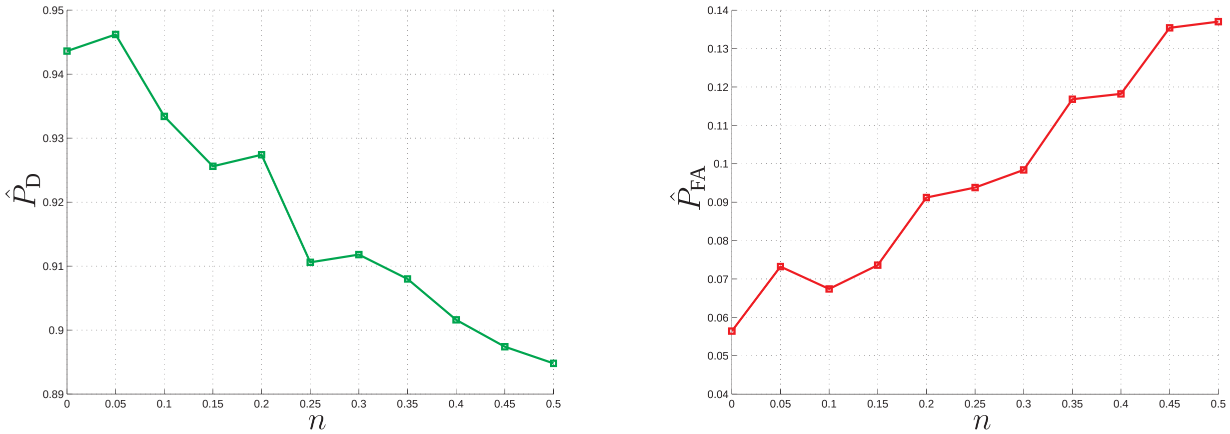


Figure 4.20: ROC curve for the classification in an error free business process with $n = 0$: $\hat{P}_{\mathrm{FA}} = 0.058$, $\hat{P}_{\mathrm{D}} = 0.944$. Compared to the isolated identification point, the performance is improved significantly by considering information from the business process layer.

however the process noise dominates, the performance of the overall system might decrease to a level that is even below an isolated identification point. For realistic process noise levels, the business layer information provides a valuable support in terms of localization.

This evaluation proofs that considering business layer information in an appropriate model is suitable for improving the localization of RFID tags in practical applications. Since HMMs can account for errors in the process, the localization reliability on a particular identification point can be significantly increased. Combined with the approach of modeling RSSI patterns and sensor signals with HMMs, the results are promising to enhance the performance of RFID systems in practical applications.



(a) True positive rate

(b) False positive rate

Figure 4.21: Performance characteristics for varying process noise levels. An increase of the process noise n yields to a degradation of the classifier performance. The detection rate constantly decreases, whereas the false alarm rate is increasing with n.

5 Use Cases

Due to the fact that the RFID market is constantly growing, more and more business processes in the industry use RFID systems for localization and tracking of goods. The developed models have been implemented in a middleware and are used in several projects in the industry. In this chapter, section 5.1 gives a general overview over the basics of an RFID middleware and describes how the mechanisms for localization can be integrated in existing software frameworks. The remainder of this chapter then describes two use cases and how the derived models have been adapted in order to fulfill the particular requirements. Due to the fact that the projects implementing the developed models are still ongoing, only preliminary results from integration tests are available.

5.1 RFID middleware

As briefly discussed in chapter 1, the middleware is the central layer of an RFID system. The core functionality is hence the integration of various hardware components, such as RFID readers and bar code scanners on the one, and interfaces to backend business applications on the other hand. To the developer of an application, hardware components need to be abstracted in a standardized way, such that differences in the various types of RFID readers are covered by a dedicated software layer. Depending on the particular application, the middleware also needs to provide management tools and interfaces to other applications as well as mechanisms for user interaction.

Considering an object-oriented approach in software development, the developed models can be implemented in encapsulated classes which can easily be integrated into an existing software framework. Since tag read events are reported in a standardized way, appropriate data structures can be used to store these events for processing. The used middleware is based on the Microsoft .NET framework and the programing language C#. Besides the object oriented approach, this language offers various comfortable features, such as multicast-delegates (also called events). The algorithms for classification using HMMs are based on a HMM implementation under the Code Project Open License (CPOL) [9].

After this brief introduction, the next two sections cover two use cases from ongoing projects in the industry.

5.2 Use Case 1: Application in fashion logistics

The first use case where the developed models have been applied deals with tracking of clothes in a logistic warehouse, where goods are transported in boxes on motorized conveyor belts. The speed of the conveyor belt is $v = 0.6\,\frac{\mathrm{m}}{\mathrm{s}}$ and packaging units contain 20 tags on average, with a maximum of up to 80 tags. The requirements to the RFID system are similar to the ones described in chapter 1:

- The RFID system needs to establish an assignment between goods and packaging units, i.e. tags and boxes. In order to check the content of packed boxes for correctness and
- perform a classification of read events in order to filter out false positive reads.

Due to the given geometry in the storage hall, the possible locations for installing an identification point are limited to a region where several conveyor belt lines are closely spaced to each other. The physical setup of the identification point consisting of the antennae A_1 and A_2 is depicted in figure 5.1. Boxes on line L_2 pass the identification point where the content is inventoried. The RFID system is ought to report the content of every box to the backend system, which compares the identified tags to a given target list. In case of a mismatch, the considered box is directed to an alternative destination, where its content is checked manually. For this

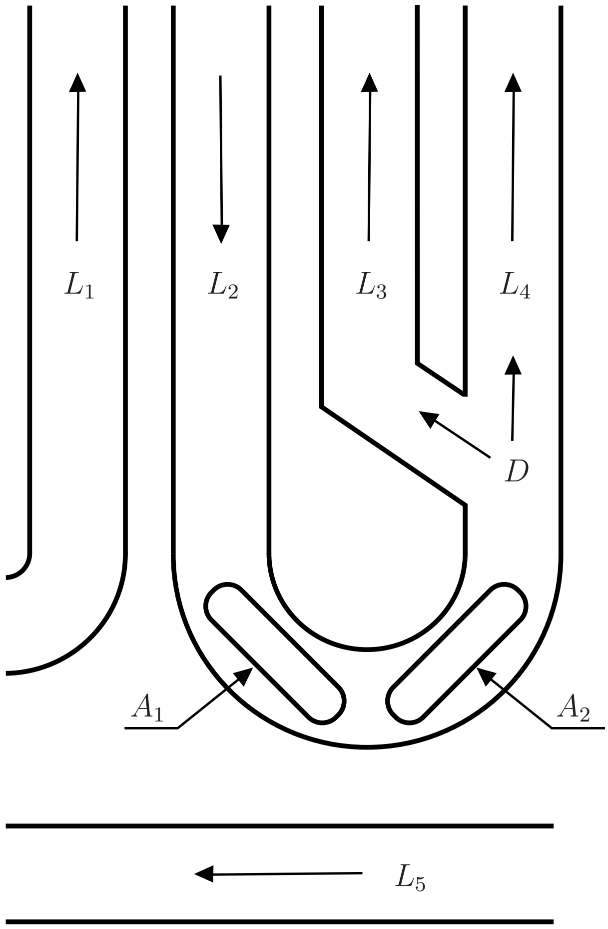
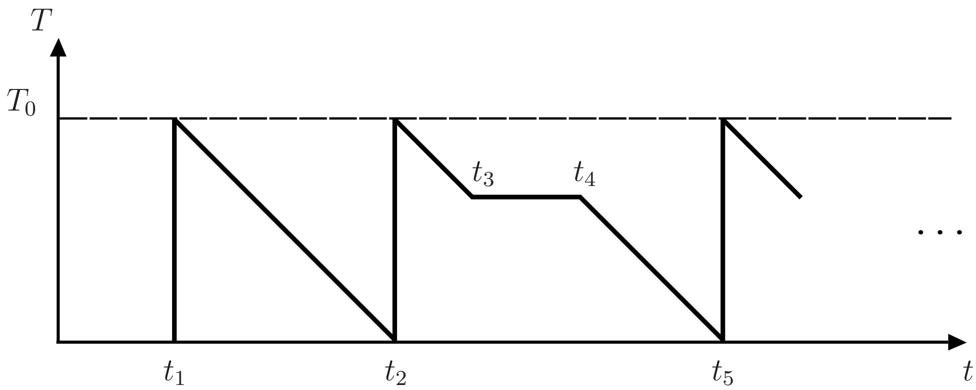


Figure 5.1: Use Case 1: Physical setup. The identification point operates in a storage hall with a total of five conveyor belt lines $L_1 \dots L_5$. The task of the RFID system is to perform a classification regarding false positive reads and to establish an assignment between tags and moving boxes.

reason, the processing time available to the RFID system is limited: The content of a box must be reported to the backend system before the considered box is passing the so called “pusher” (denoted as D in figure 5.1), an actuating device that changes the direction of the box by pushing it on another conveyor belt. Line L_4 is meant for boxes with correct content, whereas L_3 transports boxes to another destination where the content is checked manually. The approach discussed in chapter 4 for establishing an assignment between tags and packaging units can not be followed in this case, because the classification result is needed immediately after the box left the read range of the second antenna. One idea to cope with this fact is to consider tags from preceding or succeeding boxes in the background HMM that models stationary tags. As it turned out during a simulation, this does not provide satisfactory results, because the different feature sets can not be sufficiently modeled by a single HMM. For this reason, there are two more HMMs used, which represent tags from preceding and succeeding boxes. The training of these HMMs is straight forward: Consider feature sets of preceding or succeeding tags and assign the features to observation sequences which can then be used as a training data set. An unknown feature can then be classified as demonstrated in chapter 4.

An additional challenge in this setup is that the conveyor belt may stop at any time due to a congestion of boxes. On the one hand, this issue has a big disadvantage, because the Hidden Markov Model for the classification of feature sets are based on the assumption of a continuous movement of the tagged items. If this is not the case, the classification of the extracted time series will provide false results since it is based on a wrong assumption. On the other hand, this also provides the advantage that boxes which are located right underneath an antenna will provide a sufficient amount of read events for classification if the conveyor belt is stopped. Nevertheless, some modifications need to be made to the approach suggested in chapter 4 in order to ensure that the extracted feature set is according to the normalization of the time-axis when the conveyor belt is stopped.

To consider the issue of a stopping conveyor, an additional digital signal is used that indicates the current state of the conveyor belt – either running or not running. Under normal conditions, the window borders for the extraction of features are computed according to the chosen window size and the time instant when the box is passing the identification point. Given that the conveyor is stopped, the tags inside a box do not physically change their state. For this reason, the window border computation is based on a timer that starts as soon as a box enters the read range of the first antenna. This time instant is indicated by means of a light barrier signal. The timer is initialized with the window size and is constantly decreasing its value, as long as the conveyor belt is moving. The moment when the timer value attains zero indicates the timestamp of the next window border. The timer is then reinitialized and starts decrementing its value again. In case that the conveyor stops while the box is in the read range of the identification point, the timer is paused and the window border is delayed appropriately. This mechanism for adaptive window borders is shown in figure 5.2.



t_1 ...box enters read range, timer starts, t_1 is first window border
t_2 ...timer elapses, t_2 is second window border
t_3 ...conveyor stops, pause timer
t_4 ...conveyor starts, resume timer
t_5 ...timer elapses, t_5 is third window border

Figure 5.2: Adaptive window border computation

Using this mechanism allows us to consider the fact that the conveyor belt might stop during operation, without the need of modifying the Hidden Markov Model for classification. The RSSI pattern of a tag that stopped twice while passing the identification point is shown in figure 5.3. The window borders in figure 5.3 were computed according to the considerations above and are adapted properly to the periods where the conveyor belt is stopped.

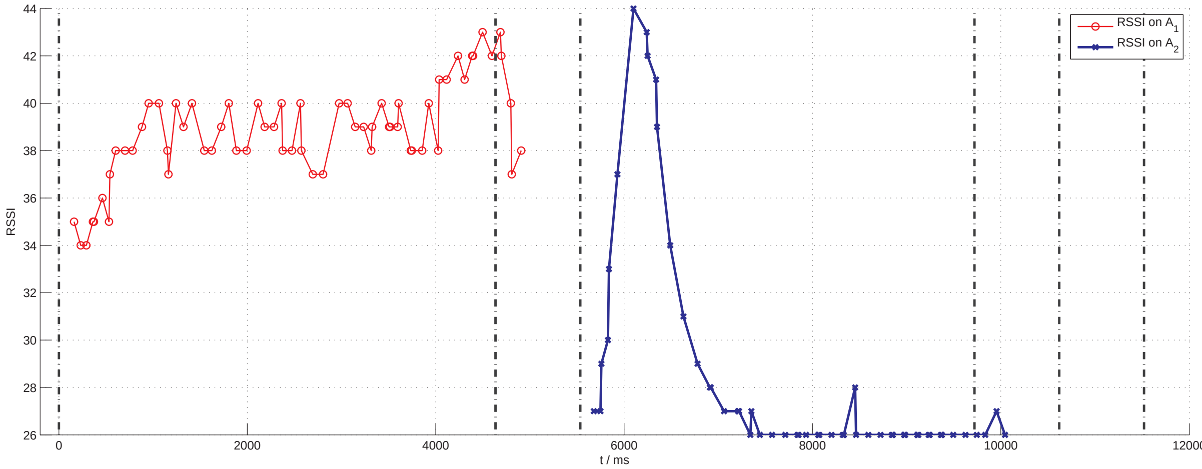


Figure 5.3: RSSI pattern in case of a stopping conveyor

Filtering out false positive reads from tags in the opposite direction on adjacent conveyor belts requires another modification. Since tags that are passing the identification point in the opposite direction will also move through the feature space in the opposite direction, it becomes apparent that the HMM for tags moving in the correct direction will output a very low likelihood. To detect tags moving in the opposite direction, a dedicated HMM can be used that represents tags moving in the backward direction. The implemented system hence uses a total of five HMMs (forward, backward, stationary, preceding and succeeding tags). The HMM that outputs the highest likelihood for a given observation sequence represents the estimated class.

This use case demonstrates the flexibility of the suggested approach. The basic idea of modeling RSSI patterns and sensor signals as discrete time series offers the possibility to account for various kinds of behavioural patterns. The classification of feature sets using HMMs allows to report reliable information to the backend system.

The results that were achieved using this approach during an on site integration test proof the concept with performance metrics equal to the results in the previous chapter. Due to the fact that the discussed use case represents a pilot project with the particular customer, the shown identification point operates in an isolated manner. For this reason, it is not possible to include business layer information to support the classification of read events. A project that deals with a whole process using RFID is described in the next section.

5.3 Use Case 2: Tracking of fruit trays

The second use case to be described covers the tracking of fruit trays during the harvesting season of the year. The considered fruit trays (containing up to 300 kg of fruit, mostly apples and pears) are used for transportation and storage. The task of the RFID system is to keep track of every single tray throughout the periodic process. For this reason, every tray is equipped with two RFID tags that have a unique identifier. On the one hand, the use of two tags is for redundancy purposes, on the other hand it also offers increased RFID readability. In addition to that, bar code labels are attached to the trays, containing the same unique identifiers.

Since the trays carry a considerable amount of fruit with high water content, the setup of identification points is quite challenging. Water in general is absorbing radio waves in the considered frequency range. This means that RFID tags can be shielded or occluded in case they are surrounded by water. For this reason, a sophisticated optimization of the geometry of identification points regarding placement and direction of antennae. is necessary in order to provide reliable RFID read events.

From the process perspective, there are several stages for every fruit tray throughout a life cycle:

Step 1: The first stage for every tagged tray is the issuing to a specific farmer. This delivery is monitored by handheld devices. The person delivering the trays has to identify all trays by means of their RFID tags and / or bar code labels. The information about which trays were shipped to a specific farmer is stored in a central database.

Step 2: After the fruit harvest, trays issued to one specific farmer are returned by a lorry to the central storage hall. The lorries are unloaded using pallet trucks that ship the fruit trays through one of four RFID portals. Every pallet truck used for unloading the lorry can carry up to six trays. The RFID portals are set out adjacently to each other in a so called dockdoor configuration, where it is necessary to determine through which portal a tray was shipped. After the identification by means of the RFID portals, the trays are weighed to determine the amount of harvested fruit.

Step 3: After the inbound of trays, they are moved to a storage hall by a forklift which is equipped with RFID hardware. Instead of using an RFID portal at each storage hall entry, RFID tags assembled in the ground floor are used to associate the current fruit tray with the specific storage hall number. When the forklift passes through the hall entry, it will read the floor tags and hence know that it shipped the currently carried fruit trays to the particular storage hall.

Step 4: After the so called interim storage, fruit trays are removed from the storage hall by RFID equipped forklifts and carried to a sorting plant where the trays are emptied and the different cultivars of fruits are stored in new trays. The old trays are cleaned and can again be shipped to a farmer, be used as storage for sorted fruits or can even be destroyed in case they are in bad physical condition. In any case, the considered trays get the specific status “unassigned”, which means that they can be reused in any part of the process where necessary.

Step 5: Sorted fruits are stored in trays which are then again moved to a certain long term storage hall by an RFID equipped forklift. After sorting, a tray contains a single cultivar of fruit.

Step 6: At the end of the storage period which is up to one year, trays are removed from the storage hall by RFID equipped forklifts and carried to packing lines. There, each tray is identified by an RFID reader, and the content of the tray is packed into smaller packaging units for reselling in supermarkets. After this step, the tray is again cleaned and will be reused in the process or destroyed.

The description above indicates that this process is rather complex in contrast to the simpler examples covered in chapter 4. Whereas most processes in the industry are linear, this process is periodic and trays can take various trajectories through the different stages. Applying the formal set of rules derived in the last chapter, this process can be mapped to a Hidden Markov Model with nine distinct states. The transition probabilities between these states have been derived in discussions with the particular customer and are based on counting absolute frequencies of the occurrence of events. The resulting HMM is depicted in figure 5.4, where the observation symbols are not shown for the sake of simplicity.

The first issue in this process is to resolve the crosstalk between the adjacently placed RFID portals at the inbound of fruit trays. These four portals have been summarized in a “Macro state” in the process HMM in figure 5.4, despite they are considered as four distinct identification points with appropriate observation symbol probabilities. Regarding the transition probabilities, the four portals are considered to be equally likely. In order to account for the crosstalk problem between adjacent portals, the RSSI signal energy of every tag moving through the portal is compared to all other portals. Based on this information and the prior knowledge that

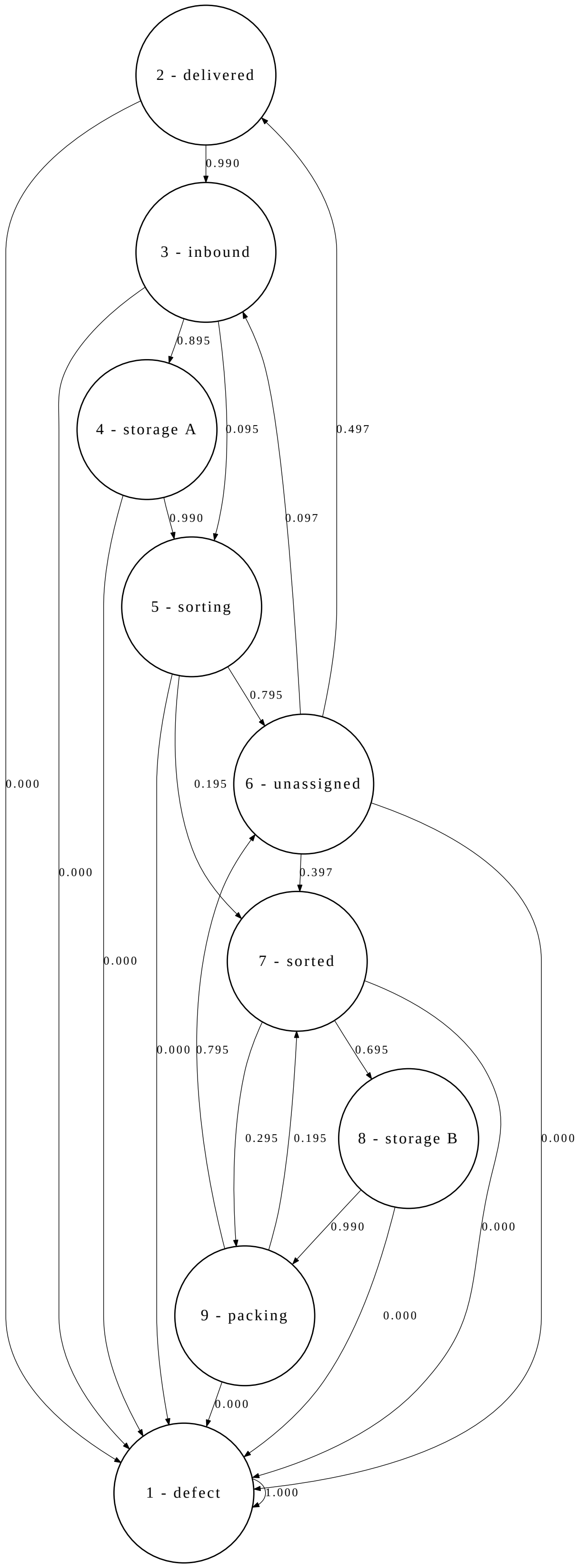


Figure 5.4: HMM representing the business process for the tracking of fruit trays. The transition probabilities have been rounded to three decimals and transitions with a probability below the noise threshold ϵ are not shown, except for the defect state.

every tray is equipped with two RFID tags, an assignment between RFID portal and fruit tray can be established. In addition to this mechanism, the identification point for incoming trays considers information provided by the business layer in order to account for stationary tags and weak RFID read events.

The next stage where a localization is performed is the inbound to a specific storage hall. This issue is solved by providing each storage hall entry with floor tags that will be read by the RFID equipped forklift. This idea was realized mainly for economic reasons: Due to the large number of storage halls, this solution is cheaper than providing every storage hall entry with RFID readers and antennae. Since it suffices to know in which storage hall a fruit tray is located, the localization in this case can be achieved by the assignment between fruit trays and the storage hall identifier represented by the floor tags. Later stages of the process perform simpler localization tasks, because trays are singulated for the process steps of sorting and packing.

This use case demonstrates that also the mapping of rather complex business processes onto Hidden Markov Models is possible, since loops and alternating paths can be considered by means of appropriate transition probabilities. The HMM is a flexible framework to consider arbitrary process chains and provides the mathematical tools to efficiently evaluate the likelihood of certain events. The results from recent integration tests of the four portal identification points for incoming trays provided satisfactory results. With the roll-out of the system in autumn 2010, long term test results will become available.

6

Concluding remarks

According to economic studies, the worldwide RFID market is steadily growing. Due to the advances in reader and tag technology combined with sophisticated standardization, tagging of objects on item level has become reality in various applications. With the increasing possibilities, also the requirements to RFID systems have become versatile. Whereas first systems where used to inventory pallets of goods in simple business processes, today's systems keep track of the location of every single item in complex and multi-layered supply chains. RFID systems operating under the EPCGlobal standard in the UHF band allow for a simultaneous identification of hundreds of items up to a distance of several meters. Besides the obvious advantages, there are also challenges and drawbacks. The large read range introduces unwanted behaviour whenever information about the exact location of goods is required. Since RFID readers in general report all tags that are currently present in the RF-field, sophisticated mechanisms for localization and tracking goods are necessary.

For this reason, the topic on localization in RFID has been extensively investigated in the last few years. Systems for localization have to cope with different challenges that result from physical phenomena on the channel, but also from the used hardware and software. As pointed out during the discussion in this work, RFID read events are subject to random fluctuations in several ways. In order to derive information about the location of RFID tags in practical applications, a model is required which takes into account this random behaviour. In contrast to previous approaches, this work combines a flexible model for RFID read events and sensor signals with information from the underlying business process. This combination results in a probabilistic framework that can be used to perform reliable localization and tracking of goods in a business process. Considering the requirements of practical systems, the localization of goods can most often be reduced to the simpler task of classification. Hence, Hidden Markov Models are used which allows for a classification of observation sequences in a flexible way. The goal is not to compute the position or trajectory of tags with respect to a coordinate system, but rather to provide information required by the business process, such as an assignment between goods and packaging units. In order to allow for a reliable localization, the presented work also includes a probabilistic process model that is able to provide quantitative information about every item in the process. Moreover, this kind of model can also deal with uncertainties and imperfections of business processes due to its flexible nature. The results presented throughout the evalution of the approach indicate that this idea is capable of fulfilling practical requirements. The developed models have been implemented in currently ongoing projects in the industry and provided satisfactory results during the preliminary integration tests.

6.1 Future work

A potential drawback of the presented model for read events and sensor signals is the requirement for a sufficiently high number of read events to perform a sophisticated and reliable classification. Especially when the number of tags in the RF-field is high, the anti-collission mechanism specified in the EPCGlobal standard introduces random behaviour and the number of read events for each tag decreases significantly. For this reason, future research could include a distinct modeling of the anti-collision mechanism. This model could be used together with a stochastic pathloss model in order to interpolate the obtained RSSI patterns. Alternatively, other methods for increasing the number of read events by performing adaptive changes in the configuration could be investigated.

Another topic for possible future research is the fusion of different classification results. As briefly discussed in chapter 4, there exist several approaches to combine different classification results in a consistent Bayesian formulation. For this purpose, a way of assigning posterior probabilities to a series of read events needs to be found. This idea is of particular interest, when more than two classifiers are combined.

A
Appendix

The appendix provides the numerical values for the model parameters of the HMM representing the exemplary business in in chapter 4.

A.1 Numerical values for the exemplary business process model

The Hidden Markov Model $\lambda_{\mathcal{B}}$ of the exemplary business process in figure 4.15 is characterized by the following parameters:

$$\boldsymbol{\pi} = \begin{bmatrix} 0 & 0.99 & 0.0025 & 0.0025 & 0.0025 & 0.0025 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1.0 & 0 & 0 & 0 & 0 & 0 \\ 1.0 \cdot 10^{-6} & 0.00667 & 0.74 & 0.24 & 0.00667 & 0.00667 \\ 1.0 \cdot 10^{-6} & 0.00667 & 0.00667 & 0.74 & 0.24 & 0.00667 \\ 1.0 \cdot 10^{-6} & 0.00667 & 0.00667 & 0.00667 & 0.74 & 0.24 \\ 1.0 \cdot 10^{-6} & 0.0025 & 0.0025 & 0.0025 & 0.0025 & 0.99 \\ 1.0 \cdot 10^{-6} & 0.0025 & 0.0025 & 0.0025 & 0.0025 & 0.99 \end{bmatrix}$$

$$\mathbf{B} = \mathbf{I}_{6\times 6}$$

where $\mathbf{I}$ denotes the identity matrix.

Bibliography

[1] G. Aires da Silva and D. R. Ferreira. Applying Hidden Markov Models to Process Mining. In *Sistemas e Tecnologias de Informação: Actas da 4a. Conferência Ibérica de Sistemas e Tecnologias de Informação*, pages 207–210. AISTI/FEUP/UPF, 2009.

[2] C. Alippi, D. Cogliati, and G. Vanini. A Statistical Approach to Localize passive RFIDs. In *Proceedings of the IEEE International Symposium on Circuits and Systems*, page 4 pp., 2006.

[3] L. E. Baum and J. A. Egon. An Inequality with Applications to Statistical Estimation for Probabilistic Functions of a Markov Process and to a Model for Ecology. In *Bull. Armer. Meterolo. Soc.*, volume 73, pages 360–363, 1967.

[4] L. E. Baum, T. Petrie, G. Soules, and N. Weiss. A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. In *The Annals of Mathematical Statistics*, volume 41, No. 1, pages 164–171, Apr. 1970.

[5] L. E. Baum and G. R. Sell. Growth Functions for Transformations on Manifolds. In *Pac. J. Math*, volume 27, No. 2, pages 211–227, 1968.

[6] M. J. Beal and Z. Ghahramani. Variational Bayesian Learning of Directed Graphical Models with Hidden Variables. *Bayesian Analysis*, 1:793–832, 2006.

[7] A. Bekkali, H. Sanson, and M. Matsumoto. RFID Indoor Positioning Based on Probabilistic RFID Map and Kalman Filtering. In *Proceedings of the Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications*, page 21 pp., Oct. 2007.

[8] T. H. Davenport. *Process innovation: Reengineering Work through Information Technology*. Harvard Business School Press, Boston, MA, USA, 1993.

[9] C. de Souza. Hidden Markov Models in C#. http://www.codeproject.com/Articles/69647/Hidden-Markov-Models-in-Csharp.aspx, March 2010. Visited: 29th September, 2010.

[10] EPCglobal Inc. Class 1 Generation 2 UHF RFID Protocol for Communication at 860 Mhz-960 Mhz, Version 1.0.9, May 2008.

[11] T. Fawcett. An Introduction to ROC Analysis. *Pattern Recognition Letters*, 27(8):861 – 874, 2006.

[12] M. Figueiredo and A. Jain. Unsupervised Learning of Finite Mixture Models. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 24-3, pages 381 –396, Mar. 2002.

[13] K. Finkenzeller. *RFID Handbook: Radio-Frequency Identification Fundamentals and Applications*. Wiley, New York, 2000.

[14] D. Hähnel, W. Burgard, D. Fox, K. Fishkin, and M. Philipose. Mapping and Localization with RFID Technology. In *Proceedings of the IEEE International Conference on Robotics and Automation*, volume 1, pages 1015 – 1020, April 2004.

[15] J. Hightower, R. Wantand, and G. Boriello. SpotON: An Indoor 3-D Location Sensing Technology Based on RF Signal Strength. Technical report, University of Washington, Department of Computer Science and Engineering, 2000.

[16] X. Huang, R. Janaswamy, and A. Ganz. Scout: Outdoor Localization Using Active RFID Technology. In *Proceedings of the 3rd International Conference on Broadband Communications, Networks and Systems*, pages 1–10, Oct. 2006.

[17] D. Joho, C. Plagemann, and W. Burgard. Modeling RFID Signal Strength and Tag Detection for Localization and Mapping. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 3160–3165, May 2009.

[18] P. Kamol, S. Nikolaidis, R. Ueda, and T. Arai. RFID Based Object Localization System Using Ceiling Cameras with Particle Filter. In *Future Generation Communication and Networking*, volume 2, pages 37 –42, Dec. 2007.

[19] M. Khan and V. Antiwal. Location Estimation Technique Using Extended 3-D LANDMARC Algorithm for Passive RFID Tag. In *IEEE International Advance Computing Conference*, pages 249 –253, 6-7 2009.

[20] R. K. L. Ko, S. S. G. Lee, and E. W. Lee. Business Process Management (BPM) Standards: A Survey. In *Business Process Management Journal*, volume 15(3), March 2009.

[21] Lawrence R. Rabiner. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. In *Proceedings of the IEEE*, volume 77, pages 257–286, Feb. 1989.

[22] A. F. Molisch. *Wireless Communications*. Wiley-IEEE Press, 2005.

[23] L. Ni, Y. Liu, Y. C. Lau, and A. Patil. LANDMARC: Indoor Location Sensing Using Active RFID. In *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, pages 407 – 415, 2003.

[24] A. Rozinat, M. Veloso, and W. M. P. van der Aalst. Using Hidden Markov Models to Evaluate the Quality of Discovered Process Models. Extended Version. Technical report, BPM Center Report, 2008.

[25] D. Ruta and B. Gabrys. An Overview of Classifier Fusion Methods. In *Computing and Information Systems*, volume 7, pages 1 – 10, 2000.

[26] T. Sanpechuda and L. Kovavisaruch. A Review of RFID Localization: Applications and Techniques. In *Proceedings of the 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, volume 2, pages 769 –772, May 2008.

[27] A.-W. W. Scheer. *ARIS - Business Process Frameworks*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1998.

[28] G. A. F. Seber. *Multivariate Observations*. John Wiley and Sons, Hoboken, NJ., 1984.

[29] J. Shibata, T. Yairi, H. Kanazaki, Y. Shirasaka, and K. Machida. Probabilistic Sensor Models for Multiple Objects Localization Problem. In *SICE-ICASE, International Joint Conference*, pages 581 –586, Oct. 2006.

[30] A. Tanenbaum. *Computer Networks*. Prentice Hall Professional Technical Reference, 2002.

[31] A. J. Viterbi. Error Bounds for Convolutional Codes and an Asymptocially Optimal Decoding Algorithm. In *IEEE Trans. Informat. Theory*, volume IT-13, pages 260–269, Apr. 1967.

[32] P. Vorst, S. Schneegans, B. Yang, and A. Zell. Self-Localization with RFID snapshots in Densely Tagged Environments. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1353 –1358, Sept. 2008.

[33] M. Weske. *Business Process Management - Concepts*. Springer Verlag, 2007.

[34] Y. Zhang, M. G. Amin, and S. Kaushik. Localization and Tracking of Passive RFID Tags Based on Direction Estimation. In *International Journal of Antennas and Propagation*, volume 2007, Article ID 17426, 9 pages, 2007.