PROBABILISTIC MODELING IN RFID SYSTEMS

Michael Goller

November 2013

Graz University of Technology

Probabilistic Modeling in RFID Systems

Doctoral Thesis

at

Graz University of Technology

submitted by

Michael Goller

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

20th November 2013

Thesis Supervisors:
Prof. Dr. Georg Brasseur
Prof. Dr. Petar Popovski
Dr. Markus Brandner

“Doubt is uncomfortable,
but certainty is ridiculous.”

Voltaire

Abstract

Passive UHF RFID is considered an enabler technology for highly flexible, yet fully transparent logistic applications. The possibility to identify and track individual items from the manufacturing stage to the end customer provides unique possibilities and a deep insight to logistic processes and supply chains. Although RFID technology has experienced significant advances during the last couple of years, there are still open issues related to data inaccuracies caused by the underlying physical principles of operation. In particular, RFID systems face challenges in terms of missing and false positive observations. Both effects lead to inconsistent data in the backend system.

This thesis addresses the problem of increasing the robustness of RFID systems in logistic applications. We develop a framework that specifically deals with false positive observations and missing tags on different abstraction levels. The presented models cover three key aspects of the RFID system: First, we present a model for RFID enabled supply chains on the process level which allows us to efficiently filter noisy observations. The model provides a well-defined way to integrate prior information about the typical behavior of goods in a supply chain and considers the spatio-temporal correlation among RFID observations. Second, we develop a probabilistic readpoint and signal model which is used in a classification approach to improve the detection performance of the readpoint. Third, we present an information fusion approach for the purpose of RFID tag localization by means of a hybrid RFID and computer vision system.

The encountered variability and the fact that RFID is an interdisciplinary field with heterogeneous system components require a consideration on an adequate abstraction level. Throughout this thesis, we show that probabilistic methods are well suited to tackle the challenges in RFID systems. Using a simulation engine and comprehensive empirical datasets from different RFID deployments, we provide an in-depth evaluation of the presented approaches and filter mechanisms and validate the underlying modeling assumptions.

Kurzfassung

Die passive UHF RFID Technologie gilt als Meilenstein für logistische Anwendungen im Hinblick auf Flexibilität und Transparenz. Die Möglichkeit, Einzelteile entlang der Lieferkette eindeutig zu identifizieren und zu verfolgen bietet eine Vielzahl von Möglichkeiten und eine Fülle von Informationen über logistische Prozesse.

Vom technologischen Standpunkt her haben sich passive UHF RFID Systeme in den letzten Jahren sehr stark weiterentwickelt. Die Genauigkeit der generierten Daten und Informationen stellt aber nach wie vor den limitierenden Faktor für den Einsatz von RFID dar. Die inhärenten Ungenauigkeiten werden durch die zugrundeliegenden physikalischen Prinzipien der Wellenausbreitung verursacht und äußern sich durch eine limitierte Leserate und das Auftreten von Falsch-Positiv Observationen.

Diese Arbeit stellt verschiedene Ansätze zur Erhöhung der Datengenauigkeit in RFID Systemen vor. Im Rahmen einer probabilistischen Formulierung wird ein Systemmodell präsentiert welches es erlaubt, RFID Daten auf Prozessebene effizient zu bewerten und zu filtern. Durch das Modell kann a-priori Wissen über den logistischen Prozess hinsichtlich des typischen Verhaltens und dem Auftreten korrelierter Vorgänge berücksichtigt werden. Ergänzend dazu wird ein Signal-Modell vorgestellt mit dem RFID Observationen auf der Ebene einzelner Lesepunkte evaluiert und klassifiziert werden können. Um den steigenden Anforderungen moderner Anwendungen gerecht zu werden, wird zudem die Fusion von RFID-Systemen mit alternativen Sensormodalitäten untersucht. Dabei steht die Lokalisierung von RFID Transpondern in praxisnahen Szenarien im Fokus.

RFID Systeme sind durch eine Vielzahl von heterogenen Komponenten gekennzeichnet. Gemeinsam mit dem Umstand, dass RFID orientierte Prozesse mitunter eine hohe Variabilität aufweisen, macht dies eine Betrachtung auf einem angemessen hohen Abstraktionsgrad unabdingbar. In dieser Dissertation wird gezeigt, dass probabilistische Methoden ein geeignetes Werkzeug sind, um die im Kontext von RFID Systemen gestellten Herausforderungen zu lösen. Zu diesem Zweck werden die getroffenen Modellannahmen und die Leistungsfähigkeit der vorgestellten Methoden auf Basis von Simulationen und umfassenden empirischen Daten evaluiert.

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

Contents

Contents **i**

List of Figures **iii**

1 Introduction **1**
1.1 Passive UHF RFID . . . 2
1.2 Motivation and Contributions . . . 5
1.2.1 Publications . . . 7

2 Related Work **11**

3 RFID System Modeling **17**
3.1 Probabilistic Process Model . . . 20
3.1.1 Process Level Localization . . . 24
3.1.2 Model Calibration . . . 26
3.1.3 System Monitoring . . . 28
3.1.4 Case Studies . . . 31
3.2 Spatio-temporal Item Correlation . . . 34
3.2.1 Modeling spatio-temporal correlation . . . 35
3.2.2 Evaluating RFID observations . . . 37
3.2.3 Experimental Evaluation . . . 40
3.3 Summary . . . 43

4 RFID Readpoint Modeling **45**
4.1 Feature Attributes . . . 46
4.1.1 Signal model . . . 51
4.1.2 Case Study: Conveyor Belt Application . . . 56
4.2 Cooperative RFID Readpoints . . . 60
4.2.1 System model . . . 62
4.2.2 Experimental Evaluation . . . 65
4.3 Summary . . . 69

5 PRISE - Probabilistic RFID Simulation Engine **73**
5.1 Related Simulation Tools . . . 74
5.2 Simulator Concepts . . . 75
5.3 Models and Implementation . . . 77
5.3.1 RFID Tag . . . 77
5.3.2 RFID Reader and Antenna . . . 78
5.3.3 Detection model . . . 79
5.3.4 Sensors . . . 80
5.4 Limitations . . . 80
5.5 Experimental validation . . . 81
5.5.1 Scenario 1 - Stationary tags . . . 82
5.5.2 Scenario 2 - Conveyor Belt Setup . . . 83
5.5.3 Scenario 3 - EAS . . . 86
5.6 Summary . . . 89

6 Information Fusion **91**
6.1 Fusion of CV and RFID . . . 92
6.2 RFID Subsystem . . . 93
6.3 Blob Detection and Tracking . . . 95
6.4 Data Association . . . 96
6.5 Calibration . . . 100
6.6 Case Study: EAS . . . 102
6.7 Summary . . . 106

7 Conclusion **107**
7.1 Summary of Contributions . . . 108
7.2 Outlook . . . 110

A Appendix **111**
A.1 PRISE - Example configuration file . . . 111

Bibliography **115**

List of Figures

1.1 RFID tags for different applications 2
1.2 RFID: Operating principle 3
1.3 Ideal *vs.* real interrogation zone 4

3.1 Exemplary supply chain 18
3.2 Exemplary retail store 19
3.3 Motion model 21
3.4 Exemplary process model: Retail store 22
3.5 Time dependent transition probabilities 23
3.6 RFID System model 24
3.7 System monitoring: Estimated detection probability 30
3.8 Case study 1: Process model for an automated sorting process 32
3.9 Case study 1: Empirical dwell time distributions 32
3.10 Case study 2: Process model for a fashion supply chain 33
3.11 Case study 2: Empirical dwell time distributions 34
3.12 Spatio-temporal item correlation: Exemplary state-space trajectories 36
3.13 Spatio-temporal item correlation: Schematic simulation setup 41
3.14 Spatio-temporal item correlation: Performance evaluation 41
3.15 Spatio-temporal item correlation: Model sensitivity 42

4.1 Experimental evaluation: Read rate 47
4.2 Exemplary low level tag responses 49
4.3 Tag velocity estimation 50
4.4 Experimental evaluation: Tag velocity distribution 51
4.5 Signal model for low level tag responses 52
4.6 Feature-space representation for low level tag responses 54
4.7 Feature space clustering 55
4.8 Commissioning and verification using RFID 56
4.9 Experimental setup: Conveyor belt 58

4.10 Experimental setup: Classification results 60
4.11 Probabilistic tag detection model: Graphical representation 63
4.12 Readpoint correlation 64
4.13 Experimental evaluation: RFID system architecture 65
4.14 Conveyor belt setup 66
4.15 Correlation on feature level 67
4.16 Cooperative readpoints: Detection probability 69
4.17 Detection probability: Temporal evolution 70
4.18 Combined detection probability: Temporal evaluation 70

5.1 Simulator Architecture 76
5.2 Tag antenna radiation pattern 77
5.3 Antenna radiation pattern 79
5.4 Experimental setup: Stationary tags 82
5.5 Experimental evaluation: Read rate 83
5.6 Experimental evaluation: Number of read events in conveyor belt setup 84
5.7 Experimental Evaluation: Simulated RSSI response 85
5.8 Experimental setup: EAS 86
5.9 Experimental evaluation: Number of read events in EAS setup 87
5.10 Experimental evaluation: Simulated *vs.* empirical RSSI responses 88

6.1 Information fusion: Block diagram 93
6.2 RFID sensor model 94
6.3 Blob detection and tracking 95
6.4 Blob detection and tracking: Camera view 97
6.5 Data association 98
6.6 Data association problem: Filtered association probabilities 99
6.7 Calibration setup 100
6.8 RFID sensor model: Contour plot 101
6.9 RFID sensor model 102
6.10 EAS scenario: Floor plan 104
6.11 EAS scenario: Camera view 104

Acknowledgements

A PhD thesis? Sounds like a lot of work and ups and downs! Enough to be sure that it can never be successful without the help and support from different people. For this reason, I'm glad to devote this special page to everyone who walked with me along the way.

My gratitude goes to my supervisor Prof. Brasseur, for his inputs and support during my PhD studies. Special thanks also to Prof. Popovski for being the co-examiner of my thesis and his valuable inputs and comments. Especially, I would like to express my sincerest gratitude to Markus: You have been my scientific mentor through the last couple of years and I cannot do more than to thank you for your personal support, the countless discussions, hints and your willingness to tackle tight schedules. I would also like to thank my colleagues Tom and Deini for numerous discussions, the obligatory coffee breaks and the right sense of humor. Special thanks goes to Axel and Christoph for their computer vision expertise and the collaboration during the *P-Detect* project.

This work wouldn't have been possible without the support from my colleagues at Enso Detego. Cordially thanks goes to Alex for the unconditional support since my first job interview. On the professional level, you have provided me with numerous (interesting!) challenges and relentlessly helped me to master them. On the personal level, I want to thank you for your support and guidance over the last five years. Sincere thanks to Thomas and Mario: You never failed to support and challenge me on a professional level and our discussions opened up my mind for alternative viewpoints. But even more importantly, I want to thank both of you for your friendship during our studies and beyond. Special thanks also to Robert for proofreading and his comments to this thesis.

Deep heartfelt thanks to my parents Elisabeth and Anton, who did everything they could to let me go my way, from the very first steps up to this moment. It was your support and effort which made my studies possible in the first place. I would also like to express my gratitude to my sisters Katharina, Maria-Lucia and Barbara for all the wonderful moments and your support. Especially, my thanks go to my nephews David and Philipp: For the things I could teach you, and for the things I learned from you.

Almost exactly two years ago, my life began to change – all because of you: Martina, I can't think of words that can nearly express how thankful I am for your love, your understanding and your support. You are my counterpart and backup, you keep me on course. My sincerest thanks for you being there.

Michael Goller
Graz, Austria, October 2013

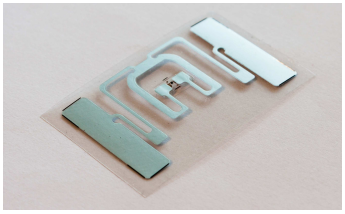
1 Introduction

Some years ago, Radio Frequency Identification (RFID) was imagined to fully replace the ubiquitous barcode technology in logistics and retail. The direct comparison between these two technologies suggests that RFID is superior in almost every aspect - except for the price. The advantages are indeed manifold: RFID chips have a considerably larger memory, can be identified in a bulk, do not require a direct line of sight to the reader, and optionally offer security features like authentication or selective memory access. However, the RFID market still hopes for the long desired increase in sales and applications like the frequently cited *Future Store*[1] are still elusive.

The RFID industry has developed various different types of RFID systems and standards tailored to the requirements of different applications [50]. However, the different system types consist of the same basic building blocks: The *reader* (*interrogator*) is designed for the communication with (low-cost) *transponders* (*tags*) by means of electromagnetic waves or inductive coupling. For high-volume applications, the Ultra High Frequency (UHF) band from 860 – 960 MHz has become the operating frequency of choice in current RFID deployments. The EPCglobal, Class-1 Generation-2 (Gen-2) standard [48] laid an important cornerstone and can be considered as the enabler for RFID systems operating on the item-level. There exists a vast variety of different types of RFID tags designed for specific product categories and environmental conditions. Three

[1] http://www.future-store.org/

examples for RFID tags are shown in Figure 1.1: An adhesive multi-purpose label, a paper tag, and a so called *hard tag* with a robust plastic housing.



(a) Adhesive label

(b) Paper tag

(c) Hard tag

Figure 1.1: RFID tags for different applications: (a) shows an adhesive multi-purpose label, (b) a standard paper tag for fashion applications, and (c) a so called hard tag with a robust plastic housing.

So when RFID technology is superior to barcodes in almost every aspect, what are the factors that still prevent a mass deployment? As stated above, the key factor is the price in comparison to well established, barcode based identification systems. In addition, the nature of passive RFID systems exhibits some peculiarities and technological challenges that make a plug-and-play deployment difficult and require special attention.

1.1 Passive UHF RFID

The unique feature that tags are remotely powered by the reader enables high volume applications since tags can be built as small and ubiquitous devices without an integrated power supply. Instead, tags draw their operating power solely from the field emitted by the reader. The operating principle of a passive RFID system is shown in Figure 1.2. The electromagnetic field emitted by the reader is used to power the tag and to transmit data and commands. Whereas the communication from reader to tag is based on amplitude modulation, information from tag to reader is

transmitted by means of a load modulation.

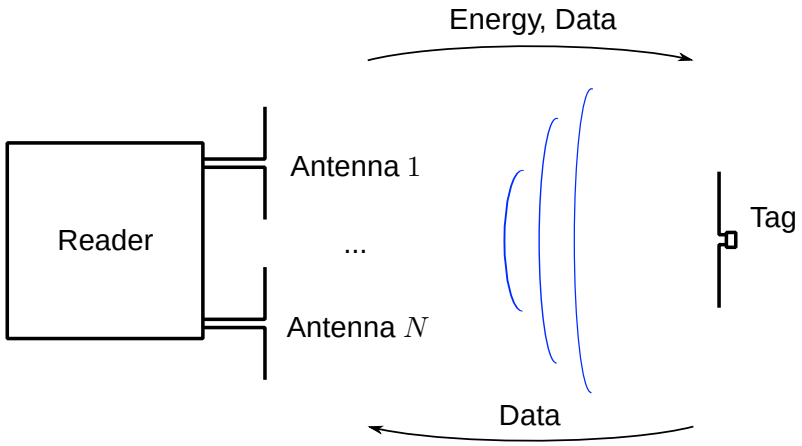
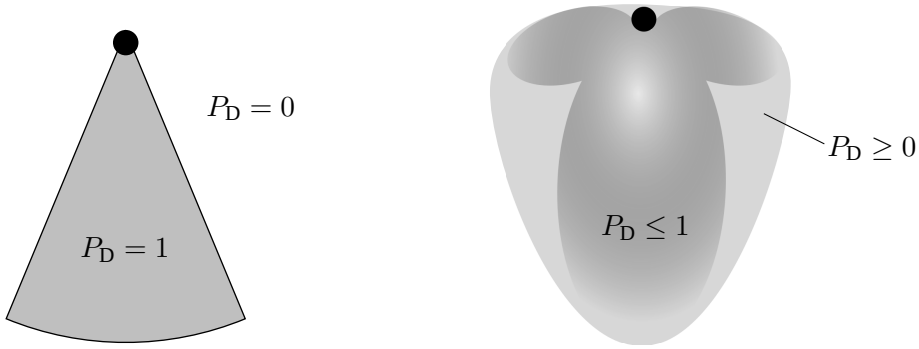


Figure 1.2: Operating principle of passive RFID systems: RFID tags are powered by the reader-field. Data and commands from reader to tag are transmitted by means of amplitude modulation, whereas data transmission from the tag to the reader is based on a load modulation.

Advances in reader and tag technology over the last years have resulted in considerable read-ranges of up to $15\,\mathrm{m}$ under ideal conditions. These advances are best demonstrated by comparing the sensitivity of the very first Gen-2 transponder chips with today's state-of-the-art: Impinj[2], a leading manufacturer of transponder chips, reader chips, and readers released the Monza/ID transponder chip with a sensitivity of $-11.5\,\mathrm{dBm}$ as one of the first Gen-2 compliant transponder chips in 2005 [85]. The latest chip, Monza 5, exhibits a nominal sensitivity of $-17.8\,\mathrm{dBm}$ [86], providing a significant increase in readrange and orientation insensitivity.

However, for typical applications, readrange is not the limiting factor. Due to the fact that communication is based on the electromagnetic wave propagation, it is not possible to define a precise *interrogation zone* for RFID readers and antennas. For an idealized system, such as shown in Figure 1.3(a), the interrogation zone is a specific volume of known dimensions in which a present tag population can be identified with a probability of $P_\mathrm{D} = 1$. In this context, P_D is referred to as *detection probability*. Outside the interrogation zone, the detection probability is $P_\mathrm{D} = 0$. However, practical systems as shown in Figure 1.3(b) do not exhibit an ideal detection probability inside the interrogation zone and the detection probability outside the interrogation zone does not vanish. This has two immediate consequences: First, a detection probability $P_\mathrm{D} \leq 1$ causes that a certain percentage of tags in the interrogation zone is not identified, leading to *false negatives* or *missing observations*. Second, the nonzero probability to detect tags outside the desired

[2] http://www.impinj.com/



(a) Ideal **(b)** Real

Figure 1.3: Ideal *vs.* real interrogation zone. Whereas an ideal RFID system has a well defined interrogation zone where the detection probability is $P_\mathrm{D} = 1$ (a), practical systems do not show this property and suffer from *false positive* and *false negative* (missing) observations (b).

interrogation zone causes so called *false positive* tag readings.

The factors that cause a non ideal detection probability inside the interrogation zone of an RFID system are manifold. First, the operating principle of passive RFID systems imposes challenging conditions in terms of the tag power supply and the communication link. The multipath channel characteristic leads to so called dead zones with insufficient energy to power the tag [13]. Second, RFID tags are required to be small and cheap and therefore have limited capabilities. This means that there is only a minimum functionality regarding power supply stabilization, stable data transmission, and anti-collision schemes. Third, the electromagnetic properties of tagged items can severely affect the identification performance. Especially, objects containing water or reflective materials introduce challenging conditions and result in severe performance degradation due to absorption or detuning phenomena [16]. Finally, the increasing item throughput imposes a limitation especially in case of moving RFID tags or readers. Although single tags can be identified at considerable velocity, increasing the number of tags also increases the required inventory time during which the tags must be present in the interrogation zone. Nikitin and Rao [152] provide a concise summary about the different impact factors in RFID systems. In terms of the missing tag problem, a careful choice of the transponder type, antenna design, and system setup is important. Typically, this involves extensive evaluations and tests during the deployment phase of an RFID system to find the optimal configuration for a given scenario.

The problem of false positive reads stems from the fact that antennas show a specific, environment dependent radiation pattern which contradicts the requirement for a well defined interrogation zone. Conductive materials with dimensions that are large compared to the wavelength reflect incident electromagnetic waves and lead to undefined interrogation zones in real-world deploy-

ments. This especially causes problems in environments where space is a scarce resource, as it is typically the case in warehouses and distribution centers.

1.2 Motivation and Contributions

The lack of a well defined interrogation zone and the non ideal detection probability have one major implication for RFID systems: Noisy data. The noise stems from the fact that a reader may fail to identify a tag inside the interrogation zone or unwanted reads from tags located outside the desired volume. Depending on the application, this has different, sometimes immediate consequences. Consider an RFID system for EAS (Electronic Article Surveillance) in a retail store. In this scenario, continuous false positives reads that trigger false alarms introduce considerable problems in terms of customer acceptance and are a serious argument against the use of RFID.

On the long run, the problem of noisy data could be partly solved by means of an accurate localization of RFID tags. Knowing the exact position of a tag enables the system to decide whether it is inside the defined interrogation zone or not. However, a precise localization of RFID tags in practical applications is difficult to achieve within the limits of the narrowband EPCglobal standard [13]. Similarly, the problem of missing tags is difficult to solve, even though readers and tags are steadily improving in performance and sensitivity.

From a practical point of view, this means that there are considerable open issues for the mass deployment of RFID systems which cannot be solely tackled by advances in reader and tag technology. Since practical applications have stringent performance requirements in terms of detection and false positive probability, additional concepts are required which help to improve the data accuracy. This motivates the use of top-down-concepts and model based approaches which are employed successfully in other fields facing similar challenges. For model based approaches, the nature of RFID systems introduces additional complexity due to the integration of several heterogeneous components. Consequently, the research questions addressed by this thesis are:

- *Can the problem of noisy data in RFID systems be mitigated by means of a top-down modeling approach?*
- *Is a probabilistic framework suitable to deal with the particular properties and heterogeneous components in an RFID system?*

The idea of a top-down modeling approach stood at the beginning of this thesis project in 2010. The goal was to establish a framework for RFID systems to improve the data quality by considering the following aspects:

- *Business process information:* The flow of products in typical applications follows certain rules. Usually, items are moving from the manufacturing stage over a network of distributors to a retail shop and finally to the end customer through a network referred to as *supply chain*. The information about the typical goods flow can be considered in a model to evaluate individual item trajectories. Another aspect on the business layer is the fact that items are usually aggregated in packaging units for easier transportation and handling. This introduces additional information by means of spatial and temporal item relationships. Logistic processes hence provide prior information about the *typical* behavior of items in the supply chain which can be integrated in a high-level RFID system model.

- *RFID system properties:* In order to account for the inherent observation noise, the modeling framework should specifically consider the properties of RFID systems in terms of false negative and false positive observations. A detailed discussion about the RFID system model that takes into account the business process information together with the particular RFID system properties is provided in Chapter 3. Due to the fact that empirical data from large scale practical applications is scarce, a simulation framework has been developed to verify and to evaluate the discussed modeling approaches. In contrast to existing simulators, this framework provides a combination of a high-level supply chain simulation and the generation of low-level RFID observations. The simulator and the underlying concepts are described in detail in Chapter 5.

- *RFID readpoint and low-level features:* From a high-level perspective, the detection of RFID tags is a binary event – either a tag is detected by a readpoint or not. RFID systems, however, provide more detailed information for every read event, which opens up several possibilities to evaluate and assess tag read events. The available information is discussed and integrated in a readpoint and signal model which is discussed in Chapter 4. The signal model forms the basis for a classification approach which can be employed at the readpoint level to identify and suppress false positive observations.

- *Information fusion:* The combination of different information sources and sensor modalities is a common approach in various technical systems that deal with noisy data. Chapter 6 gives an overview over RFID related sensor fusion approaches in the recent literature and discusses computer vision systems as an attractive fusion candidate. In particular, we develop a localization system that combines the information from a mono camera with RFID read events to determine the location of individual items in a scene.

The discussed aspects have been extensively studied during the course of this thesis project. Consequently, this lead to a number of publications which are outlined below for detailed reference.

1.2.1 Publications

The joint consideration of business process information and RFID system properties was achieved in a flexible discrete time state-space-model which was first presented at the IEEE RFID conference in 2011:

> [68] M. Goller and M. Brandner. Increasing the robustness of RFID systems using a probabilistic business process model. Poster Presentation, IEEE RFID, 2011

The ideas behind this approach were continuously refined and extended by a classification mechanism that jointly considers business process information and low-level RFID data to identify false positive and false negative observations. In order to allow for a convenient evaluation of different ideas and modeling concepts, we started the development of a Probabilistic RFID Simulation Engine – *PRISE*. In contrast to other simulation frameworks that focus on particular problems such as the UHF channel or communication protocol, PRISE allows for the simulation of large scale RFID systems and specifically considers the business process layer by simulating high-level item trajectories. The system model and the simulation engine were continuously optimized and evaluated by means of data from real-world RFID installations in different applications. This work resulted in a publication which was presented at the RFID Technology and Applications conference in 2011:

> [69] M. Goller and M. Brandner. Probabilistic modeling of RFID business processes. In *Proc. IEEE RFID-TA*, pages 432–436, 2011

The developed framework was further improved by integrating a continuous time motion model and an RFID sensor model. Using empirical data from active deployments, a comprehensive validation of the modeling assumptions was performed. The resulting state-space model and an algorithm for process-level localization were presented at the IEEE International Conference on Wireless Information Technology and Systems 2012:

> [71] M. Goller and M. Brandner. Process-level localization of RFID tags using probabilistic models. In *Proc. IEEE ICWITS*, 2012

The second part addressed in the course of this thesis project deals with RFID systems modeling on the readpoint level. In this context, the goal was to develop a probabilistic framework with the capability to evaluate and classify read events with off-the-shelf RFID hardware. Based on an analysis of different feature attributes and previous work on readpoint modeling [67] with HMMs (Hidden Markov Models), an experimental study about different strategies to evaluate low-level signal features was performed and published at the IEEE RFID conference 2011:

> [66] M. Goller and M. Brandner. Experimental evaluation of RFID gate concepts. In *Proc. IEEE RFID*, pages 26–31, 2011

The main result of this work was that the performance of individual readpoints can be significantly improved by means of a classification scheme using an appropriate system model. The readpoint model based on HMMs was gradually improved to provide a flexible framework for different feature attributes. The model has been evaluated in typical applications, both in laboratory environments and practical deployments. The results of the experimental evaluation were published at the European Conference on Smart Objects, Systems and Technologies 2012 and awarded as Best Paper:

> [70] M. Goller and M. Brandner. Evaluation of feature attributes for an RFID conveyor belt application using probabilistic models. In *Proc. SmartSystech*, 2012

The consequent next step for a general readpoint model was to investigate on more advanced sensor modalities. For logistic applications in automated environments, standard proximity sensors provide a robust and deterministic information due to well defined boundary conditions such as known object dimensions and a fixed movement speed. However, these sensor modalities are not applicable to other use cases like warehouse portals or the aforementioned EAS scenario. For this reason, computer vision systems were investigated as complementary sensor modality with the capability to provide accurate location and tracking information.

The increasing requirements in terms of detection performance and item throughput in state-of-the art logistic applications motivated further research to improve the detection performance of practical RFID deployments. With the background of a probabilistic framework, this research was focusing on diversity concepts in general, and spatial readpoint diversity in particular. Whereas the idea of spatial diversity is widely developed in wireless communication systems, it requires special attention in the context of RFID. Backed up by extensive empirical data, a model for cooperative RFID readpoints was developed, leading to a Journal publication submitted to the IEEE Transactions on Instrumentation and Measurement in 2013:

> [72] M. Goller, M. Brandner, and G. Brasseur. A system model for cooperative RFID readpoints. *Submitted to the IEEE Trans. Instrumentation and Measurement*, 2013

The problem of noisy observations due to the discussed challenges in RFID systems have been addressed by probabilistic modeling concepts on different abstractions layers. Viewing RFID systems in a probabilistic context enabled us to effectively combine different information sources and properly address the challenges introduced by the heterogeneous system components. Due to the

close cooperation with an industrial partner, the concepts developed in this thesis were successfully applied to a number of RFID projects in logistic and retail scenarios. Besides the insights gained from practical implementations, the individual projects provided us with experimental data and served as a benchmark under realistic operating conditions.

2

Related Work

RFID systems comprise a variety of heterogeneous components and therefore face diverse questions, reaching from hardware-related topics such as circuit and chip design up to software development, database design and supply chain considerations. For this reason, the literature on RFID related topics is also very heterogeneous and characterized by interdisciplinary approaches. This chapter gives an overview over the activities in the different research communities related to this thesis during the last years. In particular, we provide a system perspective view of the different approaches, discuss their advantages and limitations and highlight the open issues.

Traditionally, RFID system design and deployment is characterized by extensive experimental evaluations and feasibility studies. Due to various physical effects, the most important questions for an RFID system are still the ones about the effective detection probability, effective read range, and maximum achievable item throughput. Several authors have presented experimental studies solely devoted to these performance metrics and their dependence on environmental factors [16, 27, 39, 44, 60, 95, 103]. The reasons why this type of analysis is so popular and therefore often encountered are twofold: First, the detection probability and read range are the most basic performance metrics that are directly visible to the end-user. They can be determined without specialized measurement equipment which reduces the complexity of the experimental setup. Second, the detection probability is by far the most important and concise quantity that

describes the performance of an RFID deployment.

Although studies on this level provide a concise picture of the performance limitations, they do not explain the involved physical effects that lead to a decreased system performance. For this reason, there are various experimental studies that provide important insights on a lower abstraction level. Nikitin *et al.* have published substantial work in terms of tag performance, sensitivity and impedance measurements and the impact of different communication protocol parameters [148, 149, 153, 154, 156]. Similarly, other authors have investigated on these aspects with a focus on tag and antenna performance [29, 35, 65, 75, 189]. Since RFID systems are primarily limited in the forward link [116], tag related measurements have received a lot of attention in the community [31, 124, 129, 133, 172, 195]. Closely related to this issue is the wireless UHF channel. The wireless multipath channel in the context of RFID systems has been studied extensively by different researches and considerable effort has been devoted to channel modeling and characterization [13, 80, 131, 139].

Whereas the authors in the literature discuss evaluation scenarios on different abstraction levels and analyze different impact factors, the basic conclusions are very similar: RFID systems offer a fast and reliable method to identify tagged objects under idealized conditions but suffer from performance degradations in typical industrial environments. In particular, multipath propagation and the presence of metal or water lead to a significant decrease of the detection probability. Consequently, the challenges encountered in RFID deployments are approached on a broad basis. From a high-level perspective, these approaches can be grouped into hardware related concepts, diversity schemes, and methods targeting the protocol layer. Since the RFID tag and the forward link are the primary limitation, there are continuous advances in transponder chip technology and tag antenna design [128, 160]. There exists a huge diversity of RFID tags for different environmental requirements, for example tags that are especially designed for the application to metallic objects. The increase in sensitivity, improvements in orientation insensitivity and a certain tolerance to particular item materials are the key enabler for RFID driven supply chains in various industries. However, the physical limitations that lead to a non ideal detection performance are still an important reason that impedes the mass deployment of UHF RFID.

To overcome this issue on another system layer, RFID systems have adopted several diversity concepts [183], similar to other wireless communication systems. The most intuitive approach in this respect makes use of several RFID antennas to maximize the coverage in an interrogation zone and to increase the detection probability. Commonly, this approach is implemented by means of a simple time-multiplexing scheme for the different antennas. More advanced approaches, such as presented by Angerer et. al [11, 12] include a signal processing block as countermeasure to the undesirable properties of the wireless multipath channel. The diversity introduced by multiple

RFID tags attached to the same object is another intuitive method to increase the system performance [183], which, however, is rarely employed due to the increasing costs. Instead, temporal diversity by means of repeated inventory sessions is a frequently found approach that has been extensively studied in the literature [57, 88, 161]. On the downside, temporal diversity automatically increases the overall inventory duration and is therefore a limiting factor in terms of system throughput.

The multiple-access problem is an inherent characteristic of communication systems in general, and RFID systems in particular due to the unknown number of tags in the interrogation zone. In RFID systems, this problem is solved by an anti-collision scheme which is used to single out particular tag responses. Therefore, the anti-collision procedure is a key parameter in terms of throughput and efficiency [32, 105, 106]. Several authors have presented concepts for enhanced anti-collision schemes [115, 177, 191, 199] which show considerable throughput improvements compared to the current standard. In addition to the features specified in the EPCglobal standard [48] such as different sessions and tag muting, advanced anti-collision schemes also improve the detection probability for a given RFID deployment by providing additional time-slots for tags that are particularly hard to identify.

Although the effective detection probability is the most important performance metric for an RFID system, the problem of false positive observations is of equal importance to the data accuracy in the backend system. Characteristic scenarios that emphasize the consequences of false positive observations include the checkout-desk in a retail store or an RFID driven article surveillance. In both cases, false positive observations not only impair the data accuracy, but also have immediate impact on the end customer in terms of erroneous bills or the characteristic false alarm. Among other performance metrics, this issue is investigated in a real-life environment by Al-Kassab *et al.* [5]. The main conclusion is that the value of an RFID system is mainly determined by the achieved data accuracy. For this reason, considerable research has been conducted targeting the false positive read problem. The different approaches range from a geometric tag localization over classification schemes to high-level process modeling and the integration of different information sources.

The most popular approach to eliminate false positives is concerned with the geometric localization of RFID tags. The idea behind this approach is simple and intuitive: If a reader is able to determine the exact tag location, it is possible to decide whether the tag is located in the defined interrogation zone or not. However, tag localization in passive UHF RFID systems is a challenging task and the accuracy is impaired by several different aspects. First, the EPCglobal standard is merely designed for the identification of objects and does not explicitly address localization issues. This is reflected by the limited bandwidth available in UHF RFID systems.

Second, the harsh channel conditions in terms of multipath propagation make the localization increasingly difficult, especially in indoor environments. The localization approaches discussed in the literature are manifold and heterogeneous which makes a direct comparison difficult. A characterization can be carried out by means of the employed signal features [136]. The first type of localization methods employs the RSS (Received Signal Strength) or RSSI (Received Signal Strength indicator) from a backscattered tag response [78, 94, 173] to estimate the distance between reader and tag. In addition, several approaches using advanced tracking strategies [117, 145] have been presented by different authors. Localization using RFID has also become a popular topic in the robotic community where reference tags are frequently used to localize a mobile robot [33, 38, 185]. Although the problem in this context is to determine the location of a reader, the principles are very similar to tag localization systems. The second type of methods to determine the location of individual RFID tags is based on the phase information [150], for example with a synthetic antenna aperture [137].

Both RSS and phase-based systems have their limitations in terms of accuracy due to the difficult environmental conditions. For this reason, extended approaches including information fusion concepts and the adoption of computer vision systems [87, 143] have been developed. These systems provide a more accurate localization of individual objects due to the employed camera system. In addition, systems that work beyond the EPCglobal standard specifications (e.g., ultra-wideband systems) have been developed that show a considerable improvement in localization accuracy [123]. The capabilities of the different systems discussed in the literature are usually demonstrated in well defined setups with a low number of tags. Real-world applications which usually exhibit a considerably larger tag population are however still an open issue with respect to an accurate localization. To provide the required accuracy, future RFID standards will need to specify a considerably larger bandwidth [13] compared to current state-of-the-art systems.

Besides the geometric tag localization, there are alternative solutions to the false positive read problem. On the hardware layer, there exists a variety of options, including shielding structures or special antenna designs for particular applications [134, 135]. Whereas shielding structures are typically bulky, inflexible, and costly to install, approaches based on a specialized antenna design have the disadvantage that they are usually tailored to a very specific setup and require a lot of engineering and tuning during the deployment phase.

Somewhere in between hardware related approaches and a full-blown localization are filtering mechanisms that try to assess the characteristic of tag responses to decide whether a tag was deliberately identified or not. In this context, different filtering and cleaning mechanisms have been proposed as described in a recent survey about data processing in RFID systems by Aggarwal and Han [2]. Pioneering work in terms of data cleaning has been published by Jeffrey *et al.* [90].

The authors use an adaptive temporal windowing technique to compensate for missing reads and suppress false positive tag observations. Besides the filtering capability, this work provided substantial novelty by defining a major and minor detection region for RFID readers and by interpreting the tag detection process in a probabilistic context. Consequently, different researchers have build up on this idea [76, 101, 111, 181, 198] and provided extensions and improvements. Temporal filtering mechanisms provide an effective suppression of false positive observations in specific applications but are not directly applicable to some scenarios such as the commissioning verification of individual packaging units. However, sliding window techniques form the baseline for many RFID systems to increase the data accuracy. As extension to the discussed sliding window schemes, an approach that integrates the RSS information provided by RFID readers has been presented by Keller *et al.* [99, 100]. The authors try to extract scalar metrics that describe a particular tag response and perform a classification to distinguish between stationary and moving tags in a dock-door scenario. Their work is especially noteworthy since it is based on a large scale dataset from a real-world application, in contrast to most other publications in this field.

As a complementary idea to the discussed low-level approaches, several researchers have developed systems that use high-level prior information [98, 118]. Since RFID systems are usually employed in a certain supply chain structure [22], it is possible to consider the typical behavior of tagged items in terms of a model. The basic idea to suppress false positives is then to use the learned model and evaluate a given observation. Besides the false positive tag problem, it has been shown that an adequate supply chain model can also be used for security related aspects such as cloned tag detection and anti counterfeiting [119] as well as supply chain visualization and process analysis [83, 84]. Beyond the actual supply chain structure, other authors propose methods to integrate additional prior information such as the spatio-temporal correlation among individual read events [30, 146, 167]. These systems have in common that they rely on a probabilistic formulation and are hence suited for the integration in a general system model.

Although there are significant advances in terms of RFID system performance, the problem introduced by noisy RFID data is still an open issue. Since the original goal of RFID deployments is to provide information about a particular environment, the inherent noise impairs the information quality and therefore has impact on the value of the RFID system itself. The challenges in terms of missing and false positive observations can be addressed on different abstraction levels and there exists a huge variety of different approaches. These approaches are very heterogeneous and it is difficult to integrate them in a generic, flexible framework. This thesis aims to fill this gap by means of a generic modeling framework that is capable to handle different abstraction levels and integrate heterogeneous information sources.

3

RFID System Modeling

The peculiar properties of passive UHF RFID systems result in potentially noisy data in the back-end system. This noise is mainly caused by the lack of a well defined interrogation zone and a variety of physical effects that impair the system performance. This chapter presents top-down modeling concepts designed to tackle the problem of noisy RFID observations. First, we develop an RFID system model which is able to describe the dynamics in typical supply chain structures and specifically considers the non-ideal RFID system properties in terms of a sensor model. The model forms the basis for a process level localization mechanism and additionally allows us to evaluate the performance of an RFID system at runtime. Second, we present an approach to utilize the information about spatio-temporal item relationships as complement to the RFID system model. In particular, we develop a co-occurrence model to represent the relationship between individual items in a supply chain.

The increasing requirements for logistic systems in terms of time, costs and flexibility have led to highly complex process structures and the need for powerful backend IT-systems to provide a fast and accurate information exchange. From a high-level perspective, the task of logistic systems seems simple: Establish a goods flow to supply a specific customer demand in a timely fashion with a certain quality, low costs and the correct information. However, this usually involves long and complex network structures with different manufacturers, a distribution network and retail

stores. A simplified example of a three tier supply chain is shown in Figure 3.1. Goods are transported from different manufacturers (tier 1) over a distribution network (tier 2) to a set of retail stores (tier 3), where they are finally sold to the end-customer.

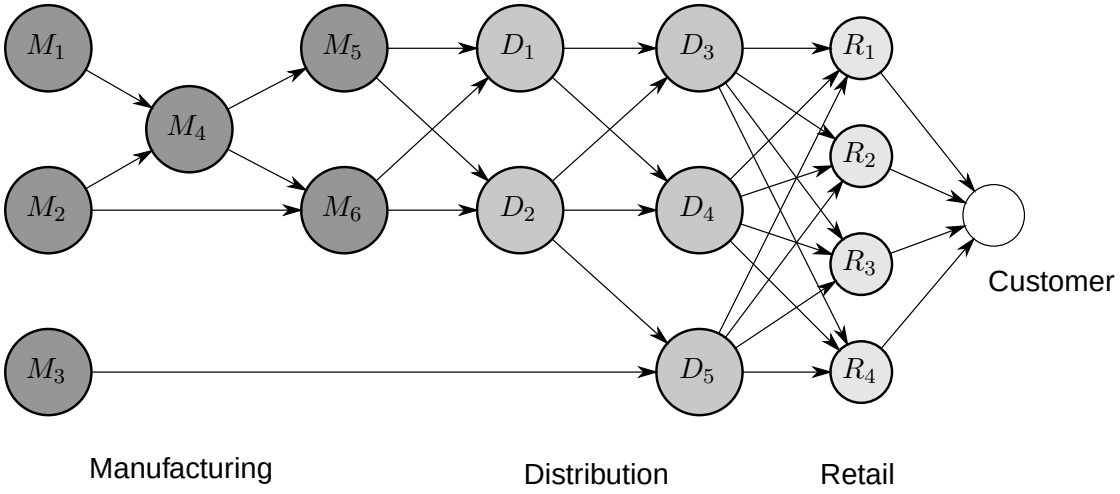


Figure 3.1: Exemplary supply chain: Goods move along typical trajectories from manufacturing over a distribution network to retail stores and are finally sold to a customer. A typical trajectory for an article could hence span from the two manufacturers M_2 and M_6 over the distribution partners D_1 and D_4, respectively, to the retail shop R_1 where it is bought by a customer.

Depending on the type of product or product mix, a supply chain can exhibit different levels of complexity. Typically, every stage involves a set of subordinate process steps, giving rise to a self-similar structure [179], as exemplary shown for a retail store in Figure 3.2. A typical retail store features a goods inbound where delivered articles are received. The goods are then placed in a storage area and are finally presented on the sales floor. Regardless of the supply chain structure and the number of hierarchy levels, inventory and order management require an accurate information flow and identification of goods and transactions. To provide information about the current location of individual items, an RFID system hence needs to automatically identify goods as soon as they enter a specific process stage. For this purpose, RFID hardware is installed at critical locations in the supply chain. In this context, we define an *RFID readpoint* $\mathcal{R}$ as a set of RFID readers, antennas and additional sensors that belong to a specific process step.

Depending on the hierarchy level at which RFID readpoints are installed, the location information can be provided at a certain discretization level. The basic functional principle, however, is independent of the hierarchy level: As soon as an item enters the interrogation zone of a particular readpoint, the RFID reader identifies the item and triggers the location update. At this level, we

can hence define a *read event* (or *RFID observation*) as a binary variable

$$z_i = \begin{cases} 1 & \text{if tag is detected by readpoint } \mathcal{R}_i \\ 0 & \text{if tag is not detected by readpoint } \mathcal{R}_i \end{cases} . \tag{3.1}$$

Illustrated by means of the exemplary retail store in Figure 3.2, an item I_1 is received and identified at the goods inbound which causes a location update to retail store R_1. Similarly, the location inside the shop is updated with every new observation. Independent of the hierarchy level in the

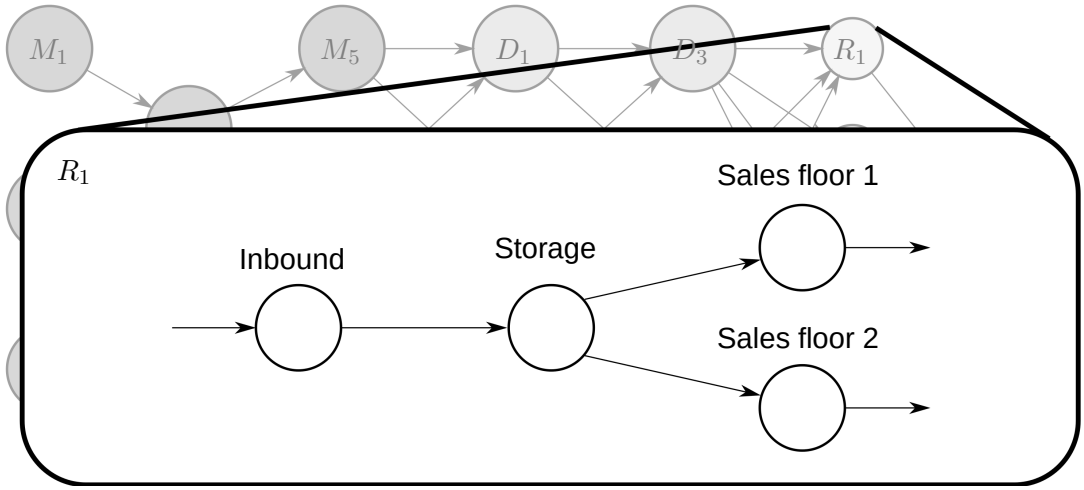


Figure 3.2: Zoom into an exemplary retail store with different process steps and locations. After receiving a set of goods in the inbound area, they are typically transferred to the storage location before being presented on the sales floor

supply chain, every stage has a set of characteristic properties:

- *Physical location:* At the highest hierarchy level, this can denote the location (address) of a manufacturing plant or distribution center. At a lower level, the location can for example refer to a specific region in a warehouse or store such as *Storage area* or *Sales floor 2*.
- *Dwell time:* Defines how long items remain in a particular supply chain stage. Depending on the process, we can identify two possibilities that characterize the dwell time: Either, it is determined by process requirements (e.g., storage or cooling period, duration of certain manufacturing steps or transportation), or external factors such as customer behavior. The first case especially covers process steps in manufacturing and distribution where the dwell time typically is determined by the duration of certain manufacturing steps. For the second case, the dwell time provides substantial information about the process for the maintainer

of a supply chain. For example, the information about how long certain products are presented on a sales floor before being sold is valuable information for management tasks and marketing.

- *Previous process steps:* Defines a set of previous stages in the supply chain from which items can be transferred to the current stage.
- *Subsequent process steps:* In analogy, this is a set of subsequent process steps to which items can be transferred from the current stage.

The discussed properties define the topological structure and the dynamic behavior of a supply chain. As an additional characteristic, logistic applications usually utilize different sorts of packaging units for item aggregation and transport. Typical examples are pallets, cardboard boxes or packs for hanging garments. Regardless of the physical properties, packaging units introduce logical item units and give rise to a certain structure and correlation in the goods flow. This is a noteworthy aspect that provides prior information about a particular supply chain.

Following a top-down approach, the remainder of this chapter presents a probabilistic process model to describe the characteristics of modern RFID driven supply chains. The model characterizes the goods flow in terms of a motion model and specifically addresses the RFID system properties by means of a sensor model. Besides the general idea, we discuss a mechanism for process-level localization and present a method to estimate the model parameters. The characteristic of spatio-temporal item relationships in a supply chain are utilized in two different ways. First, we describe an online, high-level system monitoring mechanism to estimate the detection probability of individual readpoints at runtime. Second, we utilize the spatio-temporal relationships between individual items in a co-occurrence model to increase accuracy of RFID observations.

3.1 Probabilistic Process Model

The most noticeable aspect that characterizes a supply chain is goods flow along the different stages and process steps. This flow is difficult to describe in a deterministic way due to the high dynamics and the wide variety of different products and processing steps. For this reason, an adequate approach to accurately consider these characteristics is the use of a stochastic model. In particular, Continuous Time Markov Chains (CTMCs) [169] allow for a concise description and provide a powerful mathematical framework. A CTMC

$$\lambda = (\boldsymbol{\pi}, \mathbf{P}, \boldsymbol{\mu}) \tag{3.2}$$

comprises the initial state distribution $\boldsymbol{\pi}$ and describes the dynamic behavior in terms of the transition probability matrix $\mathbf{P}$ and the dwell time parameter $\boldsymbol{\mu}$. Similar to discrete time Markov chains, a CTMC is a directed graph consisting of N states and E edges representing transitions between the states, as shown in Figure 3.3. The state-space $\mathbf{X} = (X_1, ..., X_N)$ represents the different stages in the supply chain and the transition probability matrix entries p_{ij} describe the probability for a transition from X_i to X_j for a particular item. In contrast to discrete time Markov chains,

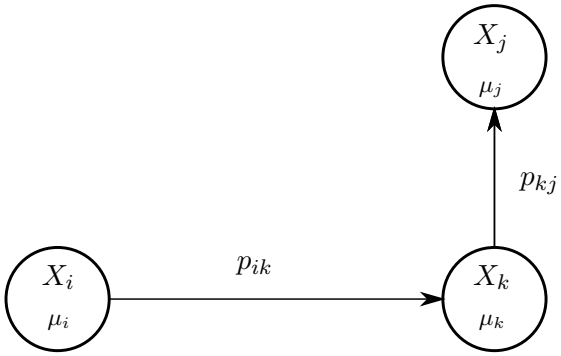


Figure 3.3: Continuous Time Markov Chain as motion model for an RFID enabled supply chain. Every stage in the supply chain corresponds to one discrete state X_i. The temporal behavior of the motion model is characterized by the transition probabilities p_{ij} and the mean dwell times μ_i in every state.

CTMCs explicitly model the duration in a particular state by means of a dwell time parameter T_i. The dwell time is defined as the time difference between leaving and entering a state

$$T_i = t^{(i)}_{\text{Leave}} - t^{(i)}_{\text{Enter}} \tag{3.3}$$

and can be characterized in terms of a probability distribution. For our purposes, we consider *time homogeneous* CTMCs, which means that the transition probabilities are independent of the absolute time

$$P(X(t) = j \,|\, X(s) = i) = P(X(t-s) = j \,|\, X(0) = i), \quad \forall\, s < t. \tag{3.4}$$

This property implies that the dwell times T_i follow an Exponential distribution

$$T_i \sim \text{Exp}(\mu_i) \tag{3.5}$$

since this is the only continuous distribution that exhibits the memoryless property [141]. Consequently, this results in the parameter $\boldsymbol{\mu}$ holding the mean dwell time for every state in the CTMC. Although the assumption of an exponentially distributed dwell time is potentially limiting the

applicability of this model, it provides a comfortable mathematical formalism and a reasonable approximation as we will show later by means of empirical data from two case studies in Section 3.1.4.

An alternative representation for a CTMC is the *generator matrix* $\mathbf{G}$ with individual entries

$$g_{ij} = \frac{p_{ij}}{\mu_i} \qquad g_{ii} = -\sum_{\substack{j=1 \\ j \neq i}}^{N} g_{ij}. \tag{3.6}$$

The generator matrix allows for an efficient computation of the time dependent transition probability function

$$p_{ij}(t) = P(X(t) = j \,|\, X(0) = i) = \mathrm{e}^{t\mathbf{G}} \equiv \sum_{n=0}^{\infty} \frac{(t\mathbf{G})^n}{n!} \tag{3.7}$$

which describes the temporal evolution of the transition probabilities. Using this framework, we can describe a wide variety of supply chain structures. For example, consider the retail store modeled as CTMC shown in Figure 3.4. The process is characterized by five states that describe the flow of goods from the inbound over a storage area and two sales floors to the end-customer. The time dependent transition probability function for a subset of states is shown in Figure 3.5. For

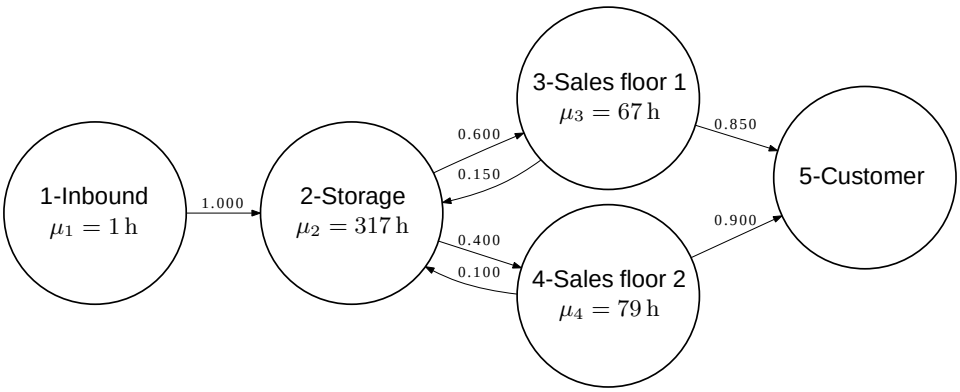


Figure 3.4: A retail store modeled as CTMC: Articles are received at the inbound and then transferred to a common storage area. From there, the articles are transferred to one of two sales floors where they are presented to the customer. The dwell time in every state is specified in terms of an exponential distribution with mean μ_i.

the interpretation of Figure 3.5, it is important to keep in mind that the transition probabilities are independent of the absolute time due to the assumption of time homogeneity. Characteristically, the probability to stay in a certain state decreases exponentially, as shown for the state *2-Storage*. In contrast, the transition to another state first increases and shows a dwell time dependent maximum.

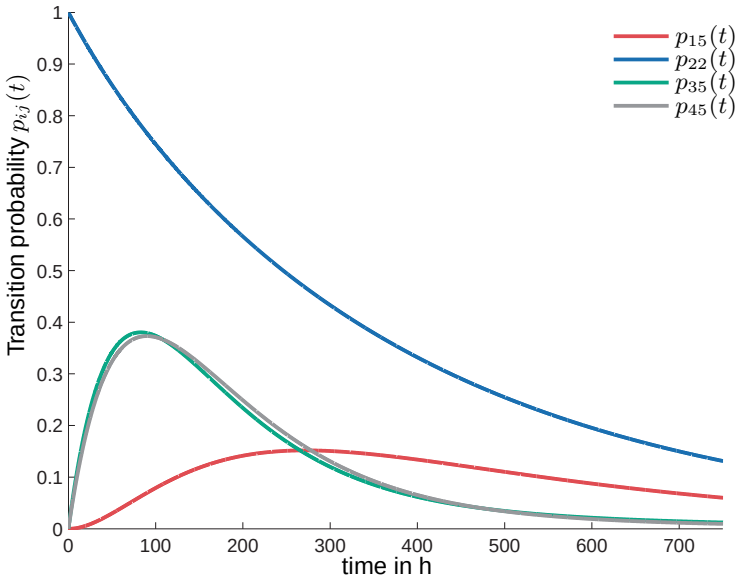


Figure 3.5: Time dependent transition probabilities for the process model shown in Figure 3.4. For the exemplary process, the probability that an item which was received at time $t = 0\,\mathrm{h}$ is sold to the customer has its maximum around $t = 271\,\mathrm{h}$ (see $p_{15}(t)$). Similarly, the transition probability from sales floor 1 to the customer has its maximum after $t = 83\,\mathrm{h}$.

In addition to the process dynamics covered by the CTMC, the RFID system model specifically considers the properties of RFID readpoints by means of a sensor model which accounts for the lack of a well defined interrogation zone and the non ideal detection probability. Specifically, the sensor model integrates the possibility for missing observations in terms of a detection probability $P_{\mathrm{D}}^{(i)} \leq 1$ and models false positive reads from tags outside the defined interrogation zone. In this context, the probability in state X_j to detect an unwanted tag currently being in state X_i is denoted as $P_{\mathrm{FA},i}^{(j)}$. The detection and false positive probabilities can be summarized in an $N \times N$ observation matrix $\mathbf{O}$, where

$$o_{ij} = P_{\mathrm{FA},i}^{(j)} \quad \text{and} \quad o_{ii} = P_{\mathrm{D}}^{(i)}. \tag{3.8}$$

The resulting state-space model as shown in Figure 3.6 consists of a CTMC as continuous time motion model and a sensor model reflecting the RFID system properties. The state-space model provides us with the following possibilities: First, we can use it as a generative data model to simulate high-level item trajectories in a wide range of possible supply chain scenarios. Second, the model can be used to robustly determine the actual item location in the supply chain based on potentially noisy read events. The inference mechanism for this high-level localization is explained in the next section.

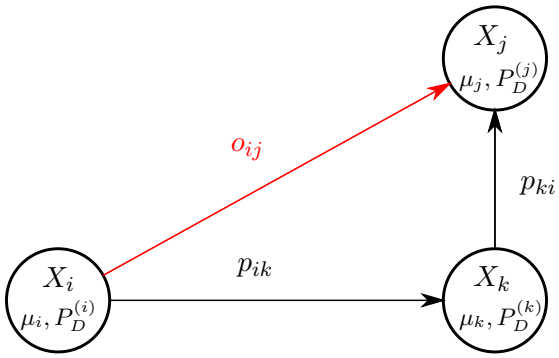


Figure 3.6: RFID system model, consisting of a CTMC and an RFID sensor model that reflects the individual readpoint properties. The sensor model specifically considers the detection probability and accounts for the possibility of false positive observations.

3.1.1 Process Level Localization

Estimating the most likely state of a system from noisy observations is commonly discussed in the literature as *filtering problem*. Given a series of observations $\mathbf{z} = z(t_1), ..., z(t_K)$ and the system model described in Equ. (3.2), the task is to compute the current state probability distribution for a particular item. In case of a discrete time, finite state-space setting, the Forward Algorithm [163] is an optimal solution for this problem. For discrete time steps k, the state probability distribution is

$$\mathbf{X}_k = \mathbf{O}_k \mathbf{P}^T \mathbf{X}_{k-1}, \quad \mathbf{X}_0 = \mathbf{O}_0 \boldsymbol{\pi} \tag{3.9}$$

where $\mathbf{O}_k$ denotes the diagonal observation matrix according to the sensor model and $\boldsymbol{\pi}$ is the initial state distribution. Equ. (3.9) can be interpreted as a recursive Bayes update to the state distribution as new observation data becomes available. The simplicity of this recursive solution stems from the first order Markov assumption, i.e., the assumption that given the present state, the past and the future are independent.

As an extension to the discrete time solution, we can consider the dwell time in every state by integrating the transition probability function $\mathbf{P}(t)$ which can be computed using Equ. (3.7) for every new observation. The state probability distribution then becomes

$$\mathbf{X}(t_2) = \mathbf{O}_k \mathbf{P}^T(t) \mathbf{X}(t_1), \quad t_2 > t_1. \tag{3.10}$$

As soon as an RFID observation occurs, the process model can be employed to update the current item location. Depending on the motion and sensor model, the observation triggers a state transition or is identified as a false positive.

The graphical representation of the system model in Figure 3.6 allows for an intuitive explanation of the filtering process described in Equ. (3.10). If we assume that a particular item is in state X_i at time $t = 0$, we can consider the probability for a valid state transition given an RFID observation z_j as defined in Equ. (3.1). In particular, the probability for a valid state transition can be formulated as

$$P(t;\, \text{Transition} \,|\, z_j) = \frac{P(t;\, z_j \,|\, \text{Transition}) P(\text{Transition})}{P(z_j)} = \frac{o_{jj} p_{ij}(t)}{P(z_j)}. \tag{3.11}$$

In Equ. (3.11), the conditional probability $P(t;\, z_j \,|\, \text{Transition}) = o_{jj}$ denotes the readpoint detection probability given that a transition to the state X_j has occurred. The transition probability, $P(\text{Transition})$ can be directly computed from the CTMC motion model and equals $p_{ij}(t)$. The posterior probability for observing a valid state transition can therefore be computed using the RFID sensor model, and the transition probability according to the CTMC motion model. Similarly, the probability for the observation to be a false positive can be expressed as

$$P(t;\, \text{False positive} \,|\, z_j) = \frac{P(t;\, z_j \,|\, \text{False positive}) P(\text{False positive})}{P(z_j)} = \frac{o_{ij} p_{ii}(t)}{P(z_j)}. \tag{3.12}$$

This reflects the possibility that an item in state X_i is identified by a readpoint in X_j. In general, there exist several states X_m that cause false positive observations in X_j. With the last observation in X_i, this requires that the transition from X_i to X_m is not detected and hence, the probability for a false positive in X_j becomes

$$P(t;\, \text{False positive} \,|\, z_j) = \frac{o_{ij} p_{ii}(t) + \sum\limits_{\substack{m=1 \\ m \neq j}}^{N} o_{mj} p_{im}(t) \left(1 - o_{mm}\right)}{P(z_j)}. \tag{3.13}$$

With a sufficiently high detection probability $o_{mm} \approx 1$ and a low false positive probability $o_{mj} \ll 1$, the summation term can be neglected which turns Equ (3.12) into

$$P(t;\, \text{False positive} \,|\, z_j) \approx \frac{o_{ij} p_{ii}(t)}{P(z_j)}. \tag{3.14}$$

The denominator in Equ. (3.11) – (3.14) corresponds to the overall probability that an observation z_j occurs and follows from the law of total probability

$$P(z_j) = o_{jj} p_{ij}(t) + o_{ij} p_{ii}(t). \tag{3.15}$$

To identify the most likely cause of an RFID observation, Equ. (3.11) and (3.14) can be combined in a decision rule

$$\gamma(t) = \frac{P(t;\ \text{Transition} \mid z_j)}{P(t;\ \text{False positive} \mid z_j)} = \frac{o_{jj} p_{ij}(t)}{o_{ij} p_{ii}(t)} \gtrless 1. \tag{3.16}$$

Using this decision rule, a readpoint can effectively identify false positive read events and suppress the corresponding state transition. An equivalent interpretation is that the RFID system model can be used to estimate the location of an item in the supply chain at a given point in time. An observed read event z_j for an item in state X_i will only be accepted by the model given $\gamma(t) > 1$, thereby providing an effective mechanism to filter noisy RFID observations.

The RFID system model addresses the introductory requirements in terms of the process dynamics and the characteristics of noisy RFID observations. To enable the desired filtering capability, the model parameters need to be adjusted according to the particular supply chain and RFID system characteristics, which can be accomplished by means of an initial calibration phase.

3.1.2 Model Calibration

The RFID system model is uniquely described by means of the prior state distribution, the generator matrix and the RFID sensor model. In order to enable the filtering mechanism described above, the model parameters need to accurately reflect the properties of the underlying supply chain and the RFID system properties. These parameters can be estimated from empirical process data which needs to be recorded during the system setup. For the calibration, we assume that logged process data in the form of read events

$$z = [t,\ \text{ID},\ X_i] \tag{3.17}$$

with timestamp t, tag identifier ID and state X_i is available. This definition of a read event is closely related to the EPCIS (Electronic Product Code Information Service) standard [47].

3.1.2.1 Transition model – CTMC

The transition model is specified in terms of the prior state distribution and the generator matrix $\mathbf{G}$ or, equivalently, the dwell time vector $\boldsymbol{\mu}$ together with the transition probability matrix $\mathbf{P}$. Consequently, the estimation of the model parameters involves two parts: First, the average dwell time μ_i for each state needs to be estimated from logged read events. For this purpose, the empirical dwell times T_i for all tags that have entered and left a particular state X_i need to be computed using the definition in Equ. (3.3). With the assumption of an exponential distribution,

the Maximum-Likelihood estimate for the mean dwell time is then

$$\mu_i = \frac{1}{N} \sum_{i=1}^{N} T_i. \tag{3.18}$$

Second, the prior state distribution and the individual transition probabilities can be estimated from logged event data by applying the Forward-Backward algorithm [163]. For this purpose, the absolute timestamp can be dropped since the Forward-Backward algorithm only evaluates the occurrence of state transitions, regardless of the temporal behavior. In order to account for possible process noise, the estimated transition probabilities and the prior state distribution require modifications to consider arbitrary start states and transitions. Following the definition by Rozinat *et al.* [171], we define a noise level ϵ that reflects the probability for an item to perform a state transition apart from the estimated probabilities. Consequently, all zero transition probabilities are set to this noise level. To ensure that $\mathbf{P}$ is a row stochastic matrix (i.e., the rows sum up to 1), we modify the matrix elements

$$p_{ij} = \begin{cases} \epsilon & p_{ij} = 0 \\ p_{ij} - \epsilon \frac{|\{(ij') \,|\, p_{ij'}=0\}|}{|\{(ij') \,|\, p_{ij'}>\epsilon\}|} & p_{ij} > 0, \end{cases} \tag{3.19}$$

where $|\{(ij') \,|\, p_{ij'} = 0\}|$ is the number of zero elements in row i and $|\{(ij') \,|\, p_{ij'} > \epsilon\}|$ denotes the number of elements in row i greater than ϵ. The prior state probability vector $\boldsymbol{\pi}$ needs to be modified in the same manner to allow for a tag to start in an arbitrary state.

3.1.2.2 RFID sensor model

Similar to the motion model, calibrating the sensor model requires a ground-truth dataset from a set of items moving along the supply chain or a set of subordinate steps. By means of that, the supply chain dynamics (i.e., the transition probability matrix $\mathbf{P}$ and the dwell times) are determined and the sensor model parameters can be considered in an isolated manner. The parameter estimation can be performed based on empirical data holding individual RFID observations from one or more calibration runs.

The detection probability for a readpoint belonging to a state X_i can be estimated as

$$\hat{P}_D^{(i)} = \hat{o}_{ii} = \frac{\#\text{ items detected in } X_i}{\#\text{ items in } X_i}, \tag{3.20}$$

where the actual number of tags in X_i in the denominator is defined by the ground-truth. Similarly,

the empirical estimates for false positive observations can be obtained using

$$\hat{P}_{\mathrm{FA},i}^{(j)} = \hat{o}_{ij} = \frac{\#\text{ false positives from } X_i \text{ in } X_j}{\#\text{ items in } X_i}. \tag{3.21}$$

The described calibration method can be carried out during the deployment of individual readpoints along the supply chain. Typically, this deployment phase is characterized by an iterative evaluation and parameter optimization. The parameters include for example the reader transmit power, communication protocol settings as well as the number and placement of individual antennas. These parameters have direct influence on the system performance and consequently also on the resulting sensor model. The detection and false positive probabilities are highly dependent on the system configuration and, moreover, cannot be considered independently. This is an issue that should be kept in mind during the deployment of an RFID system. Since there is little to no theory about how the detection probability depends on individual system parameters, a recalibration after every parameter update is vital to accurately reflect the system properties.

3.1.3 System Monitoring

After the initial calibration during the deployment phase, being aware of the current system status is a desirable feature for practical deployments in order to allow for monitoring and management tasks. Besides the possibility to react to disturbances, monitoring capabilities form the basis for an adaptation of individual system parameters driven by environmental changes. Possible changes include varying geometric conditions or, as frequently found in fashion logistics, seasonal changes in the goods assortment which has a direct impact on the RFID system performance due to varying item properties.

In general, RFID system monitoring includes several aspects: Whereas it is relatively straightforward to detect if a particular RFID reader is in operation or that an antenna is physically connected, specific information about how well an RFID deployment operates is more difficult to obtain. In this context, the detection probability of individual readpoints is the most important metric, since it concisely describes the quality of operation. To obtain exact values for the detection probability, a ground-truth for a sufficiently large set of RFID tagged items is required as described in Section 3.1.2. However, this time consuming process cannot be carried out regularly during the actual system operation. For this reason, a mechanism that allows for an online estimation based on the available RFID observations is required. The RFID system model presented in this thesis can be employed to estimate the detection probability of an RFID system at runtime without the need for an exact ground-truth. The idea is to apply a bootstrapping mechanism that combines the information from individual readpoints to obtain an estimate for the ground-truth.

In order to establish a ground-truth estimate at runtime, we can exploit the existence of logical item units introduced by the aggregation of items in packaging units. In particular, we can combine the information about a packaging unit $\mathcal{P}_j$ from different readpoints to estimate the number of items M. For this purpose, the indicator variable Z as defined in Equ. (3.1) can be employed. Let $\mathbf{Z}^{(i)} = [Z_1^{(i)},\, Z_2^{(i)},\, ...,\, Z_K^{(i)}]^T$ denote the vector of indicator variables for tags $T_1,\, T_2,\, ...,\, T_K$ in a packaging unit $\mathcal{P}_j$, stemming from the observations in state X_i. Along the supply chain, the packaging unit gives rise to a set of indicator vectors $\mathbf{Z} = (\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, ..., \mathbf{Z}^{(N)})^T$ due to the detection in different states. An estimator for the true number of items M in the considered packaging unit is found by counting the number of tags that have been detected by at least k readpoints

$$\hat{M} = |\mathbf{Z} \geq k| = |\mathbf{Z}|, \tag{3.22}$$

where $|\cdot|$ denotes the set cardinality. The following example illustrates the estimation process: Consider a packaging unit $\mathcal{P}_1$ with $M = 5$ tags. The unit proceeds through four different states in a supply chain, giving rise to the indicator vectors

$$\mathbf{Z} = \begin{array}{c} \\ T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \end{array} \begin{array}{c} \begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \end{array}. \tag{3.23}$$

The individual length of the indicator vectors $\mathbf{Z}^{(i)}$ can vary among the different supply chain states depending on whether an item is detected or not. For the example in Equ. (3.23), we obtain $\hat{M} = 5$. The estimate for the detection probability of the readpoint in state X_i is then

$$\hat{P}_{\mathrm{D}}^{(i)} = \frac{M^{(i)}}{\hat{M}} = \frac{|\mathbf{Z}^{(i)}|}{|\mathbf{Z}|}, \quad i \in [1,\, N]. \tag{3.24}$$

The ground-truth and the estimate for the individual detection probabilities are updated and refined with every processed packaging unit. Assuming that packaging units are identified at discrete time instants t, the update can be computed recursively using

$$P_{\mathrm{D},t}^{(i)} = \frac{\hat{M}_{t-1}^{(i)} P_{\mathrm{D,\,t\text{-}1}}^{(i)} + M_t^{(i)}}{\hat{M}_{t-1}^{(i)} + \hat{M}_t}, \tag{3.25}$$

where $\hat{M}_{t-1}^{(i)}$ denotes the total number of tags identified in X_i up to time $t-1$. Since this is difficult

to maintain in practical applications, it is convenient to define a temporal window for which $\hat{M}_{t-1}^{(i)}$ is evaluated.

To illustrate the estimation technique, Figure 3.7 shows a simulation result for the readpoint detection probability over time. The simulated system consists of four readpoints with an initial detection probability $P_{\mathrm{D}}^{(i)} = 0.95$. The simulation includes 2000 packaging units with $M = 50$ items each. Between $t = 1000\,\mathrm{h}$ and $t = 1500\,\mathrm{h}$, we simulate a disturbing event (malfunction of readpoint in state X_1). The obtained estimate for the detection probability are smoothed by means of a running average filter with window size $T = 50$ h to provide a convenient view. The simula-

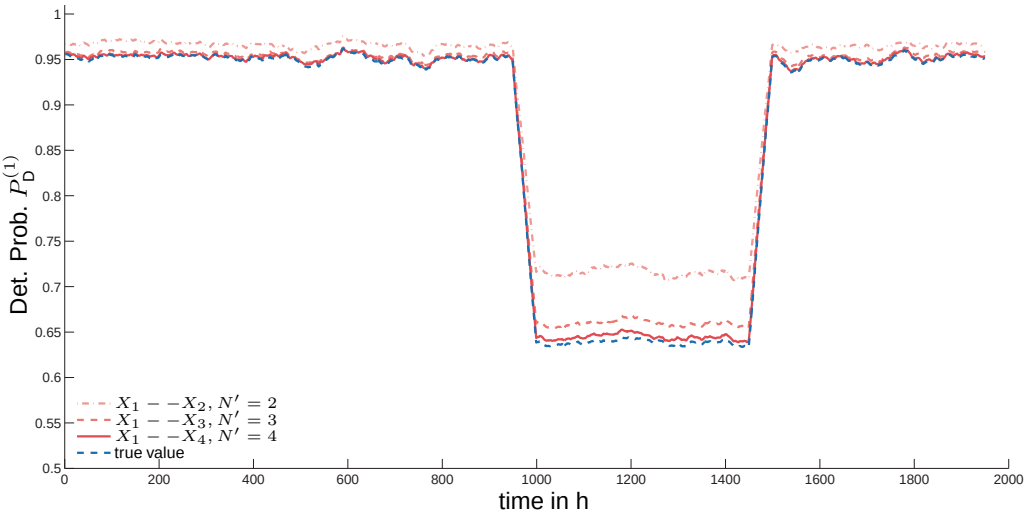


Figure 3.7: Simulation experiment: Estimated detection probability over time for readpoint $\mathcal{R}_1$ in an RFID system consisting of four readpoints. The actual detection probability in the undisturbed phase is $P_{\mathrm{D}}^{(1)} = 0.95$. Between $t = 1000\,\mathrm{h}$ and $t = 1500\,\mathrm{h}$, a disturbance leads to a considerable decrease in the detection performance. As more information (provided by the observations from subsequent states respectively readpoints) becomes available, the accuracy of the detection probability estimate increases and approaches the true value.

tion shows two interesting aspects. First, it highlights that the individual estimates are biased due to the fact that the ground-truth is obtained from uncertain information. As shown in Figure 3.7, this bias depends on the actual value of the detection probability. For the undisturbed case, the estimates closely approach the true value, whereas the bias increases during the disturbed phase. Second, the disturbed phase can be clearly identified also by means of the biased estimates, which provides the possibility to react on disturbing events in real-time. The described bootstrapping technique therefore provides an effective way to estimate the detection probability of individual readpoints at runtime. Ultimately, this forms a starting point for the adaptation of individual RFID

system parameters (e.g., transmit power) to react on small environmental changes.

3.1.4 Case Studies

The presented system model provides a flexible framework to describe RFID enabled supply chains and processes. In this section, we discuss two applications and show how the developed model can be applied to practical scenarios. By means of that we verify the model assumptions and provide important insights for the application and implementation of the RFID system model.

The first scenario is a logistic application for the processing of fruit trays. In this case, the RFID system is designed to track individual trays over the different states of an automated sorting process. The second scenario consists of a distribution center and a fashion retail store where the RFID system tracks individual articles from the shipping stage in the distribution center to the end-customer.

3.1.4.1 RFID enabled sorting plant

In this scenario, an RFID system is used to identify and track individual fruit trays through a production and sorting process. For this purpose, every tray is equipped with an RFID tag. By means of this case study, we investigate if the developed system model can be applied to automated environments. Besides developing the general model topology, we will analyze the empirical dwell times in every state to verify the assumption of an exponential distribution.

From the process perspective, three types of trays can be identified: The first type, which we will refer to as *raw trays* contain different fruit cultivars. The task of the sorting plant is basically to identify distinct cultivars and sizes and sort the fruits accordingly. The second type, called *receiving trays* are meant to be filled with one distinct cultivar. At this stage, these trays are transformed from *receiving trays* to *sorted trays*. Consequently, the trajectory of a particular tray through the process is as follows: A raw crate is transferred from the storage area to the *Sorter Feed* where it is identified by the RFID system. Subsequently, the tray is emptied and transported to a facility where the tray's physical condition is inspected. If the physical condition allows for further operation, the tray is transported as receiving tray to the *Sorter Exit* where it is again filled with one distinct cultivar. Finally, the tray is transported back to the storage area. Consequently, the process can be modeled by four distinct states as shown in Figure 3.8. The model parameters (i.e., the transition probabilities and dwell times) are estimated from empirical data over a period of three days. In particular, the dataset contains trajectories from 1.439 processed trays. The transition probability $p_{31} = 0.006$ (which corresponds to eight trays) is caused by the fact that these trays do not fulfill the inspection criteria and are therefore directly transferred back to the

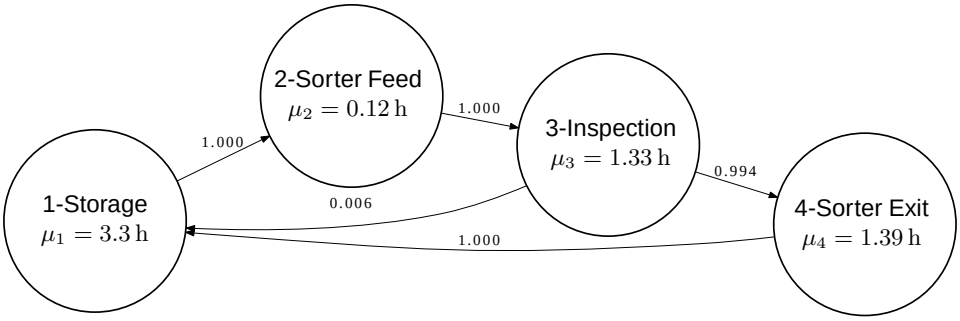
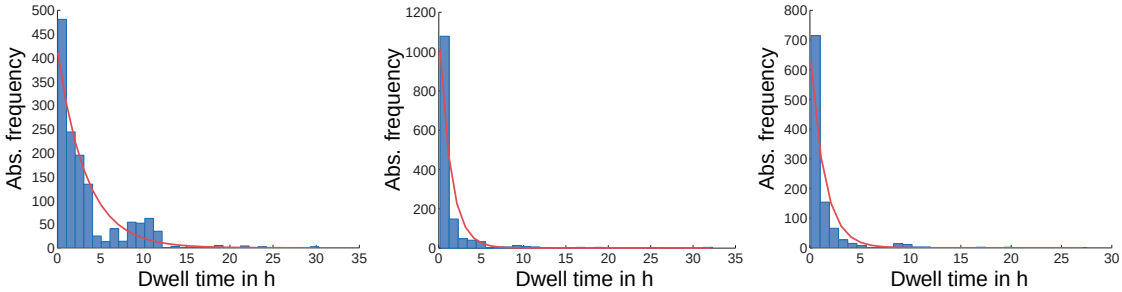


Figure 3.8: Model for an automated sorting process of fruit trays. Trays with different fruit cultivars are processed in an RFID enabled sorting plant over the stages Sorter Feed, Inspection and Sorter Exit. The transition probabilities and dwell times are estimated from selected process data containing trajectories of 1.439 trays over a period of three days. After the sorting process, the trays contain one distinct fruit cultivar and are moved back to the storage area. If an individual tray does not fulfill the inspection criteria, it is directly transferred back to the storage area, as indicated by the transition probability $p_{31} = 0.006$, which corresponds to a total of eight trays.

storage area for further processing. To verify the assumption of exponentially distributed dwell



(a) Storage, $\mu_1 = 3.3\,\mathrm{h}$ **(b)** Inspection, $\mu_2 = 1.3\,\mathrm{h}$ **(c)** Sorter Exit, $\mu_3 = 1.4\,\mathrm{h}$

Figure 3.9: Empirical dwell time distributions with fitted Exponentials for three different process stages in an RFID enabled sorting plant. Although the empirical data exhibits a certain amount of outliers, the exponential distribution provides a reasonable fit to describe the temporal behavior in the RFID system model.

times, we investigate on the empirical distributions as shown in Figure 3.9 together with their exponential fit.

The three histograms for the states Storage, Inspection and Sorter Exit indicate that the dwell time approximately follows an exponential distribution. The interpretation of the empirical data is as follows: The average tray stays in the storage area for $\mu_1 = 3.3\,\mathrm{h}$ whereas the inspection takes $\mu_2 = 1.3\,\mathrm{h}$ on average. Regardless of the outliers (for example in Figure 3.9(a), which can be explained by process interruptions such as a malfunction of the conveyor system), the exponential

distribution provides a reasonable fit to the empirical data.

3.1.4.2 Fashion supply chain

The second scenario that we consider for the evaluation of our RFID system model is a fashion supply chain, which is maybe the most prominent application of UHF RFID systems on item-level. For the evaluation of this scenario, we use an empirical dataset of 27.371 item trajectories from a pilot installation in a distribution center and a retail store. The resulting model representation is shown in Figure 3.10. The distribution center comprises a tag printing facility where

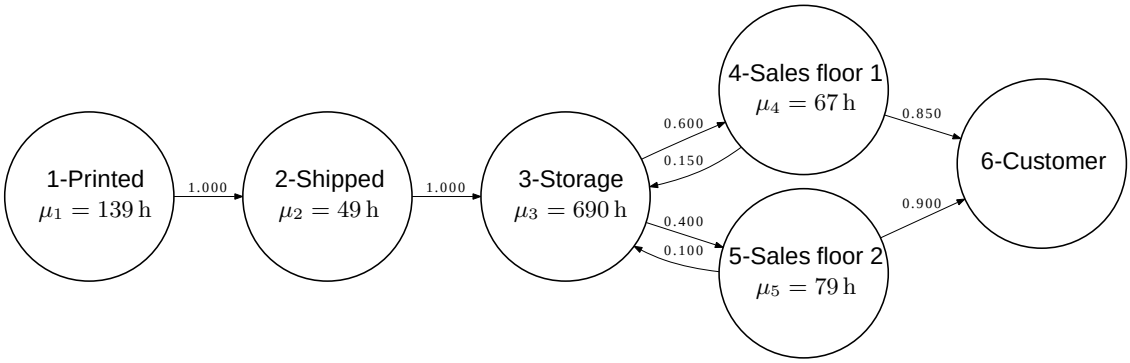
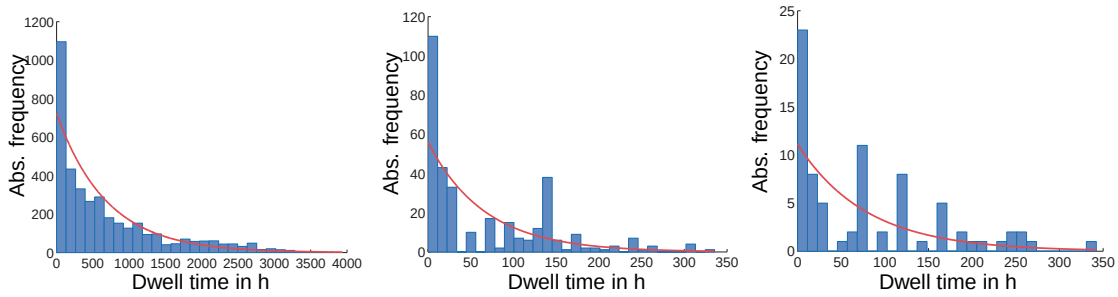


Figure 3.10: Model for a fashion supply chain, comprising a distribution center and a retail store. From the perspective of the RFID system, the supply chain comprises six distinct states. RFID tags are initialized in a printing facility and then attached to the individual articles. The articles are identified as soon as they are shipped to the retail store, causing a transition to the state *Shipped*. In the store, the articles are received and then placed in a *Storage* area. Finally, they are transported to one of two *Sales floors* where they are presented to the end-customer. The empirical data for this process stems from a pilot installation and consists of 27.371 item trajectories. The estimated dwell times are rounded to integer values.

RFID tags are initialized and attached to the individual articles. Subsequently, the articles are shipped to the store. The store features several locations which are covered by RFID readpoints. In particular, goods are received and placed in a *Storage* area, from where they are transported to one of two *Sales floors*, for women and men respectively. Finally, articles are either sold to the end-customer or transferred back to the storage area.

For the evaluation, we again focus on the dwell time distributions in the individual states. The empirical histograms for the states *Storage*, *Sales floor 1* (for women), and *Sales floor 2* (for men) are shown in Figure 3.11. The interpretation of the empirical histograms is similar to the case study of the automated sorting process. An article stays in the storage area of the shop for $\mu_1 = 690\,\mathrm{h}$ on average before it is transferred to one of the two sales floors. There, the mean dwell times are $\mu_4 = 66\,\mathrm{h}$ for women and $\mu_5 = 79\,\mathrm{h}$ for men. Compared to the automated



(a) Shop storage, $\mu_3 = 690\,\text{h}$ **(b)** Sales floor women, $\mu_4 = 67\,\text{h}$ **(c)** Sales floor men, $\mu_5 = 79\,\text{h}$

Figure 3.11: Empirical dwell time distributions with fitted Exponential for three different process stages in a fashion supply chain. Since the discussed process is mainly characterized by manual interaction, the amount of outliers is larger compared to the case study of the automated sorting plant. However, the exponential distribution still provides a reasonable fit and can therefore be used to model the temporal behavior.

environment discussed in the first case study, the empirical data in this case shows an even higher amount of outliers, but can still be approximated with an exponential distribution. The analyzed data moreover provides valuable information for the shop management. Knowing the temporal behavior of individual articles or article groups gives insight to the customer behavior and allows for an intelligent replenishment process.

The presented case studies provide us with three important insights. First, we can conclude that the temporal behavior of items in a supply chain can be described by means of a time homogeneous CTMC. The empirical data from the two analyzed scenarios suggest that the exponential distribution provides a good approximation, regardless of a certain amount of outliers. Second, the empirical data from the discussed processes shows that the presented RFID system model is able to describe a wide variety of different supply chain structures, from highly automated environments to processes that are solely characterized by human interaction. Finally, the conducted evaluation shows that RFID systems are capable to provide information for an in-depth analysis with regard to the efficiency and throughput, which forms the basis for a subsequent optimization.

3.2 Spatio-temporal Item Correlation

The fact that items are aggregated in packaging units for transportation and easier handling introduces interesting aspects for the structure of RFID read events, as already described in Section 3.1.3. Besides the possibility to approximate a ground-truth for system monitoring tasks, the structure of logical item units gives rise to a systematic correlation of read events. Under the assumption that the content of packaging units does not change over time, a packaging unit

defines a persistent item set. This means that read events in a supply chain state will occur closely spaced in time as soon as the corresponding packaging unit enters the interrogation zone of a particular readpoint. The idea is hence to utilize the information about the existence of item sets in order to compensate for missing observations and to identify false positives. For this purpose, two prerequisites are needed: First, a mathematical formulation of item sets and an adequate way of detecting co-occurring read events needs to be developed. This can be used to infer the relationships between individual items and build a probabilistic *co-occurrence* model. Second, a way to incorporate the knowledge about item sets in the detection process at the readpoint level is required to compensate for missing and false positive observations. The co-occurrence model hence needs to provide a mechanism to evaluate whether or not a particular tag belongs to an item set based on the history of joint observations.

3.2.1 Modeling spatio-temporal correlation

The identification of spatio-temporal relationships within large datasets is a frequently found problem that has received a lot of attention in the data mining community. To describe the spatio-temporal structure of RFID observations, an intuitive approach is to investigate on the co-occurrence of read events caused by individual items in packaging units as soon as they enter the readpoint interrogation zone. An example of three packaging units with 10 items each is shown in Figure 3.12. The items give rise to correlated trajectories through the state-space representing the supply chain. However, read events are typically subjected to uncertainties due to the non ideal detection performance of individual readpoints. To describe this behavior in terms of a co-occurrence model, we utilize again the indicator variable z representing a read event in a particular state. Each packaging unit gives rise to a set of indicator vectors $\mathbf{Z} = (\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, ..., \mathbf{Z}^{(N)})^T$ from which the co-occurrence of read events among individual items can be derived by counting the number of times that tag T_i and tag T_j are observed together. This concept can be formulated in terms of the co-occurrence matrix $\mathbf{C}$ with individual entries

$$c_{ij} = |z_i > 0 \cap z_j > 0|. \tag{3.26}$$

The co-occurrence matrix is a frequently found tool in data mining applications such as market basket analysis [4]. For the exemplary observations in Equ. (3.23), the resulting co-occurrence

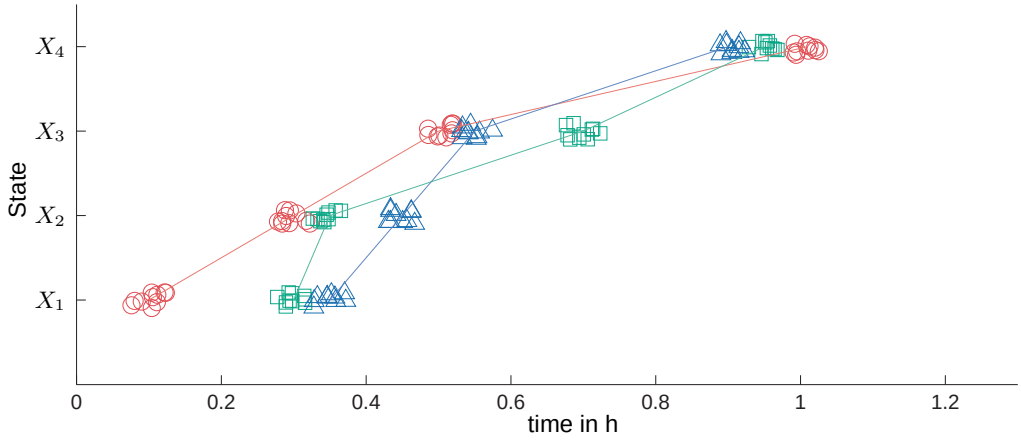


Figure 3.12: Exemplary state-space trajectories for three packaging units with 10 items each. Items belonging to a particular packaging unit give rise to neighboring read events in time and space, thereby introducing a correlation structure. This can be used to infer the relationship between individual items in terms of a co-occurrence model.

matrix is

$$\mathbf{C} = \begin{pmatrix} 3 & 3 & 3 & 3 & 1 \\ 3 & 4 & 4 & 4 & 2 \\ 3 & 4 & 4 & 4 & 2 \\ 3 & 4 & 4 & 4 & 2 \\ 1 & 2 & 2 & 2 & 2 \end{pmatrix}. \tag{3.27}$$

The diagonal elements of $\mathbf{C}$ describe how often a particular item is observed in total, whereas the off-diagonal elements describe the absolute frequency of joint observations. Consequently, $\mathbf{C}$ is a symmetric matrix, since $c_{ij} = c_{ji}$. The co-occurrence matrix can be used to derive two additional metrics. The *support*

$$s_{ij} = \frac{c_{ij}}{N}, \quad s_{ij} \in [0,\, 1], \tag{3.28}$$

where N is the total number of states (readpoints), describes the relative frequency of joint observations, which can be interpreted as the probability that tag T_i and T_j are observed together. The *confidence*

$$p_{ij} = \frac{c_{ij}}{c_{ii}} = \frac{s_{ij}}{s_{ii}}, \quad p_{ij} \in [0,\, 1] \tag{3.29}$$

is the conditional probability

$$p_{ij} = P(z_j \,|z_i) \tag{3.30}$$

of an observation z_j, given that z_i already occurred which can be interpreted as the strength of a connection between any two items. These two metrics are well suited for the integration in a

probabilistic framework, since they provide an explicit probability measure for a given observation based on previous data. The co-occurrence matrix, or alternatively the support matrix can be updated in an online manner with new observations of a particular packaging unit. For the exemplary data, the support and confidence matrices are

$$\mathbf{S} = \begin{pmatrix} 0.75 & 0.75 & 0.75 & 0.75 & 0.25 \\ 0.75 & 1.00 & 1.00 & 1.00 & 0.50 \\ 0.75 & 1.00 & 1.00 & 1.00 & 0.50 \\ 0.75 & 1.00 & 1.00 & 1.00 & 0.50 \\ 0.25 & 0.50 & 0.50 & 0.50 & 0.50 \end{pmatrix} \quad \mathbf{P} = \begin{pmatrix} 1.00 & 1.00 & 1.00 & 1.00 & 0.33 \\ 0.75 & 1.00 & 1.00 & 1.00 & 0.50 \\ 0.75 & 1.00 & 1.00 & 1.00 & 0.50 \\ 0.75 & 1.00 & 1.00 & 1.00 & 0.50 \\ 0.50 & 1.00 & 1.00 & 1.00 & 1.00 \end{pmatrix}. \tag{3.31}$$

The explicit description of item relationships in terms of a co-occurrence model enables us to evaluate particular RFID observations based on the joint previous history. Regarding the practical implementation, there are two additional aspects that need to be considered. The first is that the assumption of constant item relationships holds for a wide variety of different supply chain structures, but has its limitations when the supply chain features specific process steps such as commissioning or sorting. In such steps, the content of packaging unit changes by definition and consequently, practical implementations need to dissolve previously learned relationships. In this case, the co-occurrence model needs to be reset to a uniform distribution. The second aspect deals with the temporal characteristic of the co-occurrence model. Depending on the supply chain structure (in particular, the number of individual states), it is convenient to define a temporal window over which item relationships are learned.

3.2.2 Evaluating RFID observations

With the co-occurrence model, expressed in terms of the support and confidence matrices, we have a framework that allows us to describe the connection between two arbitrary items in the supply chain. In order to make use of this information, an evaluation method is required which can answer questions like '*Given that a particular set of tags $T_1, T_2, ..., T_N$ is observed, how likely is the absence or presence of tag T_j?*'. In other words, we want to evaluate the likelihood of an RFID observation Z_j given a set of observations $z_1, z_2, ..., z_N$ and the previously learned item relationships. This allows for the detection of missing tags if the available history suggests that a tag should actually be present in a packaging unit. Conversely, it enables us to identify false positive observations by means of their missing connection to other observed items.

The computation of the true likelihood $P(z_j \,|z_1,\, z_2,\, ...,\, z_N)$ requires the full joint distribution $P(z_j,\, z_1,\, z_2,\, ...,\, z_N)$, which is difficult to maintain in practice due to high dimensionality and

the inherent statistic dependence. Instead, an approximation can be found by combining the conditional probabilities $P(z_j \mid z_i)$ for $i \in [1,\ N]$ using a so called *opinion pool function* P_G such that

$$P(z_j \mid z_1,\ z_2,\ ...,\ z_N) \approx P_G(P(z_j \mid z_1), P(z_j \mid z_2), ..., P(z_j \mid z_N)). \tag{3.32}$$

The expression opinion pool stems from the similar problem to consolidate the opinion of several experts asked about a particular problem [1]. In our case, these experts correspond to the set of observed tags with a support value

$$s_{ij} > \gamma_O \tag{3.33}$$

larger than a given threshold. In other words, we only consider the conditional probabilities of frequent joint observations that have a sufficient support. Pooling operators have a variety of mathematical properties and there exists an elaborate framework for the description of the individual characteristics [9]. Although this mathematical treatment is beyond the scope of this thesis, there is one noteworthy property with direct implications to the considered problem. Certain pooling functions have the so called *0/1 forcing property*, which means that as soon as a single conditional probability equals zero (or one, respectively), the result of the pooling operation will also equal zero (or one). For the given problem, this property imposes a serious limitation since a single, undetected tag with a sufficient support value leads to a zero likelihood of the observation.

The different types of pooling functions can be categorized into additive and multiplicative methods. Although the latter class generally outperforms linear methods, it is not suitable for the considered problem due to the inherent zero forcing property. In contrast, the intuitive method of additive pooling

$$P_G = \sum_{\substack{i=1 \\ i \neq j}}^{N} w_i P(z_j \mid z_i), \qquad \sum_{i=1}^{N} w_i = 1 \tag{3.34}$$

does not exhibit the zero forcing property. The linear pooling operator reduces to a simple Arithmetic Bayes average when the weights w_i are chosen equally which is the case when no additional information about the observations is available.

Given a set of tag observations that are identified to belong to a particular packaging unit $\mathcal{P}_i$, the co-occurrence model and the linear pooling operator can be employed in a two-way evaluation scheme to identify missing and false positive tags. The first step involves a direct application of Equ. (3.34) for every tag observation, provided that the corresponding support is above the defined threshold γ_O. By means of that, every observation is evaluated in the light of the neighboring observations and the co-occurrence model. A potential false positive will have a significantly lower probability according to the pooling operation, due to the lack of joint previous observations. The

Step 1:
for all Packaging units $\mathcal{P}_i$ **do**
  for all Tags T_j **do**
    Compute $P_G^{(j)}$
    if $P_G^{(j)} > \gamma'$ **then**
      T_j detected, report z_j
    else
      T_j is false positive, suppress z_j
    end if
  end for
end for

Step 2:
for all Tags $T_j \in \mathcal{P}_i$ **do**
  Compute $P_G^{(j)}$
  if $P_G^{(j)} > \gamma'$ **then**
    T_j detected, report z_j
  else
    T_j is missing
  end if
end for

Algorithm 3.1: Two-way evaluation of tag observations using the co-occurrence model: Step 1 is used to identify false positive observations by means of lacking joint previous observations. In addition, step 2 uses the co-occurrence model to compensate for potentially missing observations, thereby introducing a low-pass characteristic for RFID observations.

second step is to iterate through all items previously assigned to the considered packaging unit and to compute their current observation likelihood. This step introduces a low-pass characteristic by filtering individual missed tags. The pseudo-code for the described two-way evaluation is shown in Algorithm 3.1. The final step in the evaluation is to perform an update on the co-occurrence model by means of the newly obtained observation data. Using this iterative scheme, the problem of missed and false positive detections at a particular readpoint can be effectively addressed by means of the inherent correlation structure of read events, which increases accuracy of the resulting RFID data.

3.2.3 Experimental Evaluation

To demonstrate the capabilities of the co-occurrence model, we conduct a simulation experiment involving eight distinct readpoints. The experimental setup is split into two parts: Readpoints $\mathcal{R}_1 \dots \mathcal{R}_4$ are designed such that a reliable identification of items in a packaging unit is possible, thereby enabling us to learn the co-occurrence model for the simulated set of packaging units. In contrast, readpoints $\mathcal{R}_5 \dots \mathcal{R}_8$ do not have these idealized properties. Instead, they exhibit a detection performance $P_{\mathrm{D}} < 1$ and are closely spaced such that the individual interrogation zones show a considerable overlap. By means of that, false positive observations are introduced with a certain probability. Whereas this design would not be favorable in a practical deployment, it enables us to explicitly analyze the filter capabilities of the co-occurrence model. The geometry of the experimental setup, together with the individual packaging unit trajectories is shown in the floor plan in Figure 3.13. Packaging units $\mathcal{P}_i$ are moved through the individual interrogation zones with a constant velocity of $v = 0.5\,\mathrm{m/s}$. After a packaging unit has proceeded through the interrogation zones $\mathcal{R}_1 \dots \mathcal{R}_4$, it proceeds to one of the four readpoints $\mathcal{R}_5 \dots \mathcal{R}_8$ with a probability of $p = 0.25$. This means that the simulated packaging units are distributed uniformly over the non ideal readpoints. Due to the chosen geometry, readpoints $\mathcal{R}_6$ and $\mathcal{R}_7$ show a higher false positive probability compared to $\mathcal{R}_5$ and $\mathcal{R}_8$. The simulation incorporates a total of $N = 200$ packaging units, with $M_i = 10$ items each.

For the evaluation we compare the detection and false positive probabilities for readpoints $\mathcal{R}_5 \dots \mathcal{R}_8$ to their corresponding raw values without the co-occurrence model as shown in Figure 3.14. The resulting detection and false positive probabilities indicate that the co-occurrence model provides a considerable increase in the readpoint performance. Whereas the detection rate is boosted to the ideal value of $P_{\mathrm{D}} = 1$ for every readpoint, false positive observations are perfectly suppressed. This ideal result is mainly caused by the perfect information incorporated in the co-occurrence model. Since readpoints $\mathcal{R}_1 \dots \mathcal{R}_4$ provide ideal observations, the resulting co-occurrence model accurately represents the individual item relationships. This enables the model to perfectly com-

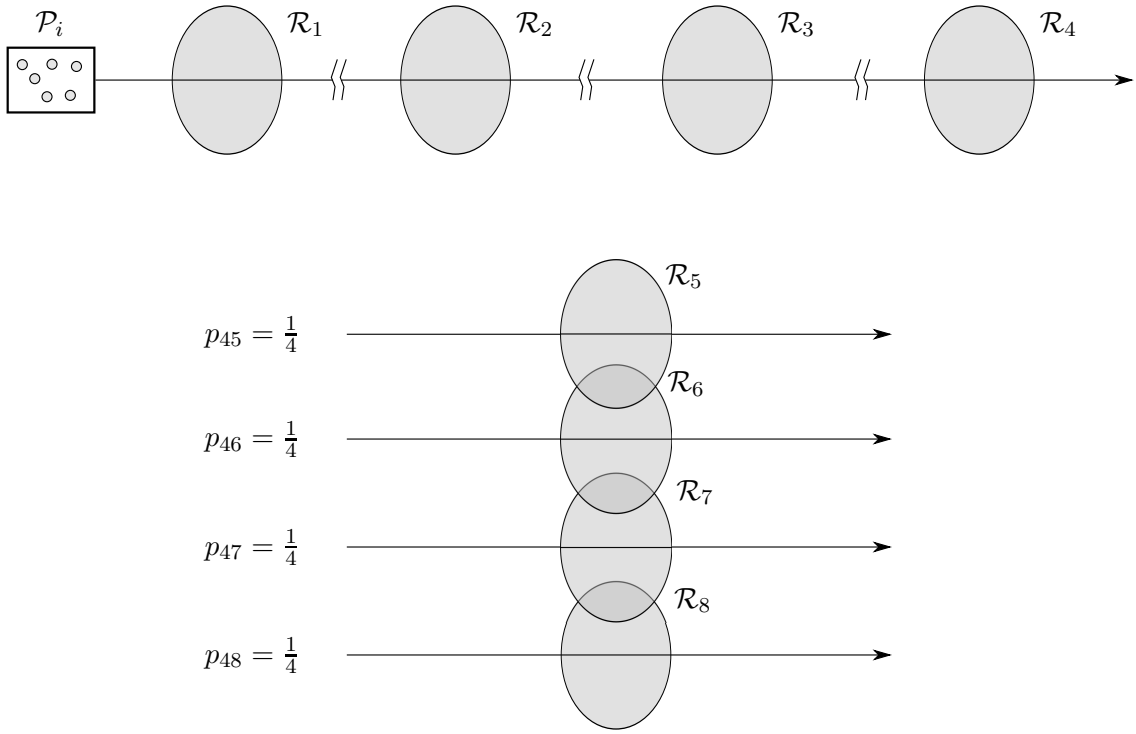
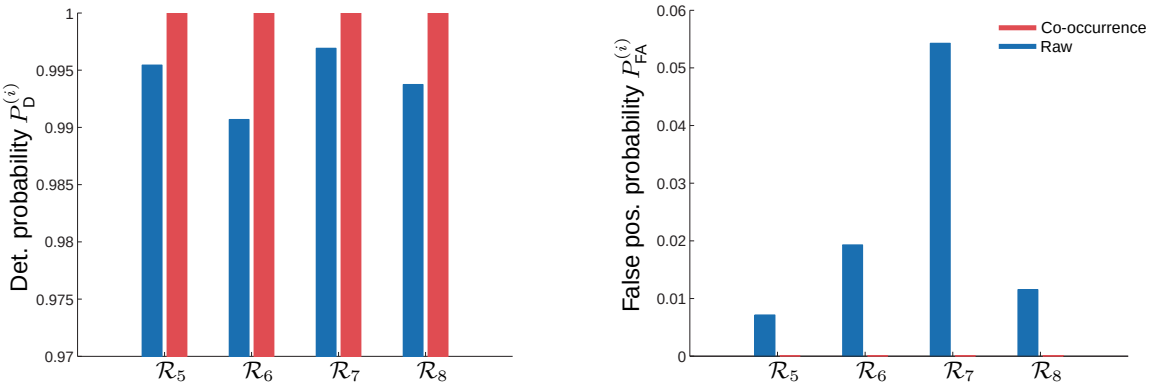


Figure 3.13: Simulation setup: Schematic representation of the simulated environment. The setup consists of a total of eight readpoints in states $\mathcal{R}_1 ... \mathcal{R}_8$. Packaging units $\mathcal{P}_i$ first travel through the interrogation zones of readpoints $\mathcal{R}_1 ... \mathcal{R}_4$ and then have a probability of $p = 0.25$ to travel through $\mathcal{R}_5 ... \mathcal{R}_8$. The first four readpoints are assumed to be ideal regarding the detection and false positive probability. In contrast, the readpoints $\mathcal{R}_5 ... \mathcal{R}_8$ suffer from false positives due to overlapping interrogation zones and have a non-ideal detection probability $P_{\mathbf{D}}^{(i)} < 1$.



(a) Detection probability

(b) False positive probability

Figure 3.14: Readpoint performance with and without co-occurrence model: With perfect information about the spatio-temporal item relationships, the co-occurrence model considerable enhances the detection probability and perfectly suppresses false positives.

pensate for missing reads and to suppress false positives in subsequent process steps.

The performance increase naturally is limited by the quality of the learned co-occurrence model. To investigate on this issue, we repeat the experiment described above under different conditions: Whereas the properties of readpoints $\mathcal{R}_5 \dots \mathcal{R}_8$ remain unchanged, the detection probability of the first four readpoints is deliberately reduced. In particular, we can define the co-occurrence model quality

$$q = \bar{P}_{\mathrm{D}}^{(1-4)} = \frac{1}{4} \sum_{i=1}^{4} P_{\mathrm{D}}^{(i)} \tag{3.35}$$

as equivalent to the average detection probability of readpoints $\mathcal{R}_1 \dots \mathcal{R}_4$. This enables us to analyze the average detection probability $\bar{P}_{\mathrm{D}}^{(5-8)}$ as a function of the model quality q, which is varied in the interval $q = [0.92, 1.00]$. The resulting detection performance together with the corresponding standard deviation is estimated in a series of $N = 10$ runs. The results of this analysis are shown in Figure 3.15.

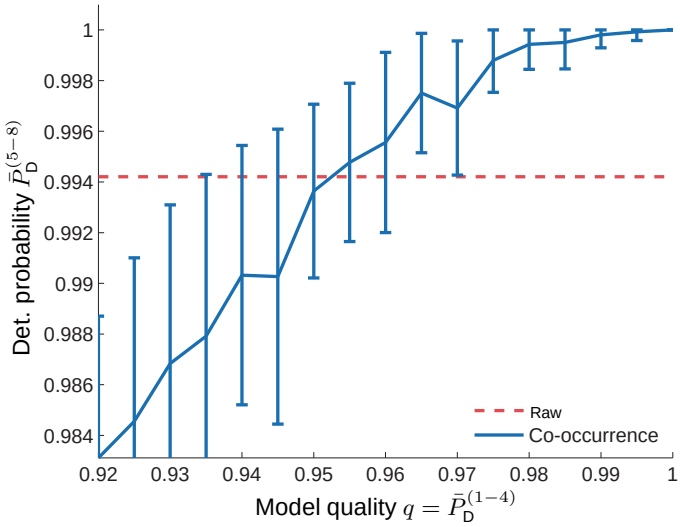


Figure 3.15: Resulting average detection probability for readpoints $\mathcal{R}_5 \dots \mathcal{R}_8$ as a function of the co-occurrence model quality q i.e. the average detection probability $\bar{P}_{\mathrm{D}}^{(1-4)}$. The co-occurrence model shows a considerable tolerance to noisy RFID observations which makes it ideally suited to integrate potentially noisy, uncertain information about existing item relationships.

The co-occurrence model is tolerant to noisy observations to a considerable extend, which makes it ideally suited to the application in RFID systems. If the model quality drops below a critical value, the overall performance is below the raw value since the co-occurrence model is dominated by uncertain or wrong information. In this scenario, a performance gain is already achieved for a model quality $q = \bar{P}_{\mathrm{D}}^{(1-4)} \geq 0.965$.

The simulation experiment presented in this section shows that the performance of individual readpoints can be improved considerably when the inherent correlation structure among RFID observations is considered. For this purpose, we have introduced a co-occurrence model which enables us to efficiently infer the relationship between items in an RFID enabled supply chain. The discussed evaluation method which is based on a linear pooling operator is an intuitive, yet powerful approach to increase the data quality in an RFID system. In particular, the described two-way evaluation can be used to identify false positive and missing observations by considering their joint history. The developed co-occurrence model is tolerant to a considerable amount of noise in the RFID observations, making it an effective way to incorporate potentially uncertain prior information about existing item relationships.

3.3 Summary

In this chapter, we have presented several high-level modeling concepts to target the problem of noisy data in RFID systems. Starting from a requirement definition, we have first developed a flexible model for RFID enabled supply chains which incorporates the dynamic behavior and the specific properties of RFID readpoints. For this purpose, we have combined a CTMC as motion model with an RFID sensor model that accounts for the possibility of missing and false positive observations. Using this model, we have discussed a process level localization mechanism which is designed to filter noisy RFID observations. Furthermore, we have employed the model in an online system monitoring mechanism to estimate the performance of individual readpoints at runtime. In order to verify the employed modeling assumptions, we have conducted an analysis of comprehensive datasets from practical applications.

The second modeling concept is based on the inherent spatio-temporal correlation structure of RFID read events in a supply chain. To utilize the information introduced by correlated observations, we have developed a co-occurrence model which describes the relationship between individual items. In particular, we have presented a framework to infer existing relationships and to evaluate RFID observations in light of the joint history. In a simulation experiment, we have demonstrated that this approach provides a significant performance increase and is tolerant to noisy observations.

4

RFID Readpoint Modeling

So far, we have defined an RFID observation as a binary event at a certain time instant which indicates that a tag is detected by a particular readpoint. At a lower abstraction level, state-of-the art RFID hardware provides more information within a particular read event such as the received signal strength and the phase angle of the tag response signal. Throughout this chapter, we will refer to this information as *low level features*. Similar to the system model approach described in Chapter 3, the goal is to establish a framework which allows us to evaluate RFID observations in order to decide whether a particular tag was deliberately identified or not. This provides us with the possibility to distinguish between true positive and false positive observations on the readpoint level, which provides more reliable RFID data to the backend system. The consideration on readpoint level is required whenever there is no possibility to employ a process-based filtering mechanism. This can be the case in specific supply chain states (e.g., verification after commissioning) or when the underlying process does not provide sufficient prior information.

To investigate on the possibilities on the readpoint level we will discuss the available low level features and establish the theoretical background behind the individual quantities. Starting from this theoretical consideration, we present a compact signal model which we integrate into a more general modeling framework that can be used in a classification mechanism. The presented approach is evaluated by means of an RFID conveyor belt scenario as typical example for a logistic

application. Complementary to the classification mechanism, we present an approach to mitigate the problem of missing tags by means of cooperative RFID readpoints. For this purpose, we employ a generalized Binomial model to consider the correlation between individual readpoints. The concept of cooperative readpoints is investigated by means of a comprehensive dataset stemming from a logistic application.

4.1 Feature Attributes

On the protocol level, the communication between reader and tag basically consists of a reader request and the tag response which, in the simplest case, contains the tag identifier and some protocol overhead. On the signal level, RFID readers measure the power and phase angle of the received tag response and provide this information to higher level software layers running on a host PC. A tag read event can therefore be characterized by a vector

$$\mathbf{e} = [t,\ \text{ID},\ r,\ \varphi,\ i] \tag{4.1}$$

comprising a timestamp t, the tag identifier ID, the RSSI r, the phase angle φ, and the logical index i of the antenna by which the inventory was performed. The antenna index is interesting for readers that have a number of antenna ports which are used in a time multiplexing scheme. If a tag remains within the interrogation zone of a reader, there will be a series of consecutive read events since the reader periodically inventories the present tag population. Besides the quantities stated in Equ. (4.1), the total number of read events (also referred to as read redundancy or read count) per unit time is an intuitive characteristic that describes how good a particular tag is identified. In this context, the *read rate*

$$\nu = \frac{\#\ \text{Read events}}{\#\ \text{Tags} \cdot \Delta t} \tag{4.2}$$

describes the total number of read events per tag and unit time. The term *read rate* is subject to conflicting definitions in the literature. Whereas some authors use it to describe the detection probability of an RFID reader or readpoint, we use this term to describe the number of read events per unit time. The basic idea behind many filtering approaches [30, 90, 132] is that tags located in the desired interrogation zone will be continuously identified and therefore provide a high number of read events per unit time, i.e., a higher read rate. The problem is, however, that the number of read events depends heavily on the total number of tags in the interrogation zone. This results from the anti-collision scheme used for the media access control. We demonstrate this dependency in an experiment where an increasing number of stationary tags is placed in the interrogation zone of an RFID reader such that each tag is inventoried periodically. The experiment is repeated $N = 5$

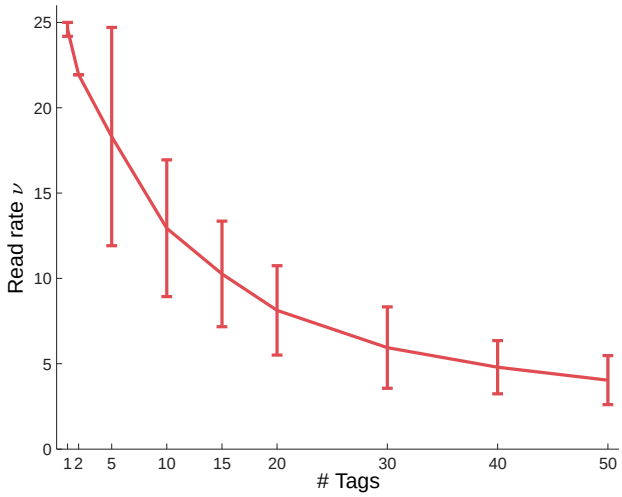


Figure 4.1: Read rate for a varying number of tags in the interrogation zone. The number of inventories per tag and unit time decreases with the overall number of present tags due to the sequential anti-collision scheme.

times for a duration of $T = 10\,\mathrm{s}$ allowing us to augment the best estimates with their respective standard deviations. As shown in Figure 4.1, the read rate for the specific reader in the experiment drops down to 25% of its maximum value already for as little as 20 tags in the interrogation zone – a number that is easily reached in practical applications.

Returning back to the actual low level features, the timestamp t of a tag read event is the most basic and intuitive quantity. Depending on the system architecture and implementation, there are several possible ways how the timestamp can be generated. Some RFID readers utilize their real-time clock (RTC) for setting the timestamp, which means that the accuracy is limited by the particular reader hardware. Another possibility is that the timestamp is generated on the host PC as soon as the tag read event is processed. In this case, the limiting factor in terms of accuracy is the non-deterministic behavior of the operating system rather than the reader RTC. In any case, the timestamp will be subjected to a jitter and drift over time. However, these uncertainties are usually in a negligible range and are thus not further investigated.

The next feature described in Equ. (4.1) is the received signal strength r, which is measured by the RFID reader for every inventory round. The signal strength is proportional to the power of the backscattered tag response given by

$$r \propto P_r = P_t G_t^2 (\mathrm{PL})^2 \sigma \tag{4.3}$$

as described by Nikitin and Rao [153]. In Equ. (4.3), P_t denotes the transmitted power, G_t is

the gain of the reader antenna, PL is the path loss of the UHF channel and σ describes the tag radar cross section. The path loss can be computed by generalizing the free-space loss model to consider multipath propagation [152]. Hence,

$$\mathrm{PL} = \left(\frac{\lambda}{4\pi d}\right)^2 \left| 1 + \sum_{n=1}^{N} \Gamma_n \frac{d}{d_n} e^{-\jmath k(d_n - d)} \right|^2 \tag{4.4}$$

where d is the distance between antenna and tag for the line-of-sight (LOS) path, Γ_n is the reflection coefficient of the nth reflecting path with distance d_n, k describes the wavenumber and N is the total number of multipath components. Equ. (4.3) is only a conceptual solution since the number of multipath components depends highly on the geometry and the surroundings and is thus unknown a priori. At the receiver front end of an RFID reader, the backscattered signal is demodulated which results in a complex valued baseband signal. This signal is characterized by an in-phase (I) and quadrature phase (Q) component, from which the RSSI can be derived in terms of

$$r = \frac{I^2 + Q^2}{Z_0}, \tag{4.5}$$

where Z_0 is the input impedance [150]. Towards the host PC, the RSSI is usually reported in units of dBm, typically with a resolution of ± 1 dBm. Besides the quantization effect, the RSSI is subject to different sources of noise, mainly caused by the multipath-channel characteristics and different object properties.

The phase angle

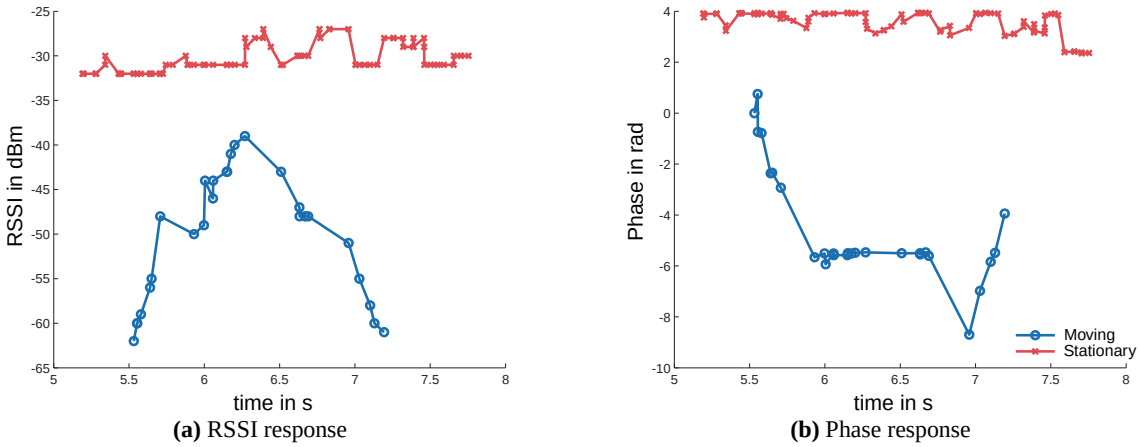
$$\varphi = -2kd \tag{4.6}$$

of the tag response depends linearly on the traveled distance d and the wavenumber k. At the receiver, the phase angle can also be derived from the I and Q components of the demodulated baseband signal using

$$\varphi = \arctan\left(\frac{Q}{I}\right). \tag{4.7}$$

Similar to the RSSI, the phase angle is perturbed by environment dependent noise. The RSSI and phase response of a tag moving through the interrogation zone of a reader with constant velocity and a stationary tag are shown in Figure 4.2. Both responses are characterized by a considerable amount of noise and the lack of a uniform sampling with respect to time.

Regardless of the inherent noise, the low level features show a characteristic behavior that can be considered in terms of a model. The approach followed by localization systems is to determine the location, velocity and/or moving direction of RFID tagged objects. From a theoretical point of view the RSSI and the phase angle provide the possibility to perform direction sensing and



(a) RSSI response **(b)** Phase response

Figure 4.2: Characteristic RSSI and phase responses for a stationary and a moving tag. Both signal features are perturbed by noise due to the environmental propagation conditions, the presence of other RFID tags and different object properties.

ranging in absence of multipath propagation and noise. Regarding the RSSI, the relationship between traveled distance and reflected power could be used directly to determine the distance between reader and tag. Similarly, the relationship between phase angle and distance in Equ. (4.6) allows for a range estimation. If the frequency f of the carrier signal is varied, the distance between reader and tag can be estimated using

$$d = -\frac{c}{4\pi}\frac{\Delta\varphi}{\Delta f}, \tag{4.8}$$

where c denotes the speed of light. However, the UHF channel is characterized by severe multipath propagation, which prohibits the use of this straight forward approach and makes localization and ranging a still desired feature for RFID systems. The phase response additionally provides the possibility to estimate the radial velocity of a tag when operating at a fixed frequency f using

$$v_r = -\frac{c}{4\pi f}\frac{\Delta\varphi}{\Delta t}. \tag{4.9}$$

This mechanism is widely used to measure the speed of moving objects by means of RF waves, which seems especially appealing for RFID systems. In most practical scenarios, tags are moved through the interrogation zone of a readpoint. In contrast the majority of false positive read events stems from stationary tags in the near vicinity. Therefore, the tag velocity is a key criterion to identify false positives. The velocity estimates over time ($v_r > 0$ for tags approaching the antenna, $v_r < 0$ for tags moving away from the antenna) for a moving and a stationary tag are

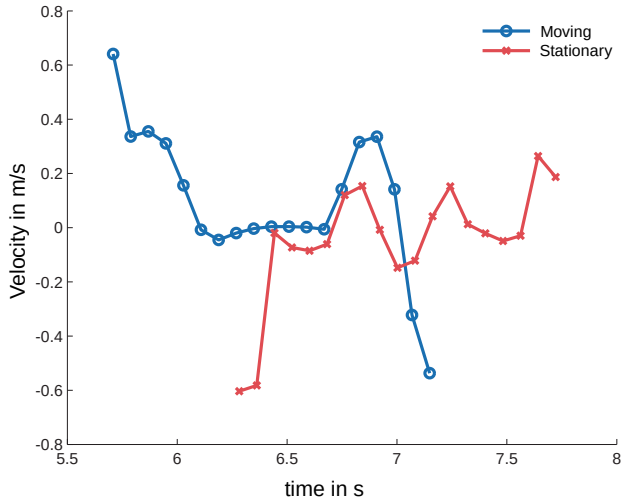


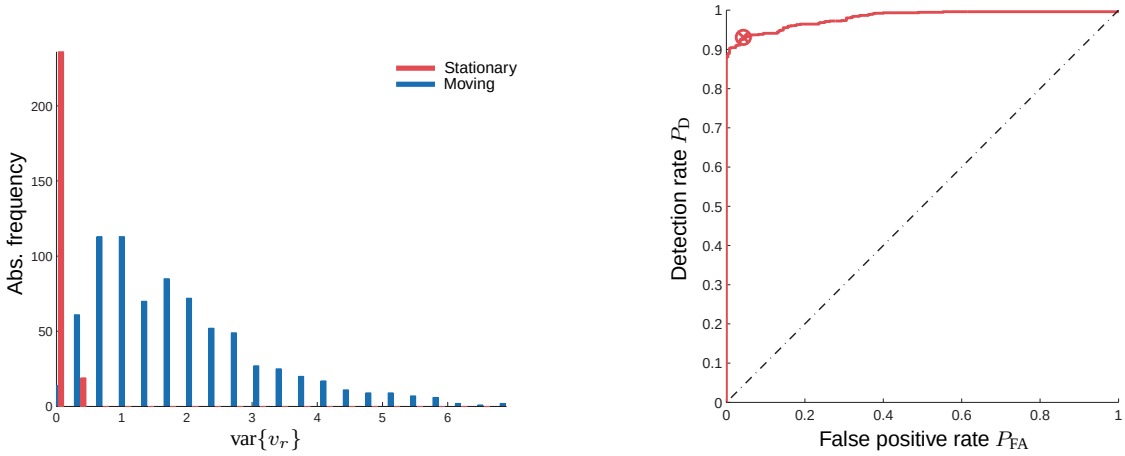
Figure 4.3: Velocity for a moving and a stationary tag. In contrast to the theoretical considerations, the velocity of the stationary tag shows considerable fluctuations and several zero crossings.

shown in Figure 4.3. The velocity response for the stationary tag exhibits random fluctuations and several zero crossings due to the noisy phase signal, making it difficult to identify that the tag is actually stationary.

The impact of the multipath propagation and the resulting noise in the phase response can be demonstrated in a simple experiment. The setup comprises an RFID reader, a set $\mathcal{S}_1$ of 765 moving and a set $\mathcal{S}_2$ of 5 stationary tags. The tags in $\mathcal{S}_1$ are placed in 51 cardboard boxes and moved through the readpoint interrogation zone with a constant velocity $v_r = 0.6\frac{\mathrm{m}}{\mathrm{s}}$. The stationary tags are placed in the interrogation zone such that they are continuously visible to the reader. For every cardboard box, the individual tag responses are recorded, yielding a total of 765 responses for moving, and 255 responses for stationary tags. The recorded phase responses are resampled and smoothed to estimate the radial velocity according to Equ. (4.9). Given a phase response consisting of N read events, we can estimate the variance of the tag velocity according to

$$\mathrm{var}\{v_r\} = \frac{1}{N}\sum_{n=0}^{N}\left(v_r[n] - \bar{v}_r\right)^2 . \tag{4.10}$$

An intuitive approach for a classification scheme based on this metric is that the absolute magnitude and the variance should be significantly smaller for stationary tags. By estimating the velocity variance from the individual phase responses and counting the number of occurrences, we can build an empirical histogram as shown in Figure 4.4(a). Due to the fact that the empirical histograms overlap, the classification performance suffers from Type I and Type II errors. In other



(a) Empirical histogram **(b)** Receiver Operating Characteristic (ROC)

Figure 4.4: Classification based on the radial tag velocity: (a) shows an empirical histogram of the velocity variance for moving (blue) vs. stationary (red) tags. Since the two histograms overlap, there is no cutoff point for an error free classification. (b) shows the ROC curve for the classification with the variance as threshold parameter. The optimal classification yields a detection rate $P_{\mathrm{D}} = 0.9306$ and a false positive rate $P_{\mathrm{FA}} = 0.0431$.

words, there is a certain percentage of false negative and false positive classification results. The Receiver Operating Characteristic (ROC) curve in Figure 4.4(b) shows the classification results when the threshold v_T is varied in the interval $[0, \mathrm{var}\{v_r\}_{\max}]$. The optimal classification yields a detection rate $P_{\mathrm{D}} = 0.9306$ and a false positive rate $P_{\mathrm{FA}} = 0.0431$ at a threshold value of $v_T = 0.0064$. The resulting error rate with this simple approach is far too high for the requirements for practical RFID systems. Taking into account that the considered experimental setup is well defined in terms of movement velocity, it is to expect that the performance will be even lower in scenarios with less stringent boundary conditions. Furthermore, the described setup is a typical near-field application and thus does not show excessive multipath characteristics. For this reason, we can conclude that the observation noise on feature level prohibits the use of straight-forward classification mechanisms based on the discussed low-level features. This motivates and justifies the use of more advanced and abstract modeling concepts which are capable of dealing with the noisy environment.

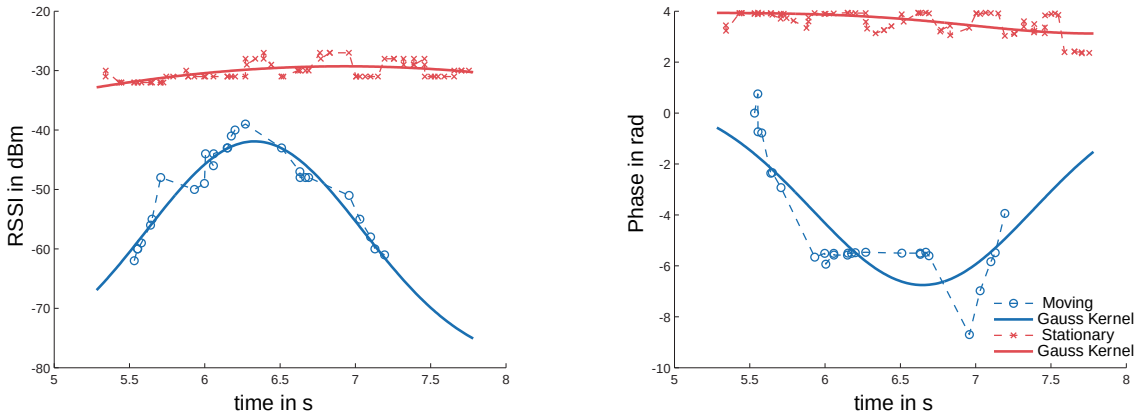
4.1.1 Signal model

The key to a reliable classification is an appropriate signal model with a sufficient robustness to the noisy low-level features. The RSSI and phase responses in Figure 4.2 suggest that both moving and stationary tags show a characteristic behavior which differs mainly in the variation

of the corresponding signals in time and amplitude. In particular, the response of a moving tag exhibits a characteristic peak that indicates when the tag was closest to the reader antenna. The idea behind the signal model and subsequent classification approaches is therefore to use a fitting function which allows us to describe and assess these characteristics. For the RSSI in free space, theory suggests that a polynomial of order $N = 4$ best describes the RSSI response. However, polynomial coefficients do not allow for a convenient assessment of the temporal location and extend of a signal. In contrast, a Gauss Kernel

$$g(t) = \frac{1}{\sigma_t \sqrt{2\pi}} \cdot e^{\frac{-(t-\mu_t)^2}{2\sigma_t^2}} \tag{4.11}$$

enables us to describe the center of gravity and the temporal extend by means of two scalar values, μ_t and σ_t. Figure 4.5 shows the RSSI and phase responses for a stationary and a moving tag with robustly fitted Gauss Kernels. As constraint for the fitting algorithm, the parameter μ_t is chosen such that it lies within the temporal limits of the observation frame. The temporal center μ_t allows



(a) RSSI response **(b)** Phase response

Figure 4.5: Characteristic RSSI and phase responses with a corresponding Gauss Kernel fit. The kernel describes the center of gravity and temporal extend by means of two scalar quantities, μ_t and σ_t.

us to efficiently determine when an RFID tag was closest to the reader antenna, i.e., actually present in the interrogation zone. This information is in particular useful when the requirement for the RFID system is to identify the content of individual, closely spaced packaging units. The temporal extend σ_t is inverse proportional to the movement velocity v_r and therefore is the key metric to distinguish between moving and stationary tags.

The signal model can be easily generalized to scenarios with more than one reader antenna. If a

readpoint features m antennas having a common interrogation zone, the estimates for μ_t and σ_t can be obtained by averaging over the individual estimates

$$\bar{\mu}_t = \frac{1}{m}\sum_{i=0}^{m}\mu_t^{(i)} \qquad \bar{\sigma}_t = \frac{1}{m}\sum_{i=0}^{m}\sigma_t^{(i)}. \tag{4.12}$$

The model described in Equ. (4.11) forms the basis for an abstract classification approach. The key requirement for the classification is a certain flexibility in terms of readpoint configuration, i.e., number and placement of antennas as well as the corresponding antenna radiation pattern. To provide this flexibility, the low-level features are considered in a state-space representation, similar to our RFID system model discussed in Chapter 3. The transformation from the signal domain to the state-space representation involves several steps. First, the fitted Gauss Kernels are abstracted by means of a *normalized state assignment function*

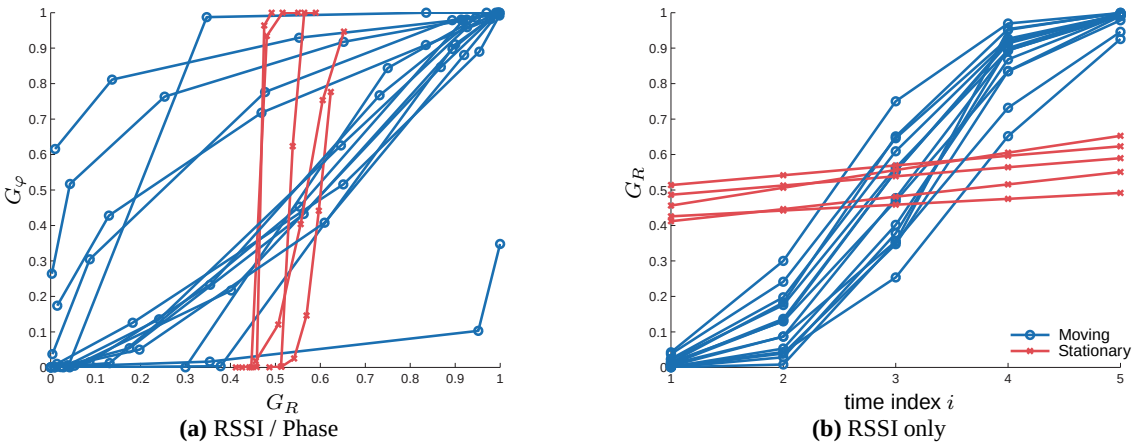
$$G(t_i) = \frac{1}{\sigma_t\sqrt{2\pi}}\int_{-\infty}^{t_i} e^{\frac{-(t-\mu_t)^2}{2\sigma_t^2}}\,\mathrm{d}t. \tag{4.13}$$

This function represents the value of the normal cumulative distribution function (CDF) of the estimated Gauss Kernel at sampling time t_i. Consequently, for each sampling time instant, we obtain a vector

$$\mathbf{f}(t_i) = [G_R(t_i),\ G_\varphi(t_i)]^T \tag{4.14}$$

in the feature space spanned by the RSSI and phase response. Periodic sampling of the feature space trajectory gives rise to a sequence $\mathbf{f}_1 ... \mathbf{f}_K$, where K is the total number of samples in the observation window. The state assignment functions for a set of 15 moving and 5 stationary tags with $K = 5$ samples are shown in Figure 4.6 for two different configurations: Whereas Figure 4.6(a) shows the trajectories in the RSSI-phase (G_R, G_φ) plane, 4.6(b) shows the trajectories over time when only the RSSI information is considered.

The next step is to translate the sampled features into a discrete set of observation symbols. This is achieved by means of a quantization step. For this purpose, a vector quantizer maps each sample to a discrete observation symbol, resulting in a sequence $\mathcal{O} = O_1 \dots O_K$. Returning back to the initial idea, the distinction between moving and stationary tags can now be formulated as a trajectory classification problem in the discrete state-space which can efficiently be performed using HMMs. HMMs provide a powerful and flexible framework to evaluate a set of given trajectories and can also be used as a generative data model [163]. Similar to the state-space model discussed in Chapter 3, HMMs comprise a motion model described by a transition matrix $\mathbf{A}$, an observation matrix $\mathbf{O}$ and a vector $\boldsymbol{\pi}$ describing the prior state probability distribution. The classification of



(a) RSSI / Phase **(b)** RSSI only

Figure 4.6: Sampled feature space trajectories for 15 moving and 5 stationary tags with $K = 5$ samples. (a) shows the trajectories in the RSSI-phase (G_R, G_φ) plane, whereas (b) depicts the trajectories over time when only the RSSI is used as feature attribute.

observation sequences involves two appropriately trained HMMs, λ_1 and λ_2 representing moving and stationary tags. The HMMs are used to compute the likelihood of an observation sequence and to assign a particular tag to the most likely class. This can be achieved by means of the likelihood ratio

$$\Lambda = \frac{P(\mathcal{O} \,|\, \lambda_1)}{P(\mathcal{O} \,|\, \lambda_2)} \lessgtr \gamma, \tag{4.15}$$

that is compared to a threshold value γ. Depending on which HMM is more likely to represent the observed sequence, the considered tag is assigned to the corresponding class.

The classification scheme described above requires a training phase to initialize the vector quantization and to estimate the parameters of the HMMs λ_1 and λ_2. Consequently, the training involves two steps: First, the vector quantizer is initialized based on the normalized state assignment functions, i.e., the sampled feature space trajectories. This can be done using unsupervised learning algorithms, such as the standard K-Means algorithm. From the sampled feature space trajectories, K-Means identifies cluster regions and allows for the estimation of the corresponding centroids $\boldsymbol{\mu}_i$ and covariance matrices $\boldsymbol{\Sigma}_i$. The results of applying the K-Means algorithm to an exemplary dataset consisting of 135 tags is shown in Figure 4.7, where the feature space is divided into $K = 5$ clusters in accordance with the number of samples in the observation window. The vector quantization assigns the sampled low-level features to one of five distinct observation symbols. The second training step is required to determine the parameters of the HMMs for moving and stationary tags. The HMM parameters $\lambda_i = (\mathbf{A}_i, \mathbf{B}_i, \boldsymbol{\pi}_i)$ can be estimated by applying the Forward-Backward algorithm to a set of labeled training sequences [163].

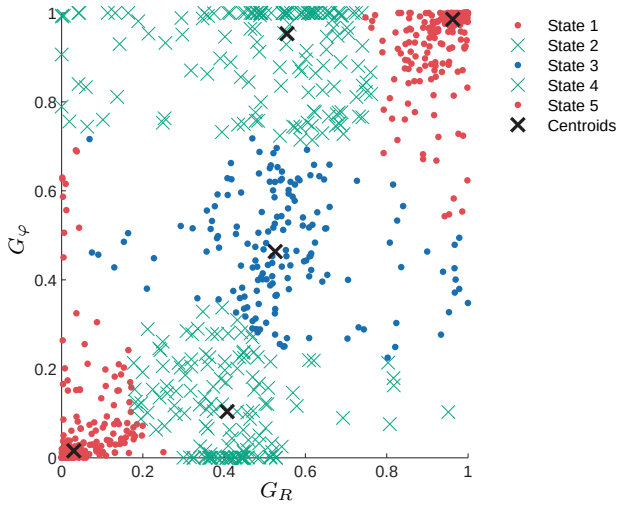


Figure 4.7: Feature space comprising RSSI and phase information for a training set of 135 tags. The feature space is clustered into $K = 5$ clusters in accordance with the number of samples in the observation window. The clusters are required to assign each low-level feature sample to a distinct observation symbol.

The number of possible and meaningful observation symbols (which defines the resolution of the vector quantizer) is strongly connected to the sampling frequency and the overall length of the observation window. An increased sampling frequency requires a higher resolution, resulting in a higher model complexity and the potential danger of over-fitting to the training samples. In contrast, a lower sampling frequency leads to a reduced number of distinct observation symbols which limits the discriminative power of the state-space model. For this reason, a compromise between model complexity respectively over-fitting and discriminative capability needs to be found for a given scenario. Considering the shape of the Gauss Kernel fits, the absolute minimum number of observation symbols is $K = 3$ to describe the signal characteristics.

The classification scheme developed in this chapter is flexible in terms of readpoint configuration and can therefore be employed in a variety of different scenarios. The necessary training phase is well suited to be integrated in the deployment of an RFID readpoint. The HMM classification framework provides an efficient way to evaluate individual tag responses. Besides the filtering of read events at the supply chain level, this forms the second important cornerstone to provide reliable RFID data for the backend system.

4.1.2 Case Study: Conveyor Belt Application

In logistic applications, a core functionality of RFID systems is to identify the content of packaging units after manual commissioning. Based on specific order information provided by an Enterprise-Resource-Planning (ERP) system, articles are manually aggregated in packaging units in a process referred to as *picking*. The task of the RFID system is then to verify the content of the packaging unit by identifying every contained item. Depending on the scan result, the packaging unit is further processed for delivery or manually corrected, as shown in Figure 4.8. The manual correction step is a time consuming and work intensive process. Considering the fact that missing a single tag triggers the manual correction of the entire packaging unit emphasizes the high requirements for an RFID system in order to allow for an efficient operation. A simple calculation example demonstrates that even small error rates cause a considerable work-load for manual correction. Consider an RFID system with a detection probability of $P_{\mathrm{D}} = 0.995$ and an ideal suppression of false positives. Assume that packaging units contain $M = 20$ items on average and that missed tags are distributed uniformly over all packaging units. The error rate on packaging unit level for this system is $p_E = (1 - P_{\mathrm{D}}) \cdot M = 0.1$, which means that 10% of all packaging units require manual correction due to errors introduced by the RFID system.

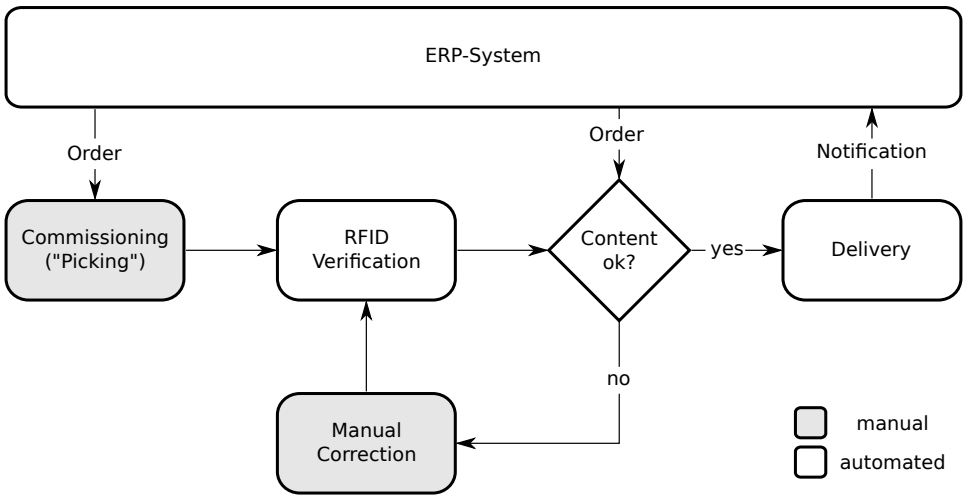


Figure 4.8: Commissioning and verification with RFID in a typical logistic application: Based on an order provided by the ERP-system, articles are commissioned (picked) in a packaging unit. The content of the packaging unit is then verified by means of an RFID scan and manually corrected if necessary.

A common approach to verify the content of individual packaging units for flat packed goods is to deploy RFID reader antennas along a conveyor belt. By means of that, the content of individual

packaging units can be identified during transportation. This has two main advantages: First, the identification usually does not require any additional time or manual interaction. Second, RFID readpoints benefit from fixed boundary conditions such as a constant movement velocity and a fixed box-to-box distance. This makes the identification more robust in comparison to processes that are characterized by manual handling.

The described use case is difficult to support with high-level modeling approaches due to the lack of prior information. Since the content of a packaging unit solely depends on the specific customer order, it is difficult to integrate prior process information as a countermeasure to the noisy RFID observations. In particular, the spatio-temporal relationship among items described in Chapter 3 cannot be utilized, since the commissioning process can be viewed as the initialization step for the co-occurrence model. For this reason, the classification of low-level read events is mandatory to provide accurate scan results. The readpoint objectives for the verification can be formulated as follows:

- *Objective 1:* Identify the content of a particular packaging unit: This requires the readpoint to perform an assignment of read events to the boxes as they move through the interrogation zone. For this purpose, the individual boxes need to be identified by means of a dedicated RFID tag or a barcode label.
- *Objective 2:* Filtering of false positive tags: Tags that do not belong to the considered packaging unit should be identified and suppressed. This includes stationary tags in the vicinity of the readpoint as well as tags from previous or subsequent packaging units on the conveyor.

The first objective, which we refer to as *tag-to-box assignment* is elegantly covered by the presented signal model since it allows us to determine the temporal signal center for every tag. Assuming a constant conveyor velocity v, we can investigate on the relation between the RSSI / phase responses and the time instants when boxes are present in the interrogation zone. For the experimental evaluation in this case study, we define two scenarios: The first scenario assumes that a proximity sensor (light barrier) in combination with a barcode reader is available. In this case, the center timestamps $t_{\mathcal{P}_i}$ of individual packaging units can be computed from the rising and falling edge of the proximity sensor signal

$$t_{\mathcal{P}_i} = \frac{t_{i,1} + t_{i,2}}{2}, \tag{4.16}$$

where $t_{i,1}, t_{i,2}$ denote the time instants of the rising and falling edge respectively. The second scenario covers the case when no proximity sensor is available. Hence, the packaging unit center

timestamps are obtained from the center timestamps of the dedicated box identifier tags

$$t_{\mathcal{P}_i} = \mu_i. \tag{4.17}$$

With the center timestamps for N consecutive boxes, the most likely box for tag T_j can be determined by finding the minimum temporal distance between the tag center timestamp μ_j and the N packaging unit center timestamps using

$$i = \arg\min_i \left(|\mu_t - t_{\mathcal{P}_i}| \right). \tag{4.18}$$

The readpoint setup for this use case, consisting of an RFID reader with a single antenna, a light barrier sensor and barcode scanner to identify the individual packaging units is shown in Figure 4.9. The experiment is designed to demonstrate the capabilities of the presented approach

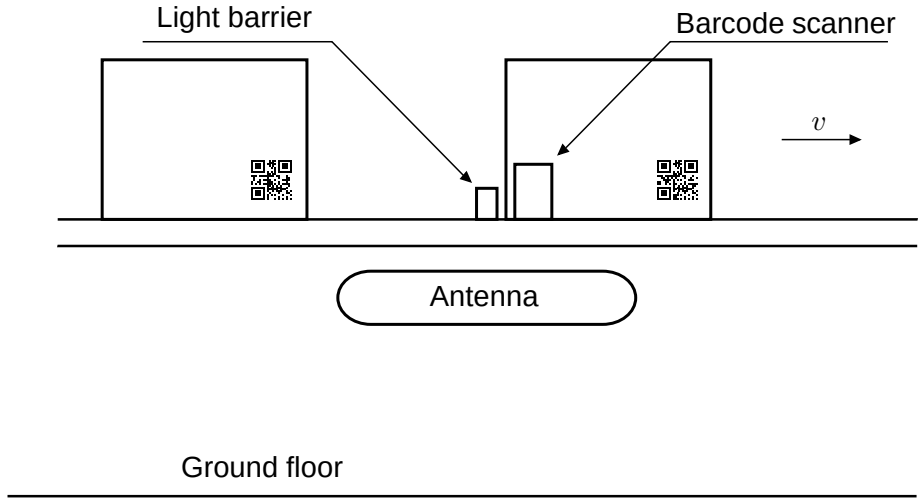


Figure 4.9: Readpoint setup comprising a conveyor belt and an RFID reader antenna mounted below the conveyor belt facing upwards. The presence of packaging units can be detected by means of a proximity sensor and a barcode scanner is used to identify the individual packaging units.

and to evaluate the individual signal features in terms of classification performance. For this purpose, a set of six feature attributes $\mathcal{F}_1 \ldots \mathcal{F}_6$ is defined, representing all possible combinations of the individual low-level features. Prior to the evaluation, the classifier is trained by means of a labeled data set. The evaluation is carried out by means of tagged items in a total of 51 boxes that are transported on a conveyor belt with a constant velocity v. Each box contains 15 items, including a dedicated box identifier tag. At the same time, five stationary tags are placed near the reader antenna such that they are continuously visible to the reader. Details about the RFID readpoint setup and reader configuration are summarized in Table 4.1.

Component / Parameter	Name / Value
Reader	Impinj Speedway Revolution R420
Antenna	Kathrein Widerange $70°/30°$
Transmit Power P_{TX}	$0.3\,\mathrm{W}$
Conveyor velocity v	$0.5\,\frac{\mathrm{m}}{\mathrm{s}}$
Box-to-box distance d	$0.5\,\mathrm{m}$

Table 4.1: Experimental conveyor belt setup: Summary of parameters and readpoint configuration

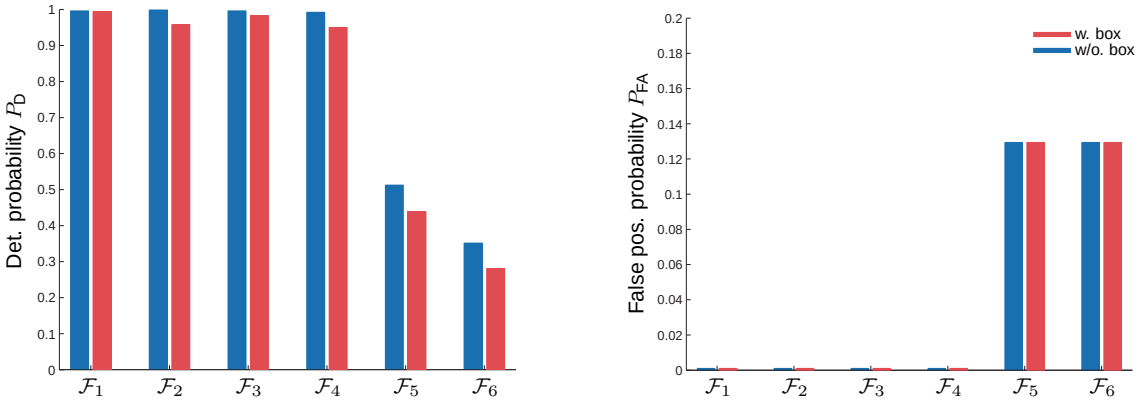
Feature attributes	w/o. box		w. box	
	$\hat{P}_{\mathrm{D}}$	$\hat{P}_{\mathrm{FA}}$	$\hat{P}_{\mathrm{D}}$	$\hat{P}_{\mathrm{FA}}$
$\mathcal{F}_1$: RSSI / LB	0.9961	0.0	0.9948	0.0
$\mathcal{F}_2$: RSSI	0.9987	0.0	0.9582	0.0
$\mathcal{F}_3$: RSSI / phase / LB	0.9961	0.0	0.983	0.0
$\mathcal{F}_4$: RSSI / phase	0.9922	0.0	0.9503	0.0
$\mathcal{F}_5$: phase / LB	0.5124	0.1294	0.4392	0.1294
$\mathcal{F}_6$: phase	0.3516	0.1294	0.2810	0.1294

Table 4.2: Performance characteristics of the classification approach using different feature attributes $\mathcal{F}$, consisting of RSSI, phase, and light barrier (LB).

The evaluation metrics for the conducted experiments are the overall detection probability P_{D} and the false positive probability P_{FA}. The detection probability considers the reader detection probability (i.e., the probability of physically detecting a tag) and the classification performance in terms of a correct tag-to-box assignment. The false positive probability accounts for stationary tags that are erroneously assigned to a packaging unit. The experiment is repeated several times to obtain a comprehensive data set. For each of the feature attributes, the classification performance is evaluated with and without the tag-to-box assignment. The results of the evaluation are summarized in Table 4.2 and visualized in Figure 4.10.

The evaluation provides two main insights. First, the proposed signal model integrated in the abstract classification scheme provides a reliable suppression of false positives and an accurate tag-to-box assignment with appropriate low-level features. The perfect suppression of false positives together with a detection probability of $\hat{P}_{\mathrm{D}} = 0.9987$ for feature attributes $\mathcal{F}_2$, consisting solely of the RSSI information demonstrate the effectiveness of this approach. The gap to a perfect detection probability stems from the fact that a single tag could not be physically identified and can therefore not be attributed to the classification scheme. However, the classification performance depends strongly on the quality of the feature attributes and drops considerably when only the phase signal is used. Figure 4.6 illustrates the reason for this result: The phase angle of the tag response is too noisy to provide meaningful information that could be exploited by the

classification algorithm. If, however, the RSSI is added to the feature attributes, the performance increases and stationary tags can be filtered out reliably as indicated by $\hat{P}_{\mathrm{FA}} = 0$ for all four test runs. The second aspect is concerned with the reliability of the tag-to-box assignment. Whereas the suppression of false positives does not benefit from the light-barrier information, the tag-to-box assignment is more accurate when taking the light-barrier signal into consideration. This can be explained by the fact that the light-barrier provides an accurate estimate of the box center timestamps.



(a) Detection probability **(b)** False positive probability

Figure 4.10: Resulting detection and false positive rate for different feature attributes $\mathcal{F}_i$, with and without tag-to-box assignment. In particular, the feature attributes are $\mathcal{F}_1$ (RSSI / LB), $\mathcal{F}_2$ (RSSI), $\mathcal{F}_3$ (RSSI / phase / LB), $\mathcal{F}_4$ (RSSI / phase), $\mathcal{F}_5$ (phase / LB), and $\mathcal{F}_6$ (phase). The combination of RSSI and light-barrier signal ($\mathcal{F}_1$) provides a perfect tag-to-box assignment and a robust suppression of false positives. In contrast, the phase signal is too noisy to provide a reliable classification in this scenario.

The issue of missing observations is a widely discussed problem for RFID systems and is still a limiting factor for the mass deployment in several applications. For this reason, there are several approaches to mitigate this problem, from an increased inventory duration to the use of multiple tags for every item. Another intuitive approach is the concept of cooperative readpoints, which will be discussed in the next section. This concept is an ideal addition to the low-level classification approach discussed above to provide a reliable detection of RFID tags together with a robust suppression of false positives.

4.2 Cooperative RFID Readpoints

Commonly, RFID systems rely on the principle of periodic inventories by making use of consecutive reader-sessions. The idea behind this approach is to increase the overall scan duration such

that all tags in the interrogation zone reply to the reader request with a sufficient probability. The need for a periodic inventory mainly stems from the necessary anti-collision scheme. Depending on the overall number of tags in the interrogation zone and the protocol parameters, a certain number of tag replies result in a collision on the air interface. In order to fully identify a given tag population, the reader hence needs to repeat the request to provide an additional slot for individual tags to reply.

Within the operation limits of the EPCglobal standard in terms of item-throughput, the main cause of the missing tag problem, however, is not associated with the temporal behavior of the anti-collision scheme. Instead, the detection probability is usually limited by the *forward link* from reader to tag, i.e., tags do not receive sufficient power to reply to a reader request [116]. The temporal diversity introduced by periodic inventory sessions can only compensate for the missing tag problem to a certain extend. Especially for stationary readpoints, the geometric constraints do not change significantly between two inventory sessions which means that additional reader requests do not necessarily improve the detection probability. For this reason, a vital step during the conception of an RFID system is a careful placement of RFID reader antennas, combined with an optimization of the parameter-setup to compensate for varying tag orientation, item throughput and item variability.

The concept of antenna diversity with respect to an RFID readpoint can be further extended to include several, cooperative readpoints. Following the definition by Fyhn *et al.* [57], cooperative means that the individual readpoints exchange information about identified tags. The overall system performance benefits from the use of additional readpoints in two ways: First, every readpoint introduces additional reader sessions, thereby increasing the overall inventory time. This decreases the probability of persistent collisions by adding additional inventory slots. Second, additional readpoints give rise to a spatial diversity which decreases the effect of dead zones and hence increases the overall detection probability. The concept of spatial diversity is widely developed in wireless communications but requires special attention in the context of RFID systems. The often tightly constrained readpoint setup in automated environments gives rise to a significant correlation between individual readpoints, even if the reader antennas are sufficiently spaced.

Taking the correlation between individual readpoints into account, this section presents a correlated Binomial model to describe the tag detection process for cooperative readpoints. The model provides an intuitive way to quantify the correlation between individual antennas or readpoints and can therefore be used to estimate the number of necessary system components for a given performance requirement. This reduces the trial-and-error characteristic often found in current conception processes and forms the basis for a more systematic approach.

4.2.1 System model

The detection performance of an isolated readpoint $\mathcal{R}_i$ can be described by means of the detection probability $P_{\mathrm{D}}^{(i)}$, which quantifies the percentage of tags in the interrogation zone that are successfully identified. In the idealized case of an RFID system consisting of N uncorrelated, cooperative readpoints $\mathcal{R}_1$, ...$\mathcal{R}_N$ with equal detection probabilities $P_{\mathrm{D}}^{(1)} = P_{\mathrm{D}}^{(2)} = ... = P_{\mathrm{D}}^{(N)} = P_{\mathrm{D}}$, the combined detection probability $\bar{P}_{\mathrm{D}}$ follows a Binomial distribution. In order to account for the inherent correlation between readpoints, we need to extend the standard Binomial model to consider the conditional probability for a detection z_i at readpoint $\mathcal{R}_i$ given previous read events z_{i-j}

$$P_{hh}^{(i)} = P(z_i = 1 \,|\, z_{i-1} = 1,\, z_{i-2} = 1,\, ...,\, z_1 = 1) \tag{4.19}$$

and

$$P_{hm}^{(i)} = P(z_i = 1 \,|\, z_{i-1} = 0,\, z_{i-2} = 0,\, ...,\, z_1 = 0). \tag{4.20}$$

In this context, the subscript h represents a successful detection (‘*hit*’), whereas a subscripted m refers to a ‘*miss*’. Consequently, $P_{hh}^{(i)}$ denotes the probability of subsequent detections (hits), for example when readpoints $\mathcal{R}_1$ and $\mathcal{R}_2$ both have successfully identified a tag. In contrast, $P_{hm}^{(i)}$ denotes the conditional probability of a detection following a miss. To keep the computation tractable, we employ the first order Markov assumption, which means that a detection by readpoint $\mathcal{R}_i$ depends only on the detection by readpoint $\mathcal{R}_{i-1}$, yielding

$$P_{hh}^{(i)} = P(z_i = 1 \,|\, z_{i-1} = 1) \tag{4.21}$$

and

$$P_{hm}^{(i)} = P(z_i = 1 \,|\, z_{i-1} = 0), \tag{4.22}$$

respectively. Similarly, the conditional probabilities for a missed tag are defined as

$$P_{mh}^{(i)} = P(z_i = 0 \,|\, z_{i-1} = 1) = 1 - P_{hh}^{(i)} \tag{4.23}$$

and

$$P_{mm}^{(i)} = P(z_i = 0 \,|\, z_{i-1} = 0) = 1 - P_{hm}^{(i)}. \tag{4.24}$$

A sequence of read events represented by the indicator variable $Z = [z_1, z_2, ..., z_N]$ for a particular tag can be interpreted as a Markov chain described by the conditional probabilities in Equ. (4.21) – (4.24). Using this set of conditional probabilities, we can construct the combined detection probability $\bar{P}_{\mathrm{D}} = P(k, N)$ in an iterative manner [114] as visualized in Figure 4.11, where each node $P_{kN} = P(k, N)$ represents the probability of k successful detections by N

readpoints. A vertex m_{kN} represents the probability of k detections by N readpoints with a miss

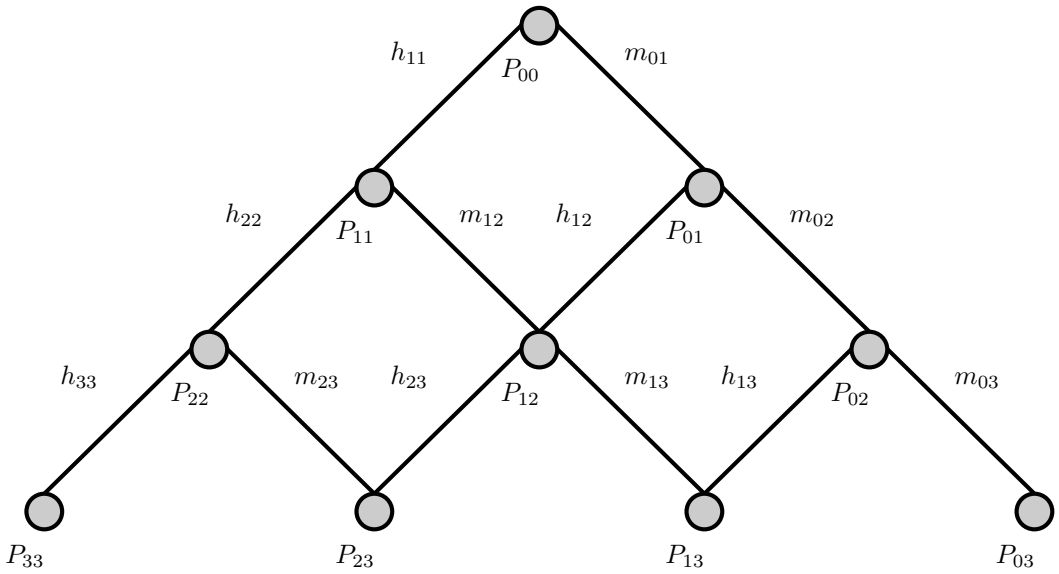


Figure 4.11: Graphical representation of the probabilistic tag detection model to compute the combined detection probability. Each vertex represents a conditional probability, whereas a node $P_{kN} = P(k, N)$ is the probability of k detections by a given number of readpoints N. The notation is adopted from Ladd [114].

on the N-th readpoint. Depending on whether the previous node was reached through a hit or a miss, this yields

$$m_{kN} = \begin{cases} P_{mm}^{(N)} & \text{if } z_{N-1} = 0 \\ P_{mh}^{(N)} & \text{if } z_{N-1} = 1. \end{cases} \tag{4.25}$$

Similarly, vertex h_{kN} is the probability of k detections by N readpoints with successful detection on the N-th readpoint, yielding

$$h_{kN} = \begin{cases} P_{hm}^{(N)} & \text{if } z_{N-1} = 0 \\ P_{hh}^{(N)} & \text{if } z_{N-1} = 1. \end{cases} \tag{4.26}$$

The combined detection probability $P(k, N)$ can be computed as the sum over all possible paths through the tree in Figure 4.11, where the first node $P(1, 1) = h_{11} = P_{\mathrm{D}}^{(1)}$ is the unconditional detection probability of the first readpoint.

From the general model, we can derive the special case for the detection by at least one readpoint

$$P(k \geq 1, N) = 1 - P(0, N) = 1 - (1 - P_{\mathrm{D}}^{(1)}) \prod_{i=2}^{N} P_{mm}^{(i)}, \tag{4.27}$$

which is the desired combined detection probability $\bar{P}_{\mathrm{D}}$. In this context, $P(0, N)$ denotes the probability that a tag is missed by all N readpoints. If we assume that the conditional probabilities $P_{mm}^{(1)} = P_{mm}^{(2)} = ... = P_{mm}^{(N)} = P_{mm}$ are equal among all readpoints, Equ. (4.27) becomes

$$P(k \geq 1, N) = 1 - (1 - P_{\mathrm{D}}^{(1)}) P_{mm}^{N-1}. \tag{4.28}$$

The combined detection probability for this simplified case is shown in Figure 4.12 for different levels of correlation, expressed in terms of P_{mm} and an unconditional detection probability $P_{\mathrm{D}}^{(1)} = 0.8$. In addition, the uncorrelated case covered by the standard Binomial distribution is shown for comparison. Although the simplifying assumption of equal conditional probabili-

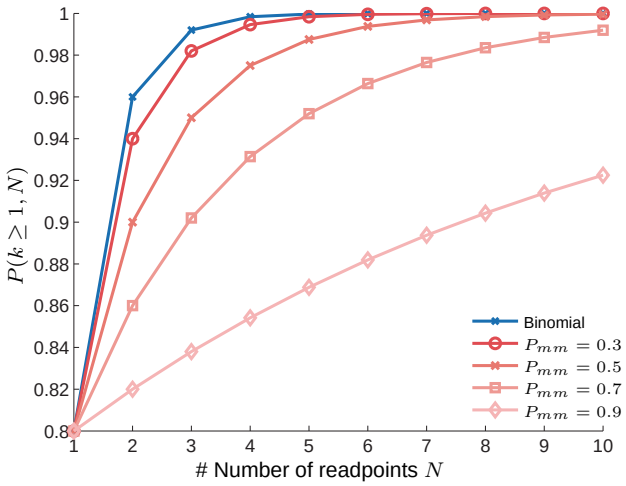


Figure 4.12: Simulated detection probability for a given number of readpoints with different degrees of correlation. The correlation is expressed in terms of P_{mm}, the conditional probability that two subsequent readpoints fail to detect a particular tag.

ties will not hold in practice, this example highlights two important aspects. First, the standard Binomial distribution readpoints forms the upper bound for the combined detection probability. For practical scenarios, this yields overly optimistic results due to the idealized assumptions. Second, the higher the correlation, the more readpoints are required to achieve a specific target performance which means that an increasing correlation limits the gain in detection performance introduced by additional readpoints. Therefore, it is important to consider measures to reduce

the correlation before performing an additional RFID scan. Such measures include an alternative orientation between reader antenna and tag, but more importantly a rearrangement of tags inside their enclosing packaging units to reduce the impact of mutual coupling and detuning.

The presented model provides a concise way to describe the correlation among individual readpoints and to quantify the combined detection probability. The model parameters can be estimated in a straight forward manner from a set of read events and can then be used to extrapolate the results in order to assess if an additional readpoint yields the desired performance improvement. To further study the effect of readpoint correlation, we use the discussed model to analyze a large scale dataset from a real-world RFID deployment.

4.2.2 Experimental Evaluation

The experimental evaluation presented in this section is part of a larger study with the goal to identify and quantify the different error sources in a practical RFID deployment. In parallel, we use the dataset to verify the assumptions of the generalized Binomial model. The data for the evaluation stems from a two month period of an active RFID deployment in a fashion distribution center. In this application, the task of the RFID system is to verify the content of individual packaging units (containing flat packed garments) after commissioning. The deployed RFID system consists of a total of $N = 4$ readpoints as shown in Figure 4.13. The individual readpoints are in-

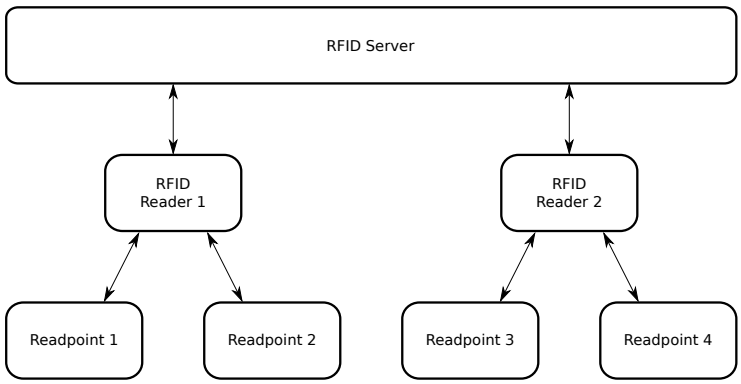


Figure 4.13: RFID system architecture: The RFID system comprises two RFID readers and $N = 4$ readpoints, each consisting of two antennas. The read events are processed on a centralized RFID server.

stalled along an automated conveyor belt and consist of two directional RFID antennas which are placed directly above and below the conveyor belt. This placement is a logical consequence of the tag dipole radiation pattern to maximize the energy transfer between reader antenna and tag. The

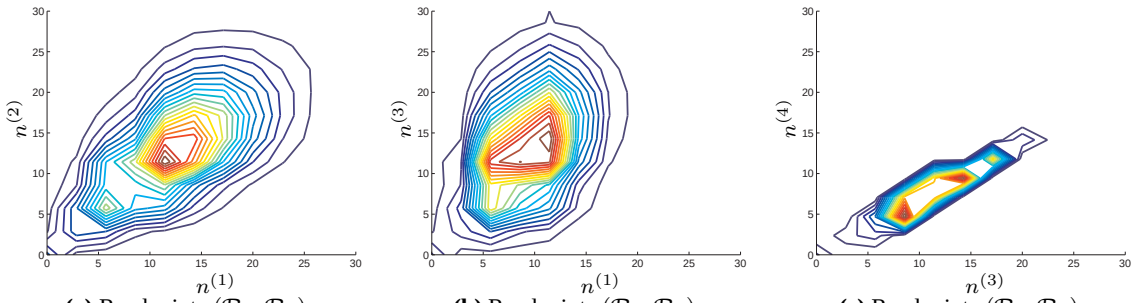
first two readpoints, $\mathcal{R}_1$ and $\mathcal{R}_2$, are shown in Figure 4.14 together with several cardboard boxes transported on the motorized conveyor belt. Due to the similar orientation and close spacing, the readpoints show a considerable correlation with respect to the detected RFID tags.

Figure 4.14: RFID readpoints installed on a motorized conveyor belt. The two antennas mounted above the conveyor belt are aligned such that the energy transfer to the tag is maximized. Due to the similar orientation and close spacing, both readpoints show a considerable correlation with respect to the detected RFID tags. Image courtesy of Enso Detego GmbH, 2013.

The raw dataset $\mathcal{D}_1$ features read events from over 500.000 tags in 23.365 packaging units during a two month period. In order to allow for an in-depth analysis, a ground-truth has been established by manually scanning the individual packaging units.

The main error source is identified in a certain percentage of defective tags, i.e., tags that do not respond to a reader query in the far field due to a hardware failure. In order not to bias the analysis, we establish a corrected dataset $\mathcal{D}_2$ by removing all defective tags from the ground-truth. The evaluation includes the following aspects: First we analyze the correlation between individual readpoints on feature level by investigating on the read redundancy (number of read events) for the individual tags. Second, we fit the correlated Binomial distribution to the raw and corrected dataset to derive the conditional probabilities and to assess the gain in detection probability introduced by additional readpoints. Finally, we carry out a temporal analysis of the readpoint detection performance to demonstrate that the spatial diversity concept provides a significant performance enhancement in terms of the missing tag problem.

For the analysis on the feature level we consider the total number of successful inventories $n_i^{(j)}$ for tag T_i on readpoint $\mathcal{R}_j$. Provided a constant conveyor velocity, the read redundancy can be interpreted as a quality indicator that describes how reliable a tag is identified. Contour plots of the joint empirical distributions for readpoint pairs $(\mathcal{R}_1, \mathcal{R}_2)$, $(\mathcal{R}_1, \mathcal{R}_3)$ and $(\mathcal{R}_3, \mathcal{R}_4)$ are shown in Figure 4.15. From the dataset, we can compute the linear correlation coefficients $\rho_{i,j}$ for all readpoint pairs, as summarized in Table 4.3. The empirical distributions and the correlation



(a) Readpoints $(\mathcal{R}_1, \mathcal{R}_2)$ **(b)** Readpoints $(\mathcal{R}_1, \mathcal{R}_3)$ **(c)** Readpoints $(\mathcal{R}_3, \mathcal{R}_4)$

Figure 4.15: Contour plots of the joint empirical read count distributions on the individual readpoints. With constant movement velocity, the read count can be interpreted as a measure of how reliable a tag is identified. Due to the similar readpoint setup, the read count exhibits a considerable correlation.

ρ	$n^{(1)}$	$n^{(2)}$	$n^{(3)}$	$n^{(4)}$
$n^{(1)}$	1.00	0.58	0.35	0.35
$n^{(2)}$	0.58	1.00	0.34	0.34
$n^{(3)}$	0.35	0.34	1.00	0.98
$n^{(4)}$	0.35	0.34	0.98	1.00

Table 4.3: Read count correlation coefficients between the individual readpoints. Readpoints $(\mathcal{R}_3, \mathcal{R}_4)$ are highly correlated on feature level due to the almost identical setup.

coefficients highlight the significant correlation between the individual readpoints. This implies that tags which are often read by a given readpoint have a high probability of being detected by subsequent readpoints in a similar fashion. In particular, readpoint pair $(\mathcal{R}_3, \mathcal{R}_4)$ shows an extraordinary correlation with $\rho_{3,4} = 0.98$. This can be explained by the similarity of the physical setup and the small spacing between the readpoints.

From the read count, we can derive the binary indicator variable

$$z_i^{(j)} = \begin{cases} 1 & \text{if } n_i^{(j)} > 0 \\ 0 & \text{if } n_i^{(j)} = 0 \end{cases} \tag{4.29}$$

which can be used to analyze the correlation on detection level. In particular, the detection probabilities of the individual readpoints are shown in Table 4.4, whereas the conditional probabilities are shown in Table 4.5. In accordance to the analysis on feature level, the conditional probabilities in Table 4.5 reflect the significant correlation among the individual readpoints. Based on

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_4$
$\hat{P}_{\mathrm{D}}$	0.9959	0.9953	0.9900	0.9885

Table 4.4: Detection probabilities for the corrected dataset $\mathcal{D}_2$. For example, readpoint $\mathcal{R}_1$ exhibits a detection performance of $\hat{P}_{\mathrm{D}}^{(1)} = 0.9959$.

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_4$
$\hat{P}_{hh}$	0.9959	0.9982	0.9905	0.9985
$\hat{P}_{hm}$	0.0041	0.0018	0.0095	0.0015
$\hat{P}_{mh}$	0.9959	0.3013	0.8746	0.0
$\hat{P}_{mm}$	0.0041	0.6987	0.1254	1.0

Table 4.5: Conditional probabilities of detecting and missing a tag on subsequent readpoints. The consistently high values for $\hat{P}_{hh}$ emphasize the correlation between the individual readpoints. For example, the probability of two successful detections on $\mathcal{R}_3$ and $\mathcal{R}_4$ is $\hat{P}_{hh} = 0.9985$.

the estimated values, we can fit the correlated Binomial to the empirical data and construct the combined detection probability $\hat{P}_{\mathrm{D}} = P(k \geq 1, N)$. The resulting detection probability over the number of readpoints is visualized in Figure 4.16. Starting from the unconditional detection probability of the first readpoint, the additional readpoints increase the combined detection probability for the raw dataset $\mathcal{D}_1$ and the corrected dataset $\mathcal{D}_2$. Although the individual detection probabilities are close to or above 99%, the performance of the isolated readpoints is not sufficient to meet the requirements in practical scenarios. In contrast, the fusion of read events provides an improvement in terms of the combined detection probability. For the corrected dataset $\mathcal{D}_2$, a perfect detection can be achieved with $N = 3$ readpoints. Since readpoints $\mathcal{R}_2$ and $\mathcal{R}_3$ show the least degree of correlation, $\mathcal{R}_3$ provides the largest gain in terms of detection performance, whereas readpoint $\mathcal{R}_4$ is redundant in this scenario. Furthermore, Figure 4.16 highlights that the standard Binomial distribution provides overly optimistic estimates, since the assumption of uncorrelated readpoints is clearly violated. In contrast, the suggested correlated Binomial model provides accurate values that fit the experimental dataset. The small deviation from the empirical data stems from the simplifying first order Markov assumption.

For the last part of the analysis, we investigate on the temporal evolution of the individual and combined detection probabilities. The readpoint detection probabilities over time are show in Figure 4.17. Although the average detection performance for every readpoint is close to or above 99%, the plot shows several significant performance dips indicating that tags in particular packaging units are difficult to identify. Whereas these performance dips correlate strongly among readpoint ($\mathcal{R}_1$, $\mathcal{R}_2$) respectively ($\mathcal{R}_3$, $\mathcal{R}_4$), the correlation between ($\mathcal{R}_2$, $\mathcal{R}_3$) is considerably less significant. This is emphasized by the combined detection probabilities for readpoints $\mathcal{R}_1 ... \mathcal{R}_4$

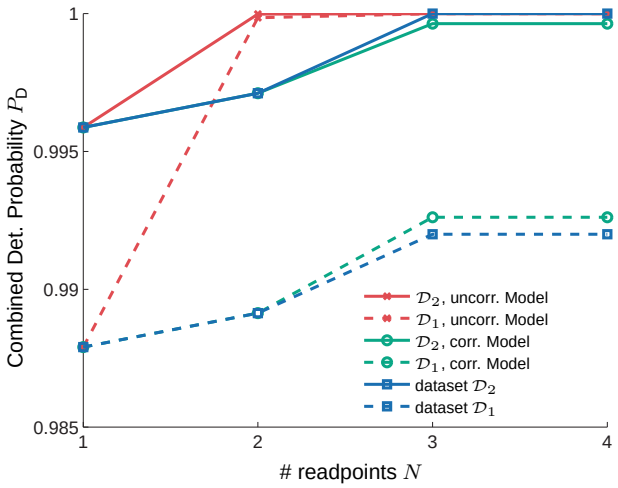


Figure 4.16: Resulting detection probability over the number of readpoints for the raw dataset $\mathcal{D}_1$ and the corrected dataset $\mathcal{D}_2$. Whereas the standard Binomial distribution provides overly optimistic estimates, the correlated Binomial model provides an accurate fit to the experimental data and can therefore be used to predict the detection rate as a function of the number of readpoints.

as shown in Figure 4.18. The combination of the individual readpoints reduces the number and extend of the performance dips. In the particular scenario, readpoints $(\mathcal{R}_1$ - $\mathcal{R}_3)$ provide sufficient diversity for the mutual compensation of missed RFID tags. The concept of spatial diversity is hence an effective way to mitigate the missing tag problem if the correlation issue between individual readpoints is properly addressed.

The generalized Binomial model provides an effective way to assess the correlation among individual readpoints and is hence a valuable tool during the conception of an RFID system. The empirical study by means of a comprehensive dataset has shown that individual readpoints suffer from random performance dips which can be significantly reduced by means of the cooperative readpoint concept.

4.3 Summary

The concepts developed in this chapter are targeted to provide an increased data accuracy for RFID systems on the readpoint level. This is particularly necessary when no prior information can be utilized in terms of a high-level system model. From this perspective, the presented readpoint model serves as complement to the high-level approach discussed in Chapter 3.

In order to address the problem of false positive observations, we have analyzed the information on readpoint level in terms of the available low level features. Using a compact signal model,

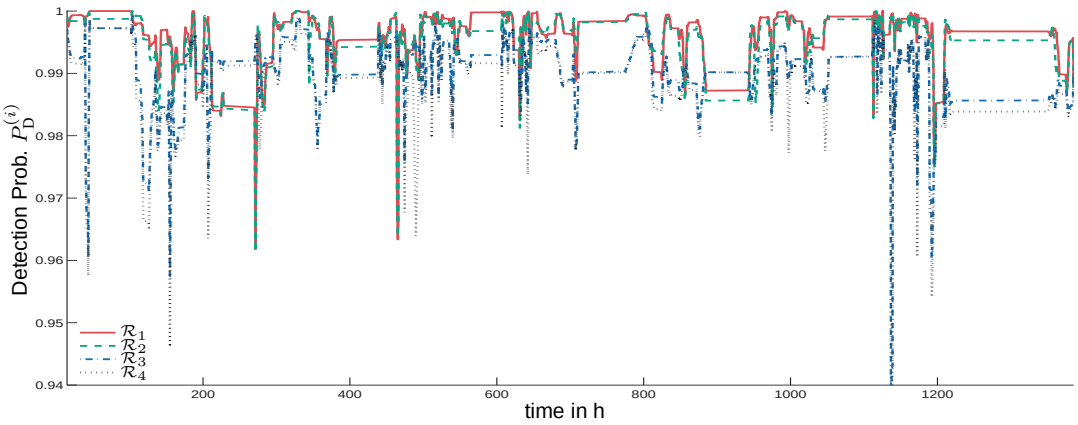


Figure 4.17: Detection probabilities for readpoints $\mathcal{R}_1 ... \mathcal{R}_4$ over time. Whereas the average detection performance is close to or above 99%, the individual readpoints suffer from random performance dips. An important aspect is that the performance dips show a certain correlation among the readpoints, which indicates that tags in particular packaging units are difficult to identify.

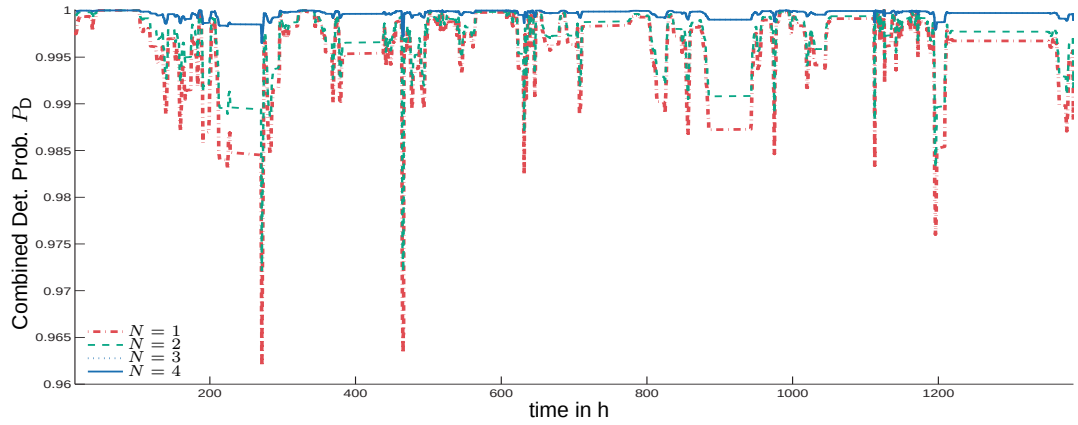


Figure 4.18: Combined detection probability of the RFID system with $N \in [1, 4]$ readpoints. Whereas individual readpoints suffer from performance dips, the fusion of read events provides a robust and reliable detection. The curves for $N = 3$ and $N = 4$ are identical which shows that the fourth readpoint is redundant in this scenario.

we have integrated these features in an abstract state-space representation which forms the basis for a robust classification mechanism. The abstraction level in the state-space representation provides the necessary flexibility and ensures that this approach can be used in a variety of different scenarios. In an experimental study, we have evaluated the individual low level features and the classification approach.

In addition to the classification mechanism, we have also presented a method to target the problem of missing RFID observations. For this purpose, we use a readpoint diversity approach which combines the RFID observations from several, independent readpoints. In this context, the correlation between readpoints has been found as the limiting factor, which is why we have explicitly considered this aspect in the presented model.

The major conclusions from this chapter can therefore be summarized as follows. First, RFID systems provide information that goes beyond the definition of a binary read event. By employing this information, a reliable suppression of false positive observations can be performed. Second, the presented classification mechanism is a flexible framework suited for a variety of different use cases. Finally, the crucial problem of missing tags can be mitigated by means of a readpoint diversity concept. The probabilistic formulation on readpoint level developed in this chapter provides a high flexibility regarding possible application scenarios and different configurations.

5

PRISE - Probabilistic RFID Simulation Engine

RFID systems cover a variety of different aspects and face complex questions from physical aspects such as wave propagation and circuit design to high-level business considerations. Due to the fact that the deployment of RFID systems is a cost and time intensive process, simulation tools for a variety of different aspects are frequently used in the conception phase. Consequently, the research community has developed various simulation and emulation frameworks addressing the different abstraction layers of an RFID system [23]. Since data from large scale RFID deployments is scarce, a simulation environment forms the necessary basis for the development and evaluation of high-level modeling approaches and filter techniques. The first part of this chapter gives an overview of different simulation frameworks for RFID systems and provides a comparison of the individual approaches. The second part discusses the concepts and ideas behind the probabilistic simulation framework PRISE and the implemented models. Finally, we present an experimental evaluation of the simulated RFID observations in three practically relevant scenarios.

5.1 Related Simulation Tools

There exist several simulation frameworks targeting the low-level aspects of an RFID system. For example, the MATLAB based simulation framework by Han *et al.* [79] focuses on the analog frontend of RFID readers and tags to verify the protocol and standard compliance of a particular tag or reader model. For this reason, the reader frontend is modeled in detail, including the digital-to-analog conversion (DAC) as well as the modulator and power amplifier subsystems. The framework furthermore provides mechanisms to specifically simulate the forward and reverse link by means of a wireless channel model and the tag reflection behavior. The simulation is performed on signal level which limits its applicability due to the considerable computational complexity. Similarly, the MATLAB based simulation framework PARIS [13–15] implements a detailed model for RFID readers, tags, and the wireless multipath channel. In addition, the PARIS framework considers the aspect of Ultra-Wideband signaling in RFID systems and is especially designed to investigate on the issue of geometric tag localization. Hence, PARIS features an abstraction level that is not adequate for the simulation of a large scale RFID system.

In terms of a high-level system simulation with various RFID readers, antennas and tags, RFIDSim [51, 52] is closest to the PRISE simulation framework presented in this thesis. RFIDSim abstracts from the signal level which reduces the computational complexity such that RFID systems with a large number of tags can be simulated. The communication between reader and tags features the mandatory commands specified in the EPCglobal standard, such as *Query*, *Ack*, *QueryRep* and *QueryAdj* [48]. The behavior of tags is implemented according to the state diagram specified in the EPCglobal standard. In terms of signal propagation, RFIDSim uses a Rician fading channel model, but does not especially consider item properties or detuning effects. A simulator that closely resembles the functionality of RFIDSim was developed by Zhang *et al.* [197], including a Graphical User Interface (GUI). In addition to the simulation frameworks discussed above, there are several other simulation and emulation approaches targeting passive UHF RFID systems, security aspects [138] and hardware-in-the-loop simulation [40]. These systems, however, do not consider high-level aspects such as the underlying supply chain structure.

The probabilistic simulation engine presented in this chapter is a system-level simulation tool that provides a compromise between high-level business process simulation and the generation of low-level RFID observations. Naturally, this requires a trade-off in the simulation accuracy due to simplified modeling assumptions. PRISE is hence not designed as a detailed reader, tag or channel simulation software, but rather as a framework to generate datasets for large scale RFID systems operated in a particular supply chain. The modular software architecture allows for the integration of more detailed models and the output of other simulators to address particular aspects

if necessary. The main differentiation to existing RFID simulators is that PRISE performs a true system level simulation, including the underlying business process which defines the abstract, high-level item movement.

5.2 Simulator Concepts

PRISE is a cross platform, C++ / Qt based software framework which provides the possibility to set up and simulate typical RFID readpoints in a supply chain. For this purpose, the simulator implements the business process model described in Chapter 3 to generate high-level item trajectories. PRISE offers a GUI allowing the user to specify the following aspects:

- *Supply chain structure setup:* Defines the structure and dynamics of the simulated supply chain in terms of the CTMC as described in Chapter 3. The setup comprises the model parameters (transition probability matrix, prior state probabilities and the individual dwell time parameters) as well as the overall simulation time and temporal resolution.

- *Item setup:* Allows for the definition of a hierarchical item structure that can be used to create packaging units and tagged items. Together with the supply chain structure, the item setup serves as input for the generation of high-level item trajectories.

- *Readpoint setup:* Defines the number, type and placement of RFID readers, connection cables and antennas in a 3D environment. Additionally, the readpoint setup covers the definition of geometric item trajectories. In order to account for variabilities in the geometric setup, a controllable degree of randomness can be added to the item trajectories.

The simulation setup is represented in an .xml file structure that serves as input to the actual simulation core. An example configuration file, representing a single reader with stationary tags is shown in Appendix A.1. The actual simulation core implements an event queue driven by a simulation clock that controls the movement of items, schedules reader inventories and updates the channel characteristics. Based on this information, a matrix holding the individual detection probabilities $p_{ij}(t)$ for the ith reader antenna and the jth tag is computed. A block diagram of the simulator architecture is shown in Figure 5.1. The simulation output is represented in terms of RFID observations

$$\mathbf{e}_{\text{Sim}} = [t,\ \text{ID},\ r,\ i], \tag{5.1}$$

with the timestamp t, tag identifier ID, RSSI value r, and antenna index i in analogy to the definition in Equ. (4.1), except for the phase angle φ. The simulated read events are provided

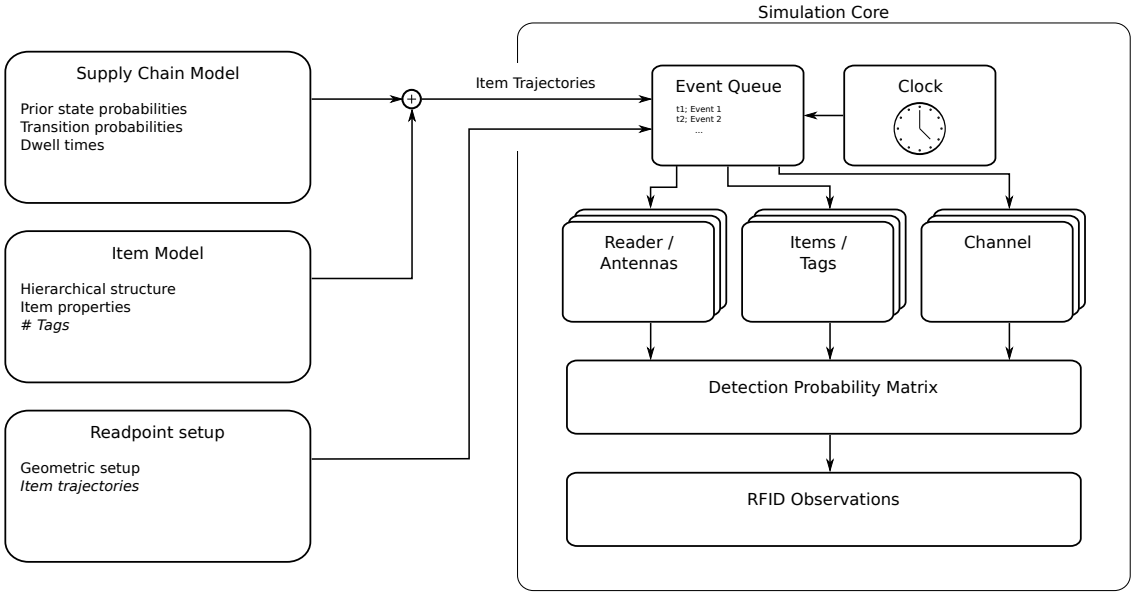


Figure 5.1: Simulator architecture: The simulation setup, comprising a supply chain model, item model, and readpoint setup is fed to the simulation core. The core is based on an event queue which is driven by a simulation clock. The implemented models for readers, antennas, items, tags, and the UHF channel are used to compute the detection probability matrix $\mathbf{P}$, where each entry $p_{ij}(t)$ represents the detection probability at time t for the jth tag on the ith reader antenna.

in a .csv file structure, hierarchically grouped by reader and tag. In addition, the output contains the ground-truth for the high-level supply chain movement and the geometric item trajectories on every readpoint.

5.3 Models and Implementation

The simulator implementation follows an object-oriented approach and currently incorporates specific RFID reader-, antenna-, and tag-models. The specification for the individual components are closely related to the corresponding manufacturer data sheets and can be subjected to a controllable amount of randomness. Currently, there exists no comprehensive dataset that provides information about the distribution of individual parameters such as reader or tag sensitivity. For this reason, standard Normal distributions with adjustable parameters are used.

5.3.1 RFID Tag

RFID tags are characterized by means of their antenna and the transponder chip. The implemented antenna model is a standard dipole with a specific antenna impedance $Z_a = R_a + \jmath X_a$. The dipole radiation pattern is shown for reference in Figure 5.2. The transponder chip is characterized by

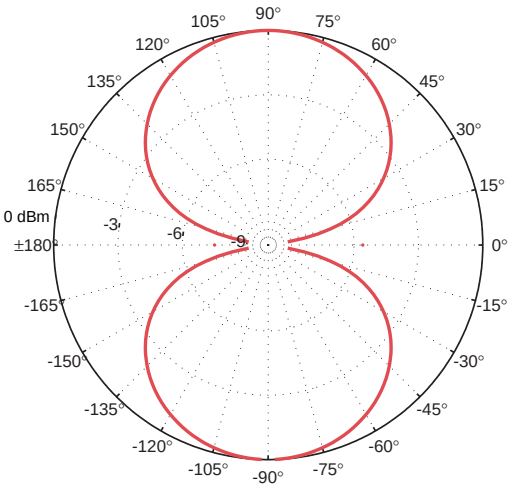


Figure 5.2: Tag antenna radiation pattern: The tag antenna features a dipole characteristic which introduces a significant orientation sensitivity.

the sensitivity value and the chip impedance $Z_c = R_c + \jmath X_c$. The individual values can be set according to the manufacturer data sheet or available measurement data [156]. The chosen

sensitivity value is interpreted as the mean value of a Gaussian distribution with a configurable standard deviation. The antenna and chip impedance are used in the simulation to compute the tag radar cross-section

$$\sigma = \frac{\lambda^2 G^2 R_a^2}{\pi |Z_a + Z_c|^2}, \tag{5.2}$$

where λ denotes the wavelength and G is the angle dependent antenna gain [153]. The radar cross-section is the primary factor that determines the backscattered power and is therefore required to compute the RSSI value during a particular tag inventory. The implemented tag model abstracts from the signal and command level and hence does not require the state machine implementation defined in the EPCglobal standard.

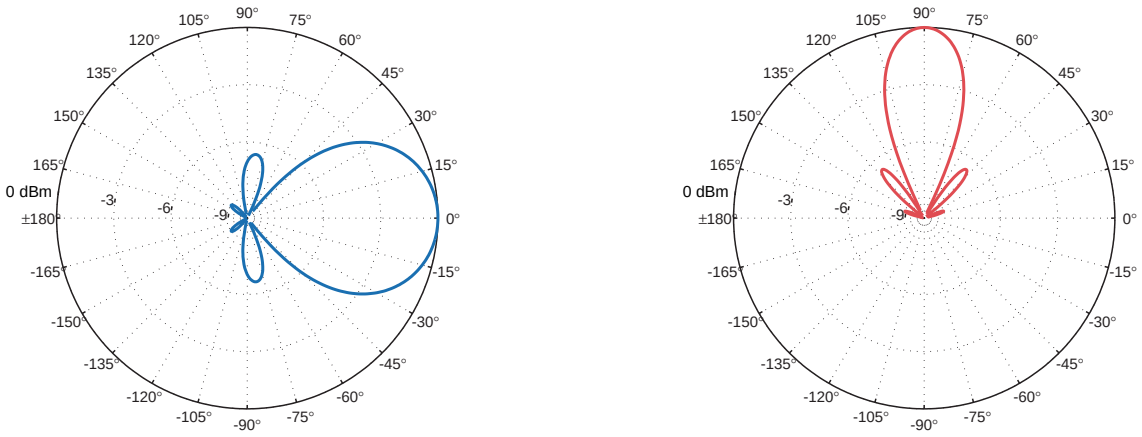
5.3.2 RFID Reader and Antenna

The implemented reader model consists of several blocks that cover different operational aspects. Considering the RF-frontend, RFID readers are equipped with a certain number (default: four) of antenna ports that are used in a time multiplexing scheme. For each port, a distinct transmit power value P_T can be configured. The reader model implements a frequency table according to the European RF regulations [49]. In terms of the backscattered tag response, the reader implements an additive noise model representing thermal noise as well as ambient RF noise. In analogy to current state-of-the art RFID readers, the backscattered signal power for each tag read event is provided in terms of an integer dBm value.

From the protocol perspective, the reader model implements the anti-collision scheme specified in the EPCglobal standard [48]. The key parameter for the the anti-collision scheme is the number of expected tags $M = 2^Q$. This quantity specifies the number of slots that are initiated in the framed slotted ALOHA protocol and therefore determines the duration of an inventory round [184]. The timing is specified by means of the data rate, in particular the Tari (Type A Reference Interval) value which determines the symbol length [48].

Antennas are connected to the reader by means of a cable with a length-dependent attenuation. In contrast to the tag model where matching is explicitly considered, the reader / antenna model is idealized and does not account for a potential mismatch. Instead, antennas are solely characterized by their radiation patterns. In particular, a standard dipole antenna model and a directional antenna model with an approximate half power beamwidth of 70° and 30°, respectively as shown in Figure 5.3 are implemented. Antennas are assumed to have circular polarization, which introduces a polarization mismatch in the link budget between reader and tag. The polarization mismatch is assumed to be constant with -3 dB for the forward and the reverse channel link. The implemented RFID reader and antenna models feature a configurable defect probability P_{Defect}

which allows us to simulate system failures in large scale RFID deployments.



(a) Horizontal **(b)** Vertical

Figure 5.3: Horizontal and vertical radiation pattern of the implemented antenna model with an approximate half power beamwidth of 70° and 30°, respectively for the horizontal and vertical axis. In both directions, the radiation pattern is characterized by a prominent main lobe and a number of additional side lobes.

5.3.3 Detection model

The detection model is based on the simplifying assumption of a free-space propagation. For this reason, the path loss

$$\mathrm{PL} = \left(\frac{\lambda}{4\pi d}\right)^2 \tag{5.3}$$

is computed according to the standard Friis equation. Instead of modeling the communication on signal level with the appropriate waveforms, the detection probability for a particular tag-antenna pair is evaluated by means of the forward and reverse channel link budgets [116]. The values for the path loss and antenna directivity are computed according to the given geometry, which is assumed to be stationary for the duration of an inventory. The link budget is then evaluated in terms of the individual tag and reader sensitivity values. If a tag is visible to the reader and successfully identified during the anti-collision scheme, the RSSI value is computed according to

$$r = ||P_T G_T^2 (PL)^2 \sigma||, \tag{5.4}$$

where $||\cdot||$ denotes the quantization operator.

5.3.4 Sensors

In addition to RFID related components, PRISE also implements a set of devices commonly found in automated environments, such as barcode scanners and proximity sensors. These devices are characterized by a maximum range (default: $1\,\mathrm{m}$) and a defect probability. To detect the presence of objects in the simulated environment, the movement engine continuously monitors if a barcode reader or sensor is triggered by a particular object. The simulated barcode scanner output

$$\mathbf{b}_{\text{Sim}} = [t,\ \mathcal{P}_i], \tag{5.5}$$

comprises a timestamp t and the identifier of a particular packaging unit $\mathcal{P}_i$. In contrast, proximity sensors are limited to a digital output

$$\mathbf{s}_{\text{Sim}} = [t,\ u], \quad u \in \{0,\ 1\}, \tag{5.6}$$

where $u = 1$ indicates that the proximity sensor is triggered. Both barcode scanners and proximity sensors are idealized with respect to duplicate readings and signal debouncing.

5.4 Limitations

As stated in the introduction of this chapter, the PRISE simulation engine is designed for a high-level RFID system simulation. For this reason, there is an inherent need for simplifications in the implemented models to reduce the computational complexity. PRISE is not designed as a profound channel, protocol, tag, or reader simulator since there exists a variety of software frameworks solely focusing on these aspects of RFID systems. Instead, PRISE establishes a compromise between an adequate high-level supply chain model and a reasonable approximation for low-level RFID observations. Consequently, there are several simplifications which are required to reduce the computational complexity.

The most important simplification in terms of the wireless channel is the assumption of a free-space propagation environment. This means that multipath propagation due to reflective materials is not explicitly considered. False positive observations are hence introduced by means of non ideal antenna radiation patterns or overlapping interrogation zones among readpoints. In addition, the channel model does not consider particular item properties (such as metal or water content) which, in practice, heavily influence the link budget and lead to detuning phenomena. Similarly, detuning effects introduced by tag-to-tag coupling are also not considered. The simplified channel model also imposes a limitation to the accuracy of the resulting RSSI responses, since the link

budgets do not consider material dependent attenuation.

In terms of reader capabilities and low-level signal features, the main simplification is that the phase response is not simulated. The reason for this is highlighted in Chapter 4: Although the phase information provides interesting possibilities from a theoretical point of view, the practical relevance for the purpose of classification is limited by the inherent noise.

Due to the modular software architecture, the limitations discussed above can be easily addressed, e.g., by integrating a more detailed channel or item model. For the research issues addressed in this thesis, the current implementation provides a fast and efficient method to generate large scale datasets for different supply chain structures and readpoint setups.

5.5 Experimental validation

For the experimental evaluation of the implemented models, we focus on the characteristics of the simulated RFID observations. In particular, the read rate and the RSSI response characteristics are analyzed. For this purpose, we define three different scenarios that are set up in a lab environment and the simulation framework. As a result, we obtain experimental and simulated datasets which allows us to perform a direct verification. In the following, the datasets and quantities from the real-world setup are marked with the subscript *Exp*, whereas the corresponding quantities from the simulated data carry the subscript *Sim*. The experimental setups for the evaluation comprise an Impinj Speedway Revolution R420 reader and Kathrein Widerange $70°/30°$ antennas. The RFID reader is connected to a host PC which controls the operation modes and stores the individual tag observations for an offline evaluation.

Scenario 1 covers a single RFID reader with an antenna that is used to identify a varying number of stationary tags in the interrogation zone. This scenario is mainly designed to evaluate the temporal behavior and the anti-collision scheme in terms of the read rate ν. *Scenario 2* features the conveyor belt scenario which we have already described in Section 4.1.2 for the identification of items in cardboard boxes. The first two scenarios are characterized by a small distance between reader antenna and tags, which means that the RFID tags are typically in the antenna near-field. The second scenario is especially chosen due to its practical relevance for logistic applications. Finally, *Scenario 3* covers an RFID reader with a total of four antennas in an EAS application. This is a typical far field scenario which is characterized by a considerable variability in terms of item trajectories and the varying orientation between tag and reader antenna. The defect probability of all components is set to $P_{\text{Defect}} = 0$ in all three scenarios.

5.5.1 Scenario 1 - Stationary tags

This scenario is used to evaluate the temporal behavior of the implemented reader and tag models. For this purpose, a varying number of tags is placed in the interrogation zone such that all tags are continuously visible to the reader, as shown in Figure 5.4. The reader is configured to perform a

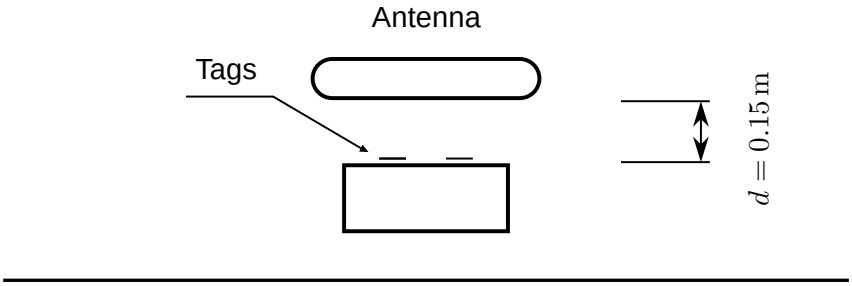


Figure 5.4: Experimental setup scenario 1: A varying number of stationary tags is placed in the interrogation zone such that all tags are continuously visible to the reader. The reader performs a periodic inventory of all tags, and the resulting number of read events is used to compute the read rate ν as a function of the total number of tags.

periodic inventory for a duration of $T = 100\,\mathrm{s}$. Consequently, the tag population is continuously identified, which allows us to evaluate the read rate ν as a function of the total number of tags. For both the experimental setup and the simulated environment, the reader is initialized with an ideal Q- value for the current tag population. The characteristic reader settings are summarized in Table 5.1. The experiment is repeated $N = 10$ times for each tag population and the mean

Component / Parameter	Value
Datarate	80 kbps
Tari	$10\,\mu\mathrm{s}$
Transmit Power P_{TX}	$0.1\,\mathrm{W}$
Number of tags M	$M \in \{1, 2, 5, 10, 15, 20, 30, 40, 50\}$

Table 5.1: Scenario 1: Setup and reader configuration

values of the resulting read rates are estimated together with the corresponding standard deviation. The comparison between the experimental data and the simulation data is shown in Figure 5.5. The simulated read rate shows a characteristic decrease with an increasing number of tags in the interrogation zone. The experimental dataset is characterized by a higher variability but otherwise provides an accurate match to the simulated data. The higher variability is introduced by physical effects that are not explicitly considered in our channel model, such as dead zones introduced by destructive interference.

Regarding the simulation time, we analyze the overall duration of the last experiment with $M = 50$

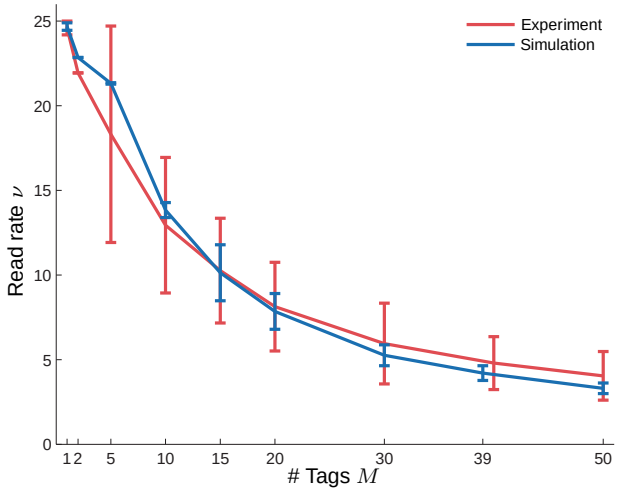


Figure 5.5: Read rate ν (number of RFID observations per tag and unit time) for the simulated and experimental setup over $N = 10$ runs for each tag population. The simulated read rate provides an accurate fit to the empirical data, which means that the timing parameters and the anti-collision procedure are in accordance with the EPCglobal standard. Characteristically, the read rate decreases with an increasing number of tags visible to the reader.

RFID tags. The average duration over the $N = 10$ runs is $\bar{T}_{\text{Sim}} = 3.073\,\text{s}$ on a standard PC with an Intel Core i5-2520M CPU running at $2.5\,\text{GHz}$. The abstraction from the signal level hence provides a considerable speed-up compared to the experimental setup.

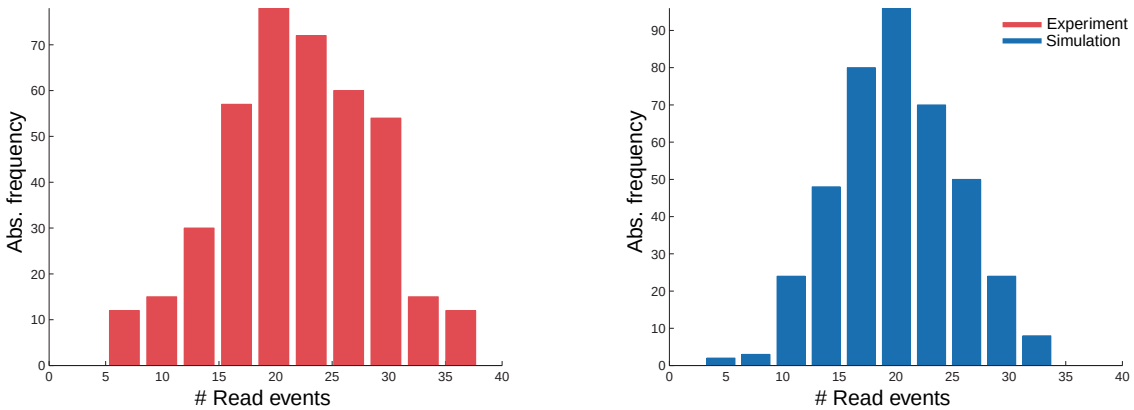
5.5.2 Scenario 2 - Conveyor Belt Setup

The second scenario is of considerable importance for practical applications, since the verification of commissioned packaging units is a key application for RFID systems. The key criteria for the analysis in this case are the low-level signal features, in particular the characteristic signal shape which forms the basis for the signal model in Chapter 4. In addition, the analysis covers the read count per tag as general quality metric. This aspect is investigated by means of the empirical read count distributions over all identified tags. The signal shape is analyzed qualitatively and in terms of the dynamic range (i.e., minimum and maximum RSSI values).

Similar to the evaluation of Scenario 1, this near-field scenario is set up in a test lab and the simulator environment. The configuration parameters are shown in Table 5.2 and the experimental setup is the same as depicted in Figure 4.9. The resulting empirical distributions for the read count are shown in Figure 5.6 with mean values $\bar{n}_{\text{Exp}} = 22.08$ and $\bar{n}_{\text{Sim}} = 20.07$. This means that on average, the experimental setup yields slightly more RFID observations. Furthermore,

Component / Parameter	Value
Datarate	$80\,\mathrm{kbps}$
Tari	$10\ \mu\mathrm{s}$
Transmit Power P_{TX}	$0.3\,\mathrm{W}$
Conveyor velocity v	$0.5\ \frac{\mathrm{m}}{\mathrm{s}}$
Number of tags per box	$M_i = 15$
Total number of tags	$M = 405$
Box-to-box distance d	$0.5\,\mathrm{m}$

Table 5.2: Scenario 2: Experimental conveyor belt setup. Parameters and readpoint configuration



(a) Experiment **(b)** Simulation

Figure 5.6: Empirical distribution of the number of read events for a conveyor belt scenario. The mean values for the read redundancy are $\bar{n}_{\mathrm{Exp}} = 22.08$ and $\bar{n}_{\mathrm{Sim}} = 20.07$, whereas the corresponding values for the standard deviation are $s_{\mathrm{Exp}} = 6.71$ and $s_{\mathrm{Sim}} = 5.17$. The empirical distribution of the experimental data shows a slightly higher mean value and higher tails. The deviations are due to the simplifications in terms of the propagation channel and tag-to-tag coupling effects.

the experimental data exhibits a higher variability (standard deviation $s_{\text{Exp}} = 6.71$ *vs.* $s_{\text{Sim}} = 5.17$), which results in higher distribution tails. The deviations between the experimental and simulated data are mainly due to the simplified channel model. Since multipath propagation is not considered, the simulated dataset lacks a certain percentage of read events that occur over non line-of-sight paths. However, the simplifications in terms of the channel model still allow for a reasonable approximation.

Regarding the quantitative analysis of the simulated low-level features, Figure 5.7 shows an exemplary RSSI response which features a peak resulting from the characteristic tag movement in the interrogation zone. Besides the main peak, the RSSI response exhibits two additional maxima which are caused by the side-lobes of the reader antenna radiation pattern. Regarding the temporal aspect, the response is characterized by non uniform sampling with respect to time due to the non-deterministic inventory procedure as discussed in Chapter 4.

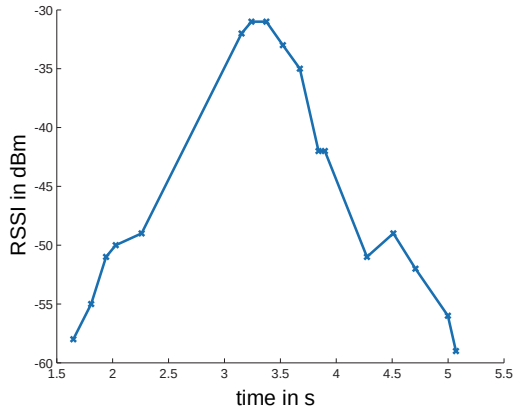


Figure 5.7: Simulated RSSI response: The quantitative comparison between simulated and measured response shows that the implemented models provide an accurate fit in the considered near-field scenario.

	Experiment	Simulation
$\min\{r^{(i)}\}$ in dBm	-66	-68
$\max\{r^{(i)}\}$ in dBm	-28	-28

Table 5.3: Dynamic range of the RSSI response for the experimental and simulated data.

The dynamic range of the empirical and simulated RSSI responses is analyzed by finding the minimum and maximum value over all individual signal values as shown in Table 5.3. The simulated RSSI values are in a comparable range to the empirical data despite the simplified channel model, since the scenario does not include specific materials that influence the channel link budget. For this reason, we can conclude that the implemented models provide a reasonable approximation for this scenario in terms of the generated RFID observations.

5.5.3 Scenario 3 - EAS

Whereas the two previous scenarios are designed to evaluate the simulator performance in near-field applications, Scenario 3 is used to investigate on the far-field behavior. RFID systems for EAS are a popular but challenging application in different fields. The basic task of an RFID readpoint used for EAS is to identify items that are removed from a shop or sales floor without a preceding transaction at the checkout desk. For this purpose, RFID antennas are usually mounted near the shop exit, as depicted in the floor-plan in Figure 5.8. Four antennas are mounted at a height of $z_A = 3\,\mathrm{m}$ and spaced by a distance of $d_x = d_y = 2\mathrm{m}$ in each direction. Hence, the

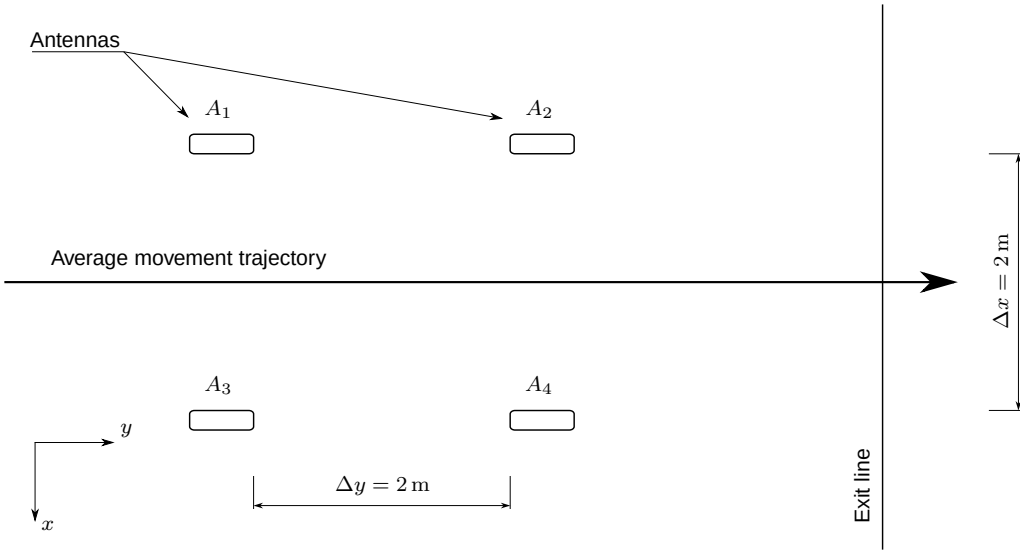


Figure 5.8: Floor plan of the experimental setup for scenario 3: Four RFID antennas are mounted at a height of $z_A = 3\,\mathrm{m}$ to identify tags that leave the shop through the exit door. To account for the variability, the trajectory is specified by a set of randomized way points

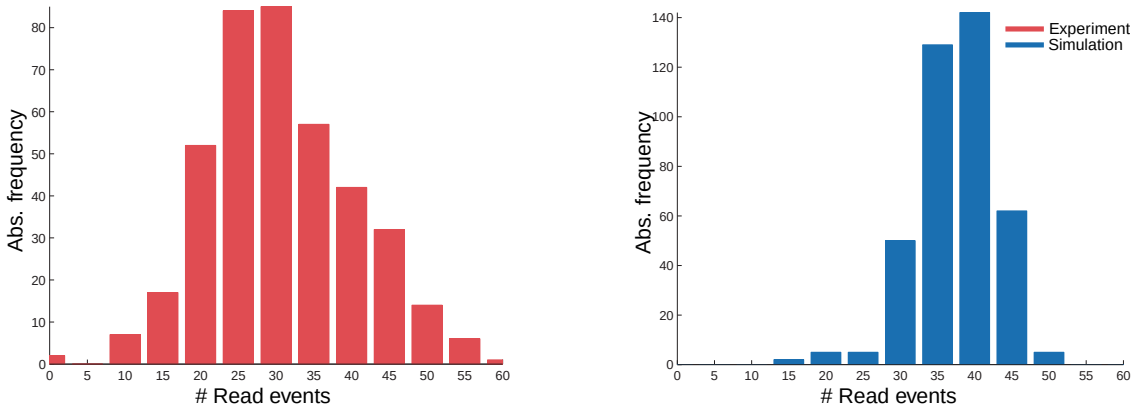
scenario differs from the first two applications in a number of aspects: First, the distance between reader antenna and tag is considerably larger and tags are typically located in the antenna far-field. Second, the link budget suffers from the absorption of RF waves due to the water content in the human body. As a consequence, the channel characteristics are considerably more difficult and result in a decreased detection performance. Third, this scenario shows a considerably greater variability due to the lack of well defined boundary conditions in terms of movement speed and geometric trajectories. Finally, the multi-antenna setup introduces another aspect that needs to be considered in the simulation: RFID-readers use a temporal multiplexing scheme and activate the connected antennas sequentially. For this reason, the carrier signal needs to be switched off and on again for every antenna which introduces an additional time delay compared to the

single antenna case. This delay is a reader dependent property and needs to be considered in the simulation. The experimental setup and the simulation parameters are summarized in Table 5.4. For the experiment, four tagged items are carried in a shopping bag through the interrogation zone in a total of $N = 100$ trials. The empirical and the simulated dataset hence comprise read events from $M = 400$ tags each.

Component / Parameter	Value
Datarate	$80\,\mathrm{kbps}$
Tari	$10\,\mu\mathrm{s}$
Transmit Power P_{TX}	$0.75\,\mathrm{W}$
Item velocity v	$1.5\,\frac{\mathrm{m}}{\mathrm{s}}$
Total number of trials	$N = 100$
Total number of tags	$M = 400$

Table 5.4: Scenario 3: Experimental EAS setup. Parameters and readpoint configuration

The evaluation of the simulated environment incorporates the same metrics as for Scenario 2. In particular, the number of RFID observations and the dynamic range of the RSSI responses are investigated. The corresponding empirical histograms for the read redundancy are shown in Figure 5.9. The simplified channel model is the limiting factor in terms of accuracy for this scenario.

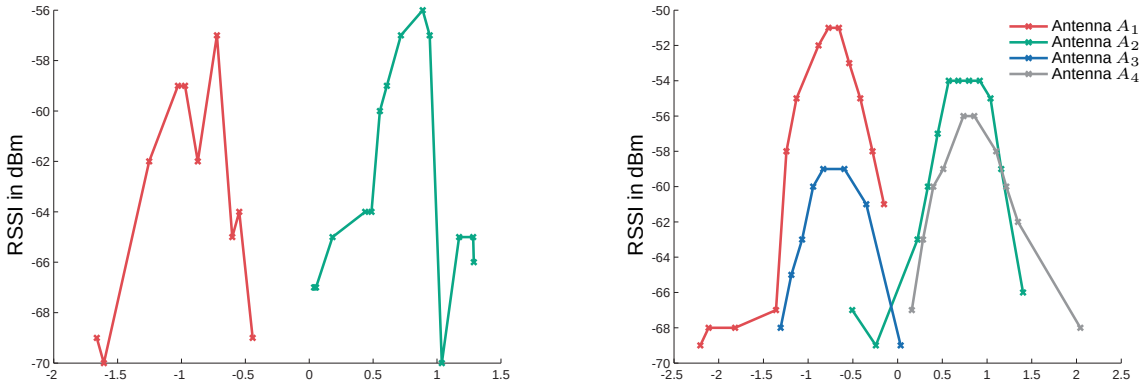


(a) Experiment **(b)** Simulation

Figure 5.9: Empirical distribution of the number of read events for the EAS scenario. The mean values for the read redundancy are $\bar{n}_{\mathrm{Exp}} = 30.61$ and $\bar{n}_{\mathrm{Sim}} = 37.51$, the estimates for the standard deviation are $s_{\mathrm{Exp}} = 9.72$ and $s_{\mathrm{Sim}} = 5.24$. The simplified channel conditions (absence of absorption effects) cause a considerably higher number of read events due to the more optimistic link budget. Furthermore, the experimental data is characterized by a higher standard deviation due to the greater variability.

The simplification results in an overly optimistic link budget and hence yields a considerably higher number of read events. This effect is highlighted when looking at the RSSI responses in

Figure 5.10. Since the human body (as carrier for the tagged items) is not explicitly modeled, there is no absorption of RF waves which increases the tag detection probability and the number of read events. In contrast, the experimental data shows that the shielding effect introduced by the human body leads to a considerable performance degradation. The dynamic range of the sim-



(a) Experiment **(b)** Simulation

Figure 5.10: Experimental (a) and simulated (b) RSSI response, shifted around the temporal center. The simplified channel model leads to an overly optimistic link budget which results in a higher number of RFID observations per tag, since absorption effects are not considered. For this reason, the shown experimental data lacks read events from the opposite antennas A_3 and A_4.

ulated RSSI responses, however, shows an accurate fit to the experimental data. This means that the implemented reader and tag models, together with the backscattering characteristic provide a reasonable approximation also in the far-field. The small deviations on signal level are caused

	Experiment	**Simulation**
$\min\{r^{(i)}\}$ in dBm	-70	-69
$\max\{r^{(i)}\}$ in dBm	-45	-45

Table 5.5: Scenario 3: Dynamic range of the RSSI response for the experimental and simulated data.

by the simplified assumptions, especially in terms of the channel model. Considering the simulator design goals, the implemented models, however, provide a reasonable compromise between accuracy and computational complexity.

5.6 Summary

The presented RFID simulation engine allows us to perform an efficient simulation of large scale RFID systems. This includes the generation of high-level item trajectories through a supply chain, which is the major distinction to existing simulation frameworks. Whereas other simulators focus on particular aspects in an RFID system, PRISE operates at a higher abstraction level which significantly reduces the computational complexity. The evaluation presented in this chapter demonstrates that the simulated RFID observations provide an accurate fit to experimental data, especially in near-field scenarios.

For this reason, PRISE is a powerful tool that can be used to speed up the conception and implementation phase of an RFID system. In particular, the software implementation on the backend side can benefit from the efficient possibility to generate RFID observations in a controlled manner instead of relying on time consuming prototypes and experimental setups. Furthermore, the possibility to generate large scale datasets provides a convenient way to perform profound stress tests for implemented algorithms and database structures.

The current limitations are mainly due to the simplified modeling assumptions which are necessary to reduce the computational complexity. The modular software architecture allows for a seamless integration of more advanced models which provides the possibility to detail on specific aspects if necessary.

6

Information Fusion

Under optimal conditions, RFID systems provide a reliable identification of individual items up to a range of several meters. However, the precise localization of items is a challenging task due to the limited bandwidth and the multipath channel characteristic as we have discussed in Chapter 2. In certain applications such as EAS or adjacently placed warehouse portals, the limited localization capabilities are the major source of noisy observations due to false positive read events. Commercially available EAS systems try to minimize false alarms by specialized antenna designs with an extremely narrow radiation pattern. Whereas this indeed reduces the number of false positive read events, multipath propagation can still lead to unwanted observations which is a problem that cannot be solely solved by antenna design considerations.

The probabilistic view of RFID systems in this thesis provides a considerable increase of data accuracy in RFID systems from the readpoint level up to large scale deployments along a supply chain. The developed mechanisms increase the reliability in terms of detection performance and false positive observations by considering the typical behavior on different abstraction levels. For scenarios like a retail store with an RFID driven article surveillance, this approach has its limitations due to the lack of prior information and the considerable variability. In this case, geometric tag localization is a more promising approach to the problem of false positives. The established probabilistic framework is well suited for the integration of different sensor modalities

which can be used to increase the localization accuracy. This chapter discusses the fusion of computer vision (CV) and RFID systems to provide the required localization accuracy for typical applications. For this purpose, we first review the related literature in this field and provide a comparison of different tag localization systems. Second, we introduce an information fusion approach for tag localization and tracking and provide an experimental evaluation by means of a typical EAS scenario.

6.1 Fusion of CV and RFID

CV systems have progressed to a technological state where they allow for a reliable tracking of individual objects or people from image sequences at low costs. Several authors have investigated on the combination of RFID and CV systems for localization and tracking in different applications. For example, Germa *et al.* [64] have developed a fusion system consisting of an RFID reader and a camera on a mobile robot platform. In particular, the authors combine the sensor information from the CV and the RFID system using a Particle Filter framework to track and follow individual people in a scene. A similar system [18] uses a fixed camera and a moving RFID reader. Nick *et al.* [143–145] have also investigated on the localization of passive UHF tags in combination with a CV system and an Unscented Kalman Filter. In this case, the target is to track moving objects with a fixed RFID reader and camera setup. For this purpose, the authors employ a deterministically found RSS model and a template based object detection to identify and track tagged items in a warehouse portal scenario. In order to further improve the localization accuracy, the authors integrate prior information in terms of a known object height in the scene. The presented studies show a reasonable accuracy for a very specific setup but do not include the case of multiple tags and the related data association problem [127, 166], i. e., the problem of associating the observations from several tags and objects in the scene.

To tackle the problem of a multi-target tracking scenario, a combined CV and RFID system requires four major building blocks as shown in the block diagram in Figure 6.1. First, the RFID system needs to provide at least a rough location estimate $\hat{\mathbf{x}}_{T_i}$ of individual tags. Second, the CV system needs to provide an estimate $\hat{\mathbf{x}}_{B_j}$ for the current location and trajectory of moving objects in the scene. For this purpose, a mechanism to detect and track moving objects (hereafter referred to as blobs) is required. Third, a way to combine the individual location estimates needs to be found, which can be formulated in terms of a data association problem. The combined information can then be interpreted in terms of a high-level reasoning scheme such as a trajectory classification block.

The discussed system architecture provides a flexible way to combine two complementary sensor

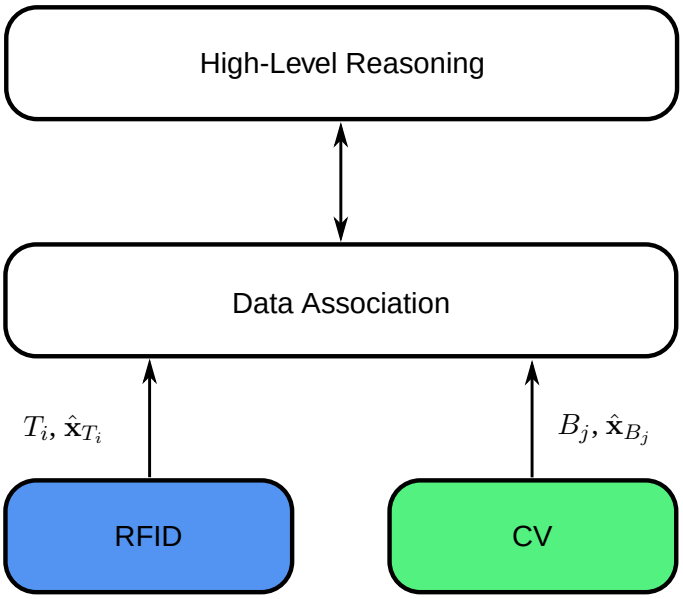


Figure 6.1: Block diagram for a combined RFID and CV localization system. The RFID and CV subsystems detect and localize tags T_i and blobs B_j, respectively. The data association layer subsequently finds the most probable match between identified tags and detected blobs in the scene. Based on the combined information, a high-level reasoning scheme can be employed to identify specific trajectories and provide abstract information.

modalities in a probabilistic framework. This forms the basis for an elegant fusion approach and furthermore enables us to integrate prior information, for example in terms of a floor plan.

6.2 RFID Subsystem

The task of the RFID system in this context is not only to identify RFID tags, but also to provide a location estimate $\hat{\mathbf{x}}_{T_i}$ in order to track individual tags over time. Specifically, the system is required to provide a location estimate for a number of tags $N > 1$. As we have discussed in Chapter 4, the available information for each tag per unit time decreases as the total number of tags increases. For this reason, elaborate localization schemes requiring a high number of read events cannot be employed in this scenario. Instead, we choose a model-based location by proximity approach to provide location estimates for a given number of tags. Naturally, this leads to a compromise between the number of tags that can be handled by the system and the achievable localization accuracy. Since the goal is to combine the information from the RFID system with other sensor modalities, a formulation for the location estimate together with the associated uncertainty is required. The presented probabilistic framework enables us to interpret

the RFID related information in a Bayesian sense and forms the basis for a recursive update scheme when new observation data becomes available.

The idea behind the localization approach is to use several antennas $A_1 \ldots A_K$ that cover a specific region of interest. The antennas are set up such that the individual interrogations zones overlap. For the localization approach, we approximate the individual antenna interrogation zones in the horizontal $x - y$ plane by means of a two dimensional Gauss Kernel g, specified in terms of the antenna position $\boldsymbol{\mu}_i$ and the covariance matrix $\boldsymbol{\Sigma}_i$. Consequently, the entire region of interest can be modeled using a Mixture of Gaussians (MoG), where each mixture component represents one particular antenna. For every RFID observation $\mathbf{z}$, the tag location $\hat{\mathbf{x}}$ can be expressed using the mixture model

$$P(\hat{\mathbf{x}} \,|\, \mathbf{z}) = \sum_{i=1}^{K} w_i g(\hat{\mathbf{x}} \,|\, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i). \tag{6.1}$$

In Equ. (6.1), $w_i \propto \bar{r}_i$ are the weights of the mixture components proportional to measured RSSI values on each antenna. For a simplified scenario with $K = 3$ antennas, a qualitative RFID sensor model is shown in Figure 6.2 with the individual mixture components for a given RFID observation. Since the considered tag is closest to antenna A_1, the RSSI value (and consequently, the weight of the mixture component) for this antenna is dominating over antennas A_2 and A_3.

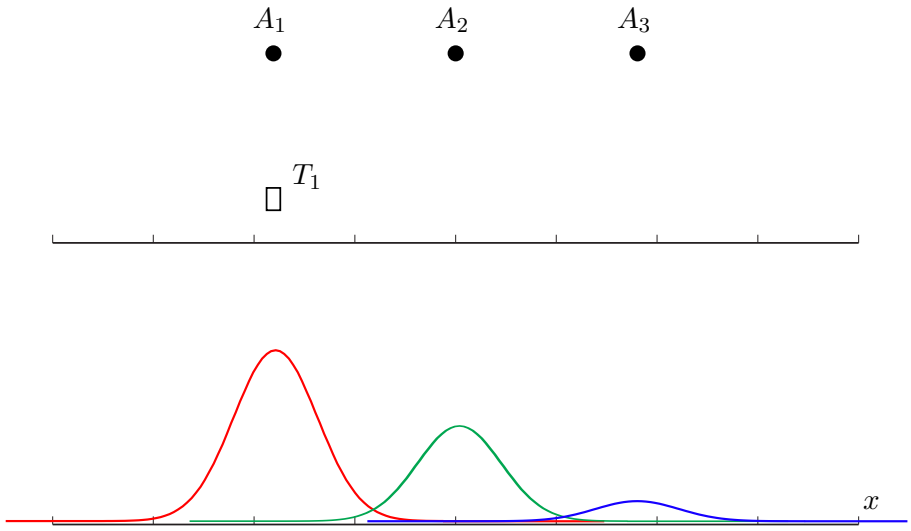


Figure 6.2: RFID sensor model for a simplified 1D case with $K = 3$ antennas. The sensor model incorporates a Mixture of Gaussians (MoG) to provide a likelihood for the location of tag T_1 based on the mean RSSI values. The parameters of the mixture model depend on the antenna characteristics (position and radiation pattern) and need to be learned in a calibration step.

The chosen modeling approach has several implications in terms of complexity, accuracy and

practical aspects. First, the model requires a parameter initialization to estimate the characteristics of the Gaussian mixture components. This parameter estimation needs to be carried out during the deployment phase in order to accurately reflect the characteristics for a given scenario. Second, the resulting abstract representation of the interrogation zone removes the need for an accurate channel model and hence forms a compromise between accuracy and model complexity. Finally, the chosen approach is relatively robust to changes in tag orientation, since we do not use the RSSI values directly to estimate the distance between tag and antenna, but rely on the ratio of the individual RSSI values.

6.3 Blob Detection and Tracking

As complementary sensor modality, the CV system is used to monitor the region of interest in order to detect and track moving blobs. For this purpose, we use a monocular camera with a wide-angle lens to maximize the field of view in a given scenario. For indoor scenarios, it is advantageous to mount the camera on the ceiling to provide a bird's-eye view of the scene. In addition, this allows us to align the camera field of view with the RFID interrogation zone such that we can establish a common reference frame.

The blob detection and tracking mechanism includes several processing steps that are applied to each recorded frame as shown in Figure 6.3. The first block implements a foreground / back-

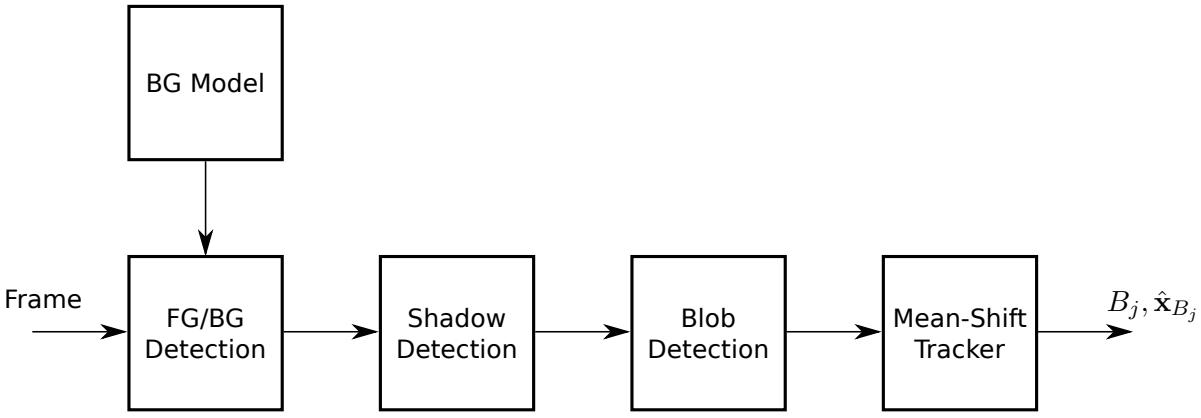


Figure 6.3: Blob detection and tracking: Each frame is first segmented into a foreground (FG) and background (BG) region using a MoG model which represents the RGB values of each pixel $\mathbf{p}$. Subsequently, a shadow detection mechanism is applied to suppress false positive blobs caused by moving shadows in the scene. The actual blob detection incorporates several constraints regarding blob size and velocity. Each detected blob is tracked over consecutive frames using the mean-shift algorithm. Consequently, the blob detector provides an annotated trajectory for every detected blob in an image sequence.

ground segmentation which is based on a MoG model representing the RGB (red, green, blue) values of each pixel $\mathbf{p}$ [201, 202]. This segmentation scheme requires an initial training phase with a static scene to establish a background model BG. Based on the learned model, we process a new frame and classify each pixel by computing the likelihood ratio

$$R = \frac{P(\mathrm{BG} \,|\, \mathbf{p})}{P(\mathrm{FG} \,|\, \mathbf{p})} = \frac{P(\mathbf{p} \,|\, \mathrm{BG}) P(\mathrm{BG})}{P(\mathbf{p} \,|\, \mathrm{FG}) P(\mathrm{FG})} \lessgtr \gamma_R, \tag{6.2}$$

and comparing the value to a threshold γ_R. Since there is usually no prior information about foreground objects, a pixel is assigned to the background if the likelihood

$$p(\mathbf{p} \,|\, \mathrm{BG}) > c_\mathrm{t} \tag{6.3}$$

exceeds a certain threshold c_t. The segmentation scheme implements an online update mechanism and can therefore adapt to slow illumination changes in the scene. This is of particular importance for practical applications without strictly defined lighting conditions. In addition, a shadow detection block [162] is used to suppress false positive blobs caused by moving shadows. This mechanism is based on a non-parametric approach which introduces the two additional classes *highlighted* and *shadowed* background for the classification of each pixel.

The actual blob detection operates on the segmented foreground region and incorporates constraints regarding the size and velocity of blobs. To track the individual blobs over a given image sequence, a Mean-Shift tracking algorithm is employed which is based on the histogram of gray scale values in the blob region. The histogram is updated with every new frame and the blob center $\hat{\mathbf{x}}_{B_j}$ is identified as the mode of the histogram. Hence, for each blob B_j, we obtain an estimate for the current position $\hat{\mathbf{x}}_{B_j}$ in image coordinates and can construct the trajectory $\hat{\mathbf{x}}_{B_j}(t)$ over time. For an exemplary scenario in a lab environment, the camera view with a person detected as blob is shown in Figure 6.4. The visualization shows the estimated blob dimensions (ellipse) as well as the moving direction and velocity (arrow). The presented approach provides an efficient way to identify and track moving objects or people in a scene. The CV system can be easily integrated in an existing RFID deployment and does not require knowledge about the intrinsic camera parameters.

6.4 Data Association

For every processed frame, the two described subsystems provide a set of tags $\mathbf{T}$ and a set of blobs $\mathbf{B}$, incorporating the identifier and location estimates for every tag and blob, respectively.

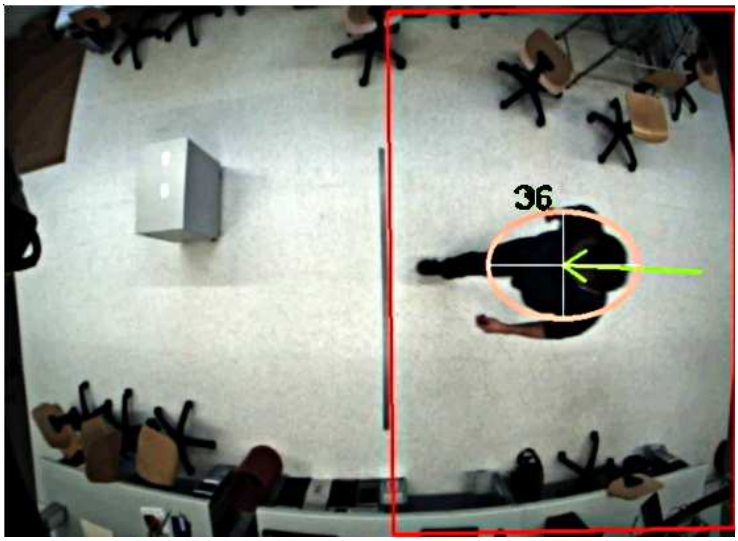


Figure 6.4: Blob detection and tracking: Camera view of a lab environment. In this scenario, a moving person is identified as blob with ID 36. The arrow visualizes the movement direction and velocity.

From these two sets, we want to establish an assignment between individual tags and blobs such that a subset

$$(\mathbf{T}_i \subset \mathbf{T}) \rightarrow B_j \tag{6.4}$$

is assigned to a particular blob B_j. Furthermore, we need to consider the possibility of stationary tags, i.e., tags that belong to the scene background rather than a particular blob. The described problem can be formulated in a data association context which considers the spatial distance $d_{i,j} = \sqrt{x_{i,j}^2 + y_{i,j}^2}$ between tag T_i and blob B_j, as shown for one tag and two blobs in Figure 6.5.

The spatial distance can be transformed into a probability measure

$$p_{i,j} = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(\frac{-x_{i,j}^2}{2\sigma_x^2} + \frac{-y_{i,j}^2}{2\sigma_y^2}\right) \tag{6.5}$$

using a Gauss Kernel with zero mean and a specific covariance

$$\mathbf{\Sigma}_d = \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{pmatrix} \tag{6.6}$$

according to the localization uncertainty of the RFID system. In addition, we explicitly consider the possibility of stationary tags in terms of a velocity based background model. In particular, we

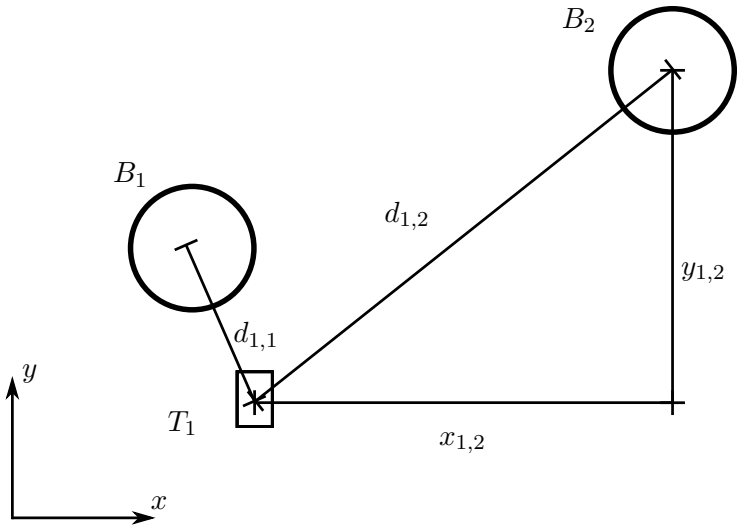


Figure 6.5: Data association problem for a tag T_1 and two blobs B_1, B_2. The goal is to find the most likely assignment between this tag and the detected blobs B_j. For this purpose, the spatial distance $d_{i,j}$ is estimated and transformed to a probability measure $p_{i,j}$ by means of a Gauss Kernel. In addition, we account for the dynamic of tag-blob assignments by incorporating the association history. Stationary tags are considered using a velocity based background model.

estimate the tag velocity

$$\hat{\mathbf{v}}_i = \frac{\mathrm{d}\hat{\mathbf{x}}_i(t)}{\mathrm{d}t} \tag{6.7}$$

as derivative of the trajectory with respect to time. In analogy to the spatial distance, the velocity is transformed to a probability measure by means of a Gauss Kernel with zero mean and a covariance matrix $\mathbf{\Sigma}_v$. For each observed tag, we can hence build an assignment matrix

$$\mathbf{M} = \begin{array}{c} \\ T_1 \\ T_2 \\ \vdots \\ T_N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{BG} & B_1 & B_2 & \cdots & B_M \end{array} \\ \begin{pmatrix} p_{1,\text{BG}} & p_{1,1} & p_{1,2} & \cdots & p_{1,M} \\ p_{2,\text{BG}} & p_{2,1} & p_{2,2} & \cdots & p_{2,M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{N,\text{BG}} & p_{N,1} & p_{N,2} & \cdots & p_{N,M} \end{pmatrix} \end{array} \tag{6.8}$$

holding the individual probability measures for each tag $\leftrightarrow$ blob pair and the background BG. A tag T_i can then be assigned to the most likely class (blobs B_j or background) by finding the maximum value in each row.

The quality of the data association depends on the RFID observations and hence needs to be tolerant to noisy and incomplete data. To provide the required robustness, we integrate the history

of previous assignments by considering the individual blobs and the background in a discrete state-space setting

$$\mathcal{B} = [\text{BG } B_1 \; B_2 \; ... \; B_M]. \tag{6.9}$$

The idea is that the tag $\leftrightarrow$ blob assignment is rather stationary, i.e., a particular tag T_i is typically connected to a blob B_j for a given observation period. Based on the defined state-space, we can hence employ a HMM $\lambda_A = (\boldsymbol{\pi}, \; \mathbf{A}, \; \mathbf{B})$ to filter the assignments. Since there is usually no prior information available, the prior state distribution $\boldsymbol{\pi}$ represents a uniform distribution over all blobs and the scene background. The requirements for the transition model $\mathbf{A}$ are twofold: First, it needs to consider the discussed stationary assignment characteristic by means of appropriate self-transition probabilities. Second, the transition model also needs to allow for a transition from one blob to another, for example when a tag is handed over. For this reason, a compromise between a robust assignment and the capability to follow the scene dynamics needs to be found. The observation model $\mathbf{B}$ accounts for the uncertainty regarding missing RFID observations and the limited localization accuracy.

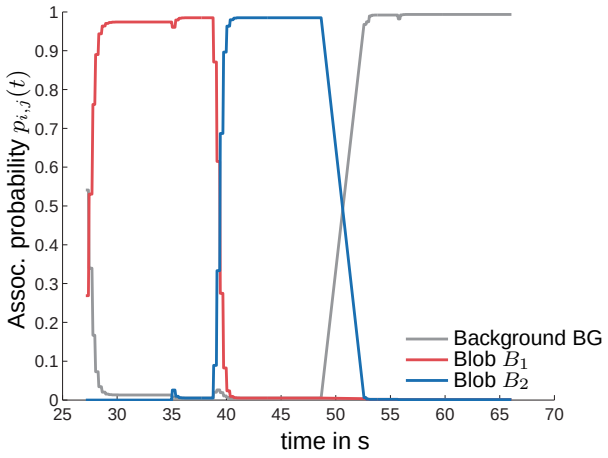


Figure 6.6: Data association problem: Filtered association probabilities in a scene with one tag and two blobs. The scenario is as follows: The considered tag is first carried by blob B_1 and then handed over to blob B_2. Finally, the tag is left stationary in the region of interest while both blobs leave the scene. The HMM based filtering approach adds a low-pass characteristic to the association probabilities and hence provides a certain robustness to noisy and incomplete observations.

Based on the described HMM, we can filter the estimated assignments by means of a recursive Bayes update which is implemented in the Forward algorithm [163]. An example for the filtered assignment probabilities $p_{i,j}$ over time for the exemplary scenario with one tag and two blobs is shown in Figure 6.6. In this scenario, the tag is first carried by blob B_1 and then handed over to blob B_2. Finally, the tag is left stationary in the region of interest. The association probabilities

show a smooth behavior since the HMM filter introduces a certain low-pass characteristic based on the history of previous associations. This provides a considerable robustness for the data association problem and reduces the negative impact of noisy observations.

6.5 Calibration

The combination of the two different sensor modalities requires a calibration procedure consisting of two main steps. First, a common reference frame for the RFID and the CV system needs to be established. For this purpose, the antenna and camera positions $\boldsymbol{\mu}_i$ as well as the camera's field of view need to be determined. Second, the characteristics of the Gaussian mixture components, represented by the covariance matrices $\boldsymbol{\Sigma}_i$ need to be estimated. These characteristics directly reflect the spatial extend of the antenna interrogation zone and are the crucial parameter for the location by proximity approach.

To estimate the characteristics of the interrogation zone, a 2D tag grid in the region of interest is required for which RFID read events are recorded over a specific time frame. Figure 6.7 shows a schematic representation of the calibration setup, consisting of the RFID system and an equidistant tag grid. From the read events, we obtain an observation statistic for every tag-antenna pair.

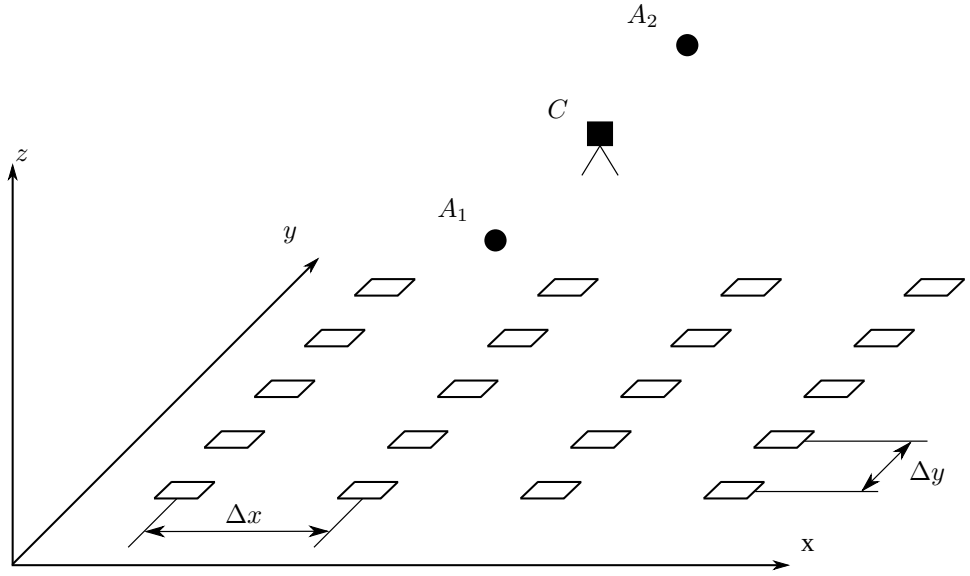
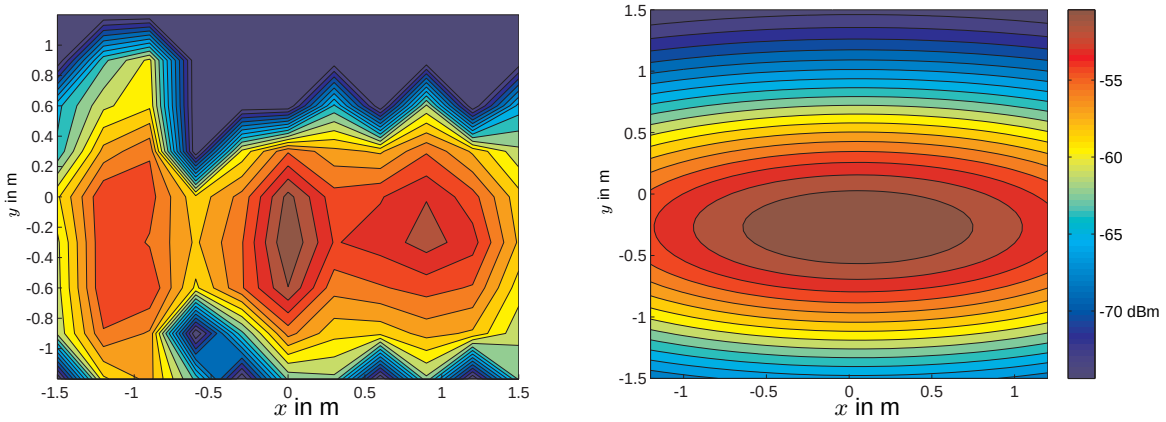


Figure 6.7: Schematic representation of the calibration setup. A tag grid with equidistant spacing is used to obtain an observation statistic for each tag-antenna pair. The observation statistic comprises the mean RSSI value and the corresponding standard deviation. The statistic is used to estimate the parameters of the Gaussian mixture model that represents the spatial extend of the interrogation zone.

In particular, the statistic comprises the mean RSSI and the corresponding standard deviation for each tag and antenna. Since the K antenna positions $\boldsymbol{\mu}_i$ are known with respect to the reference frame, we need to fit a set of K Gauss Kernels to the recorded calibration data. This directly yields an estimate for the covariance matrices $\boldsymbol{\Sigma}_i, i = 1 \dots K$. The measured interrogation zone and the corresponding Gaussian approximation for a single antenna (Kathrein Widerange with 70° / 30° half power beamwidth) are shown in Figure 6.8. The measured interrogation zone exhibits local maxima and minima (dead zones) as a direct consequence of the antenna radiation pattern and the multipath channel environment. Furthermore, the interrogation zone shows a certain asymmetry due to reflections caused by a concrete wall located at $x = -1.5\,\mathrm{m}$. Figure 6.9 shows the mea-



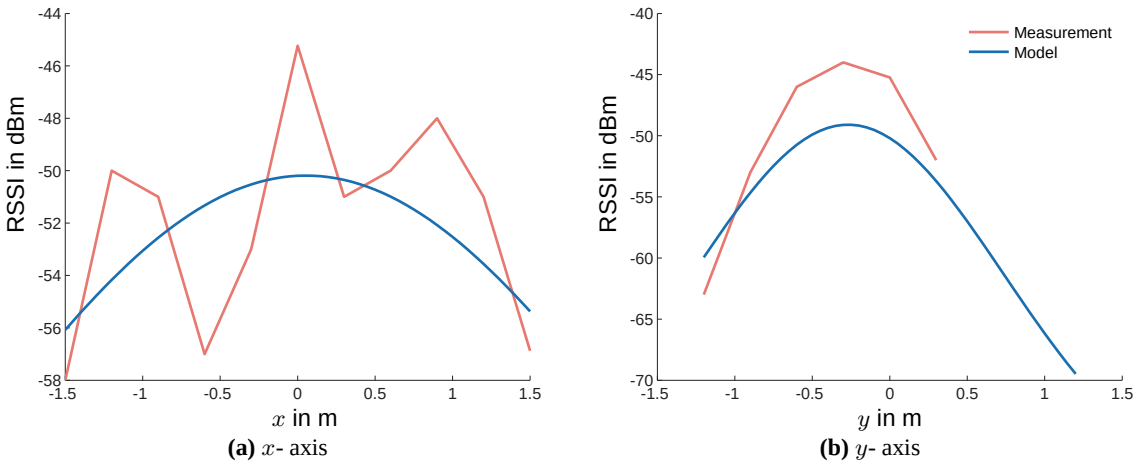
(a) Measurement **(b)** Model

Figure 6.8: RFID sensor model: Contour plot. (a) shows the measured interrogation zone (RSSI values) for a $30 \times 30\,\mathrm{cm}$ tag grid and (b) shows the approximation by means of a 2D Gauss Kernel. The considered system comprises a Kathrein Widerange antenna with 70° / 30° half power beamwidth, mounted at a height of $h = 2.5\,\mathrm{m}$.

sured and modeled interrogation zone for the principal antenna axes in x- and y- direction. The antenna radiation pattern which is characterized by a prominent main-lobe and two side-lobes introduces local RSSI maxima in the x- direction. Due to the difference in halfpower beamwidth, the interrogation zone shows a considerably larger extend in the x- direction.

The Gauss Kernel is a reasonable approximation for the interrogation zone and can be efficiently estimated from the calibration data. In a setup with K equivalent antennas, the combined sensor model can be obtained as a superposition of the individual Gaussian components. This allows for a concise formulation of the likelihood function for the tag location in terms of the mixture model.

The presented information fusion concept provides a concise, yet powerful way to incorporate the two different sensor modalities. To study the performance in a practical environment, the next



(a) x- axis **(b)** y- axis

Figure 6.9: RFID sensor model: RSSI values along the principal antenna axis for the x-direction in (a) and the y- direction in (b), together with the Gaussian approximation. Along the x- axis, the interrogation zone shows a characteristic peak and two local maxima from the side-lobes of the antenna radiation pattern.

section presents a case study for an EAS scenario.

6.6 Case Study: EAS

To evaluate the presented information fusion concept, we investigate on an EAS scenario which is a typical application for RFID systems in retail. The main challenge for RFID driven EAS systems is the occurrence of observations that trigger a false alarm. Due to the lack of prior information and the inherent variability, the low-level approach of geometric tag localization is promising to significantly reduce the percentage of false positives in this scenario.

The functional requirement for an EAS system is to trigger an alarm as soon as an article leaves the shop without a preceding transaction at the checkout desk. With the presented localization and tracking capability, the idea is to track the tag in the region of interest and trigger the alarm as soon as it crosses an imaginary line marking the shop exit. More generally, we can define an *exit region* which a moving tag is not allowed to enter. In addition to the localization and tracking task, this adds a high-level reasoning step which can be elegantly integrated in the proposed framework. In contrast to current RFID driven EAS solutions, this provides a considerable degree of robustness in practical applications.

From the physical point of view, not only the localization, but also the tag detection is challenging in this scenario. As briefly discussed in Chapter 5, the limiting factor in this case is the forward

link budget. The power transferred from the reader to the tag suffers from the absorption due to the water content in the human body and arbitrary tag orientation. The evaluation is hence focused on the two key performance metrics for an RFID system, the detection and false positive probability, rather than the geometric localization uncertainty. For the detection probability, we define two measures: On the tag-level, the detection probability P_{D} is defined as the ratio of correctly detected tags over the total number of tags that move through the scene and enter the exit region. On the test-run level, $P_{\mathrm{D,\,Run}}$ denotes the percentage of runs where at least one tag has been successfully identified as stolen. The false positive probability denotes the ratio of false positive observations over the number of stationary tags in the scene.

The evaluation is based on an RFID system comprising an Impinj Revolution R420 reader and $K = 3$ Kathrein Widerange antennas together with an monocular camera with unknown intrinsic parameters. The camera features an $1/3$" Aptina CMOS sensor, coupled with a $1.8\,\mathrm{mm}$ wide-angle lens to provide an appropriate field of view. The chosen image resolution is $752 \times 480\,\mathrm{px}$ at a framerate of $100\,\mathrm{fps}$. For the positioning of the individual system components, several aspects need to be taken into account. First, the crucial part in terms of RFID localization is the placement of the individual RFID antennas to provide a good coverage over the region of interest with a sufficiently high detection probability. Second, the CV system needs to be placed such that the camera field of view is aligned with the antenna interrogation zones. This enables us to define a common reference frame and estimate the RFID sensor model parameters by means of the proposed calibration step. Finally, there are certain restrictions regarding the antenna placement due to shop design and spacing considerations. For the evaluation, we follow the common approach to mount the RFID antennas on the ceiling. The floor plan of the geometric setup and the camera view of the evaluation environment are shown in Figure 6.10 and Figure 6.11, respectively.

For the evaluation, we define different scenarios with an increasing level of complexity. First, we investigate on the detection performance in a single person, single tag scenario. This scenario gives insight to the combined detection performance of the blob detector and the RFID system and does not consider false positive observations, since no stationary tags are placed in the region of interest. The second and third part of the evaluation cover a single person, multiple tag scenario with and without stationary tags in the region of interest. Finally, we evaluate a multi person, multi tag scenario with stationary tags. In each scenario, the tags are carried through the region of interest using a standard shopping bag along different, randomly displaced trajectories. The stationary tags are placed in the region of interest such that they are continuously visible to the RFID system.

For the first scenario (single person, single tag), we perform $N = 150$ trials to estimate the

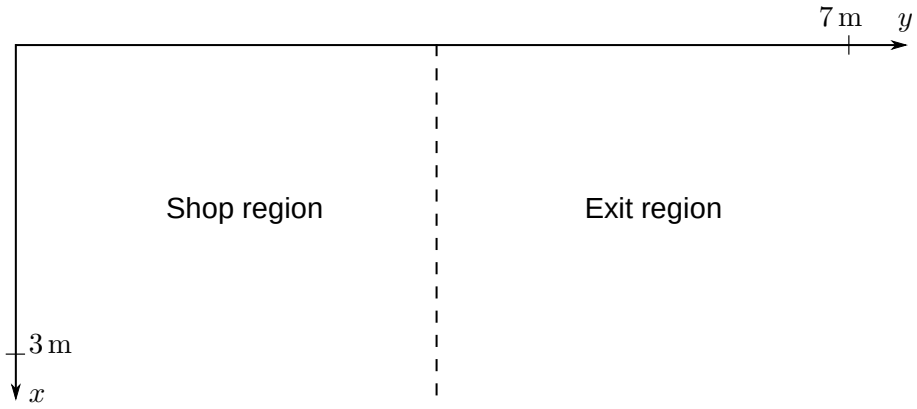


Figure 6.10: EAS scenario: Floor plan. The considered scenario features a $3 \times 7\,\mathrm{m}$ region of interest, split into a *shop region* and an *exit region*. The task of the EAS system is to trigger an alarm as soon as a tag enters the exit region without a preceding transaction at the checkout desk. In addition to the localization and tracking of individual RFID tags, this requires a high-level reasoning step to decide whether a tag has crossed the imaginary line to the exit region.

Figure 6.11: EAS scenario: Camera view with highlighted *exit region*. The camera view shows a considerable distortion due to the use of a wide angle lens which is required to provide an adequate field of view.

detection probability. The resulting performance is

$$P_{\text{D, Run}} = P_{\text{D}} = 0.9733$$

which means that the combined RFID and CV system missed as little as four tags. In this scenario, all tags have been detected by the RFID system, but could not be identified as stolen. This is caused by the CV system failing to identify a moving blob or the background model which erroneously considers a particular tag as stationary.

For the second scenario (single person, multiple tags) we perform an evaluation with $N' = 2$ and $N' = 4$ tags per run. In this case, the results are

$$N' = 2: \quad P_{\text{D, Run}} = 1 \quad P_{\text{D}} = 0.975 \tag{6.10}$$

and

$$N' = 4: \quad P_{\text{D, Run}} = 1 \quad P_{\text{D}} = 0.9125 \tag{6.11}$$

These results agree to the intuitive assumption that the detection probability depends on the scenario complexity, which in this case is determined by the overall number of tags. This is due to the fact that the number of read events per tag and unit time decreases as the number of tags increases. The number of read events has direct impact on the localization accuracy and hence on the overall system performance. For the considered scenario with $N' = 4$ tags, the achieved detection probability of over 90% allows for an efficient system operation since the theft event is detected for all performed test runs.

To further increase the scenario complexity, we add $N_S = 4$ stationary tags to the region of interest. Consequently, we can evaluate the detection and false positive probability. The resulting performance for a total of 80 test runs with $N' = 2$ tags per run is

$$P_{\text{D, Run}} = 1 \quad P_{\text{D}} = 0.9312 \quad P_{\text{FA}} = 0.0188, \tag{6.12}$$

which means that 149 tags could be identified as stolen with six false alarms caused by the stationary tags. Although the detection probability on the tag level decreases, the actual theft event is robustly detected by the proposed system.

Finally, the multi person, multi tag scenario features two people in the region of interest: Whereas one person is walking around at random, the other carries $N' = 2$ tags per run. For this scenario, the resulting performance is

$$P_{\text{D, Run}} = 1 \quad P_{\text{D}} = 0.8438 \quad P_{\text{FA}} = 0.0219. \tag{6.13}$$

The performance decrease compared to the previous cases is mainly due to the complexity in the multi person scenario which imposes an additional challenge to the data association layer. This is especially the case when the two people in the scene are closely spaced. An erroneous data association is hence the main cause for false positives. For this reason, we can conclude that the system performance is directly affected by the scenario complexity and limited by the RFID system rather than the CV system. Even in this complex scenario, the proposed system provides a considerable detection performance on the tag level and robustly suppresses false positives. Since the increased number of tags also adds a certain kind of diversity, the actual theft event can be reliably detected in the defined setup. The presented information fusion approach can be directly applied to the discussed EAS scenario and shows a robust performance under realistic environmental conditions.

6.7 Summary

RFID tag localization can be considered as the key enabling step for RFID systems in applications that require an accurate location information. The limited system bandwidth together with the multipath channel characteristics impose serious limitations to the achievable accuracy. The information fusion concept presented in this chapter is a powerful approach to combine the localization capabilities of CV systems with the strengths of an RFID system.

With the use of a flexible, yet compact RFID sensor model and a location by proximity approach, we can estimate the location and velocity of individual RFID tags in a considered region of interest. The fusion of RFID and CV data results in a data association problem which we formulate in a probabilistic context. The information fusion approach improves the location estimate for a particular RFID tag and forms the basis for high-level reasoning schemes. The experimental evaluation shows that the presented approach provides a considerable detection probability and is robust to false positives. The performance depends on the complexity of the application scenario and decreases with the number of tags. This is mainly caused by the anti-collision scheme which limits the information per tag and unit time. For the demonstrated EAS scenario, the information fusion approach is superior to currently established systems which only try to minimize the number of false positive observations by means of specialized antenna designs and therefore cannot deal with stationary tags in the interrogation zone.

7 Conclusion

Passive UHF RFID has the potential to provide full visibility in supply chain applications, from the manufacturing plant over distribution networks and retail stores to the end customer. Besides the insights that are gained from the additional information on item level, the ability to uniquely identify a particular product also opens up new possibilities in terms of customer experience and marketing channels.

From the technical point of view, RFID systems are still concerned with debates about *read rates* (detection probabilities) and the widespread problem of false positive observations. The effects leading to missed and false positives are an inherent property of passive RFID systems and thus cannot be solely mitigated by means of advanced reader, antenna or tag design. The consequences of missing and false positive observations are manifold and range from inconsistent and noisy data in the backend system to false alarms in case of an RFID driven EAS system.

In this thesis, we have developed several concepts to target the problem of noisy and inconsistent data based on a top-down modeling approach. In order to deal with the peculiar properties of passive RFID systems, we have investigated on the applicability of probabilistic models to the considered problem. The results of the conducted research are best summarized by answering the research questions that stood at the beginning of this thesis project.

Can the problem of noisy data in RFID systems be mitigated by means of a top-down modeling approach? As we have demonstrated in the course of this thesis, a model based approach on different abstraction levels can be used to increase the data accuracy and thereby the value of an RFID system. The targeted abstraction levels range from high-level supply chain considerations down to the individual readpoint and antenna. By means of comprehensive experimental data from practical deployments, we have shown that the presented modeling concepts considerably increase the performance of an RFID system and reduce the inherent observation noise.

Is a probabilistic framework suitable to deal with the particular properties and heterogeneous components in an RFID system? Considering the structure of RFID systems, the versatile requirements and inherent variability, a probabilistic approach seems to be the only way to incorporate the information provided by the heterogeneous system components. Processes on the different abstraction levels show a considerable variability which motivates and justifies the consideration in a probabilistic context.

7.1 Summary of Contributions

The main contributions of this work follow the discussed top-down view of RFID systems. For the first time, we have presented a probabilistic model for RFID enabled supply chains which considers three important characteristics. First, the model incorporates the *typical* behavior of goods in a supply chain by means of a continuous time motion model and thus provides an elegant way to incorporate prior process information. Second, the model takes into account the RFID system properties in terms of the detection and false positive probability of individual readpoints. The proposed state-space representation forms the basis for a process level localization scheme and can be easily implemented at the readpoint level to evaluate a given RFID observation. Finally, we have integrated the characteristic of logical item units as source of prior information. The different modeling concepts and the underlying assumptions have been evaluated by means of extensive simulations and comprehensive empirical datasets from different applications.

On the readpoint level, we have developed a signal model that is able to describe the characteristic low-level features in a concise way. This signal model forms the basis for an abstract classification scheme which can be employed to identify false positive observations. The consideration on readpoint level is necessary when the underlying application does not provide sufficient prior information that can be integrated in terms of a high-level model. Typical application scenarios are hence the verification of commissioned packaging units where neither the transition, nor the co-occurrence model can be employed. The performed experimental evaluation has provided important insights in terms of the individual signal features and has demonstrated the effectiveness

of the presented classification approach.

The developed RFID simulation engine *PRISE* is a valuable tool for the simulation of large scale RFID deployments. The ability to generate realistic, high-level item trajectories together with low-level RFID observations provides the unique capability to study the structure of RFID enabled supply chains. Furthermore, it allows us to effectively validate filtering mechanisms on the different abstraction levels of an RFID system. The implemented models are a compromise between accuracy on the signal level and computational complexity in order to enable a large scale system simulation. PRISE relies directly on the developed RFID system model and provides a fast and efficient possibility to generate large scale datasets.

Consequently following the probabilistic view of RFID systems, we have presented an information fusion approach for the geometric localization of RFID tags. For this purpose, we combine an RFID and a computer vision (CV) system to localize and track tags in practically relevant scenarios. The RFID system incorporates a location by proximity approach which is significantly enhanced by means of the computer vision system. The probabilistic formulation of the localization problem provides a high degree of flexibility and allows for an easy integration into a specific setup. The evaluation by means of a typical EAS application proved the efficiency of the presented approach in several scenarios with different complexity.

The empirical data analysis conducted in the course of this thesis has provided several important insights. First, we have shown that RFID related applications are characterized by a considerable degree of variability which can be efficiently described in terms of a probabilistic framework. Second, the presented analysis highlights that RFID deployments benefit from well defined boundary conditions, such as a constant movement velocity, defined movement trajectories or a fixed tag orientation. Applications that require highest detection probabilities are hence best set up in an automated environment which can guarantee for defined operating conditions. Finally, the empirical data suggests that an increasing item throughput is a limiting factor for the performance of RFID systems. Although the EPCglobal standard is theoretically capable of identifying several 100 tags per second, practical applications suffer considerably from an increased item throughput.

Since the basic task of an RFID system is to collect information about a particular process, the value of an RFID deployment is directly connected to the quality of the provided information. The ultimate goal is hence to increase the data quality in order to provide a reliable basis for high-level decisions and process optimizations. The presented models and filtering schemes have the main advantage that they abstract from the technology layer towards a more general view. Although the individual models might require slight adaptations to consider new technological aspects, we are confident that the fundamental ideas and mechanisms provide a general and technology independent framework.

7.2 Outlook

Beyond the goal of increasing the data accuracy in RFID systems, we believe that the developed models can be used as basis for an in-depth analysis and monitoring of RFID driven supply chains. Possible applications range from intelligent replenishment reports over accurate order management to the efficiency analysis of individual processing steps. In this context, we see interesting research challenges for the data mining community to extract relevant information from the collected data and draw conclusions from a series of observations.

Similarly, the probabilistic top-down interpretation of RFID systems can be the starting point for the design of smart, interactive processes. By introducing a feedback loop that provides immediate information to the user, processes in manufacturing and logistics can be optimized to increase the flexibility and efficiency. For this reason, Artificial Intelligence and Machine Learning concepts are required which can interpret the gathered data to formulate an adequate feedback.

Regarding the characterization of individual readpoints, fundamental research is required to establish a theory for the tag detection process under different environmental conditions and readpoint setups. Despite the fact that detailed channel models and simulation environments are available, the design of an RFID system is still dominated by extensive experiments and a trial-and-error characteristic. This is mainly due to the fact that available simulators face a considerable computational complexity and hence cannot be directly used to simulate practically relevant scenarios.

On the hardware layer, we expect that advanced reader and antenna systems will provide interesting new features such as antenna beamforming, switching polarization, advanced anti-collision schemes and additional sensing capabilities. Considering the flexibility of the presented concepts, future research should specifically consider the technological advances to further improve the data accuracy and enable new application ideas.

A

Appendix

A.1 PRISE - Example configuration file

```
<?xml version="1.0" encoding="UTF-8"?>
<ProjectConfiguration>
   <ProcessModel>
      <Name>MyProcess</Name>
      <NumberOfStates>1</NumberOfStates>
      <StateDescriptions>
         <State Description="State 1"/>
      </StateDescriptions>
      <ContinuousTimeMarkovChain>
         <InitialProbabilities>
            <pi0>1</pi0>
         </InitialProbabilities>
         <DwellTimes>
            <t0>3600</t0>
```

```
15          </DwellTimes>
16          <TransitionProbabilities>
17             <r0>
18                <c00>1</c00>
19             </r0>
20          </TransitionProbabilities>
21       </ContinuousTimeMarkovChain>
22       <SensorModel>
23          <r0>
24             <c00>0</c00>
25          </r0>
26       </SensorModel>
27    </ProcessModel>
28    <Sites>
29       <Site>
30          <Name>State 1</Name>
31          <Comment></Comment>
32          <Readers>
33             <RfidReader>
34                <Identifier>State 1_Reader_1</Identifier>
35                <DefectProbability>0</DefectProbability>
36                <Dimensions x="0.5" y="0.5" z="0.05"/>
37                <Location x="0" y="0" z="0"/>
38                <Sensitivity>9.03361e-12</Sensitivity>
39                <InitialQ>3</InitialQ>
40                <OperationMode>Miller 2</OperationMode>
41                <DataRate>80000</DataRate>
42                <Tari>10e-06</Tari>
43                <Antennas>
44                   <Antenna>
45                      <Identifier>A000</Identifier>
46                      <Type>Directional3050</Type>
47                      <DefectProbability>0</DefectProbability>
48                      <AntennaImpedance Real="50" Imag="0"/>
49                      <Dimensions x="0.7" y="0.35" z="0.15"/>
```

```
                        <Location x="0" y="1" z="1"/>
                        <LookAt x="0" y="0" z="-1"/>
                        <RfPower>0.1</RfPower>
                    </Antenna>
                </Antennas>
                <Cables>
                    <Cable>
                        <CableType>RG58</CableType>
                        <CableLength>2</CableLength>
                        <Attenuation>20</Attenuation>
                    </Cable>
                </Cables>
            </RfidReader>
        </Readers>
        <Sensors/>
    </Site>
</Sites>
<Trajectories>
    <Trajectory>
        <Name>State 1_t</Name>
        <Velocity>0.0001</Velocity>
        <Variance>0</Variance>
        <waypoint x="-0.005" y="1" z="0.75" yaw="0" pitch="0"
            roll="0"/>
        <waypoint x="0.005" y="1" z="0.75" yaw="0" pitch="0"
            roll="0"/>
    </Trajectory>
</Trajectories>
<SimulationParameters>
    <SimulationEndTime>72000</SimulationEndTime>
    <SimulationTimeResolution>0.001</
        SimulationTimeResolution>
</SimulationParameters>
<ItemTemplates>
    <ItemTemplate>
```

```
82          <TemplateName>Box</TemplateName>
83          <NumberOfItems>1</NumberOfItems>
84          <IdentifierPrefix>Item</IdentifierPrefix>
85          <Dimensions x="0.4" y="0.4" z="0.005"/>
86          <Tags min="5" max="5"/>
87          <TagIdentifierPrefix>Tag</TagIdentifierPrefix>
88       </ItemTemplate>
89    </ItemTemplates>
90 </ProjectConfiguration>
```

Listing A.1: Example configuration file

Bibliography

[1] A. E. Abbas. A Kullback-Leibler view of linear and log-linear pools. *Decision Analysis*, 6(1):25–37, 2009.

[2] C. C. Aggarwal and J. Han. A survey of RFID data processing. In *Managing and Mining Sensor Data*, pages 349–382. Springer, 2013.

[3] C. C. Aggarwal and P. S. Yu. A framework for clustering uncertain data streams. In *Proc. IEEE ICDE*, pages 150–159, 2008.

[4] R. Agrawal, T. Imieliński, and A. Swami. Mining association rules between sets of items in large databases. In *ACM SIGMOD Record*, volume 22, pages 207–216, 1993.

[5] J. Al-Kassab, F. Thiesse, and T. Buckel. RFID-enabled business process intelligence in retail stores: A case report. *Journal of Theoretical and Applied Electronic Commerce Research*, 8, August 2013.

[6] J. J. Alcaraz, J. Vales-Alonso, and J. Garcia-Haro. RFID reader scheduling for reliable identification of moving tags. *Trans. Automation Science and Engineering*, 10(3):816–828, 2013.

[7] C. Alippi, D. Cogliati, and G. Vanini. A statistical approach to localize passive RFIDs. In *Proc. IEEE ISCAS*, pages 843–846, 2006.

[8] F. Alkoot and J. Kittler. Experimental evaluation of expert fusion strategies. *Pattern Recognition Letters*, 20(11):1361–1369, 1999.

[9] D. Allard, A. Comunian, and P. Renard. Probability aggregation methods in geoscience. *Mathematical Geosciences*, 44(5):545–581, 2012.

[10] P. M. Altham. Two generalizations of the binomial distribution. *Applied Statistics*, 27:162–167, 1978.

[11] C. Angerer, R. Langwieser, G. Maier, and M. Rupp. Maximal ratio combining receivers for dual antenna RFID readers. In *Proc. IEEE IMWS*, pages 1–4, 2009.

[12] C. Angerer, R. Langwieser, and M. Rupp. Experimental performance evaluation of dual antenna diversity receivers for RFID readers. In *Proc. EURASIP RFID*, 2010.

[13] D. Arnitz. *Tag Localization in Passive UHF RFID*. PhD thesis, Graz University of Technology, May 2011.

[14] D. Arnitz, U. Mühlmann, T. Gigl, and K. Witrisal. Wideband system-level simulator for passive UHF RFID. In *Proc. IEEE RFID*, pages 28 – 33, Apr 2009.

[15] D. Arnitz, U. Mühlmann, and K. Witrisal. The PARIS simulation framework. *www.tinyurl.com/paris-osf*, 2010.

[16] S. R. Aroor and D. D. Deavours. Evaluation of the state of passive UHF RFID: An experimental approach. *IEEE Systems Journal*, 1(2):168–176, 2007.

[17] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on Particle Filters for online nonlinear/non-Gaussian Bayesian tracking. *Trans. Signal Processing*, 50(2):174–188, 2002.

[18] Z. Babic, M. Ljubojevic, and V. Risojevic. Indoor RFID localization improved by motion segmentation. In *Proc. IEEE ISPA*, pages 271–276, 2011.

[19] M. A. Bart, J. Fligner and W. I. Notz. *Sampling and Statistical Methods for Behavioral Ecologists*. Cambridge University Press, December 1998.

[20] U. N. Bhat and R. Lal. Number of successes in Markov trials. *Advances in applied probability*, 20:677–680, 1988.

[21] R. Bhattacharyya, C. Floerkemeier, and S. Sarma. Low-cost, ubiquitous RFID-tag-antenna-based sensing. *Proc. of the IEEE*, 98(9):1593–1600, 2010.

[22] H. H. Bi and D. K. Lin. RFID-enabled discovery of supply networks. *Trans. Engineering Management*, 56(1):129–141, 2009.

[23] M. Bolic, D. Simplot-Ryl, and I. Stojmenovic. *RFID systems: Research trends and challenges*. Wiley, 2010.

[24] A. Bouzakis and L. Overmeyer. Position tracking for passive UHF RFID tags with the aid of a scanned array. *Int. Journal of Wireless Information Networks*, pages 1–10, 2013.

[25] M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. Van Gool. Robust tracking-by-detection using a detector confidence particle filter. In *Proc. IEEE ICCV*, pages 1515–1522, 2009.

[26] M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. Van Gool. Online multiperson tracking-by-detection from a single, uncalibrated camera. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(9):1820–1833, 2011.

[27] M. Büttner and D. Wetherall. An empirical study of UHF RFID performance. In *Proc. ACM MobiCom*, pages 223–234, 2008.

[28] S. Caizzone and G. Marrocco. RFID grids: Part II–experimentations. *Trans. Antennas and Propagation*, 59(8):2896–2904, 2011.

[29] L. Catarinucci, D. De Donno, M. Guadalupi, F. Ricciato, and L. Tarricone. Performance analysis of passive UHF RFID tags with GNU-radio. In *Proc. IEEE APS/URSI*, pages 541–544, 2011.

[30] H. Chen, W.-S. Ku, H. Wang, and M.-T. Sun. Leveraging spatio-temporal redundancy for RFID data cleansing. In *Proc. ACM SIGMOD*, pages 51–62, 2010.

[31] X. Chen, F. Lu, and T. Ye. Mutual coupling of stacked UHF RFID antennas in NFC applications. In *Proc. IEEE APS/URSI*, pages 1–4, 2009.

[32] T. Cheng and L. Jin. Analysis and simulation of RFID anti-collision algorithms. In *Proc. IEEE ACT*, volume 1, pages 697–701, 2007.

[33] B.-S. Choi, J.-W. Lee, J.-J. Lee, and K.-T. Park. A hierarchical algorithm for indoor mobile robot localization using RFID sensor fusion. *Trans. Industrial Electronics*, 58(6):2226–2235, 2011.

[34] J. S. Choi, H. K. Kang, J. E. Jung, J. E. Kim, and D. H. Lee. Impact of operating frequency on passive UHF RFID based localization system. In *Proc. IEEE ICWITS*, 2012.

[35] J. S. Choi, M. Kang, R. Elmasri, and D. W. Engels. Investigation of impact factors for various performances of passive UHF RFID system. In *Proc. IEEE RFID-TA*, pages 152–159, 2011.

[36] J. S. Choi, B. R. Son, H. J. Park, and D. H. Lee. Consideration for performance of passive UHF far-field RFID tag to tag interference based localization system. In *Proc. IEEE ICWITS*, 2012.

[37] H.-T. Chou, S.-P. Liang, and S.-Y. Wei. Characteristic examination of RFID tag implementation on a vehicle outer surface. In *Proc. IEEE ICWITS*, 2012.

[38] G. Cicirelli, A. Milella, and D. Di Paola. RFID tag localization by using adaptive neuro-fuzzy inference for mobile robot applications. *Industrial Robot: An International Journal*, 39(4):340–348, 2012.

[39] R. H. Clarke, D. Twede, J. R. Tazelaar, and K. K. Boyer. Radio frequency identification (RFID) performance: The effect of tag orientation and package contents. *Packaging Technology and Science*, 19(1):45–54, 2006.

[40] V. Derbek, C. Steger, R. Weiß, D. Wischounig, J. Preishuber-Pflügl, and M. Pistauer. Simulation platform for UHF RFID. In *Proc. IEEE DATE*, pages 918–923, 2007.

[41] J. L. Devore. A note on the estimation of parameters in a Bernoulli model with dependence. *The Annals of Statistics*, 4:990–992, 1976.

[42] T. Deyle, C. C. Kemp, and M. S. Reynolds. Probabilistic UHF RFID tag pose estimation with multiple antennas and a multipath RF propagation model. In *Proc. IEEE IROS*, pages 1379–1384, 2008.

[43] E. DiGiampaolo and F. Martinelli. A passive UHF-RFID system for the localization of an indoor autonomous vehicle. *Trans. Industrial Electronics*, 59(10):3961–3970, 2012.

[44] D. M. Dobkin and S. M. Weigand. Environmental effects on RFID tag antennas. In *Proc. IEEE MTT-S*, pages 135–138, 2005.

[45] Z. Drezner and N. Farnum. A generalized Binomial distribution. *Communications in Statistics - Theory and Methods*, 22(11):3051–3063, 1993.

[46] D. W. Engels and S. E. Sarma. The reader collision problem. In *Proc. Systems, Man and Cybernetics*, volume 3, pages 6–pp, 2002.

[47] EPCglobal Inc. EPC Information Services (EPCIS) Specification, Version 1.0.1, September 2007.

[48] EPCglobal Inc. Class 1 Generation 2 UHF RFID Protocol for Communication at 860 Mhz-960 Mhz, Version 1.2.0, October 2008.

[49] European Telecommunications Standards Institute. ETSI EN 302 208-1 v.1.4.1, November 2011.

[50] K. Finkenzeller. *RFID-Handbuch*. Hanser München, 2002.

[51] C. Floerkemeier and R. Pappu. Evaluation of RFIDSim-a physical and logical layer RFID simulation engine. In *Proc. IEEE RFID*, pages 350–356, 2008.

[52] C. Floerkemeier and S. Sarma. RFIDSim–a physical and logical layer simulation engine for passive RFID. *Trans. Automation Science and Engineering*, 6(1):33–43, 2009.

[53] J. L. M. Flores, S. S. Srikant, B. Sareen, and A. Vagga. Performance of RFID tags in near and far field. In *Proc. IEEE ICPWC*, pages 353–357, 2005.

[54] V. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello. Bayesian filtering for location estimation. *Pervasive Computing*, 2(3):24–33, 2003.

[55] F. Fuschini, C. Piersanti, F. Paolazzi, and G. Falciasecca. Analytical approach to the backscattering from UHF RFID transponder. *Antennas and Wireless Propagation Letters*, 7:33–35, 2008.

[56] F. Fuschini, C. Piersanti, L. Sydanheimo, L. Ukkonen, and G. Falciasecca. Electromagnetic analyses of near field UHF RFID systems. *Trans. Antennas and Propagation*, 58(5):1759–1770, 2010.

[57] K. Fyhn, R. M. Jacobsen, P. Popovski, and T. Larsen. Fast capture–recapture approach for mitigating the problem of missing RFID tags. *Trans. Mobile Computing*, 11(3):518–528, 2012.

[58] K. Fyhn, R. M. Jacobsen, P. Popovski, A. Scaglione, and T. Larsen. Multipacket reception of passive UHF RFID tags: A communication theoretic approach. *Trans. Signal Processing*, 59(9):4225–4237, 2011.

[59] K. Gabriel. The distribution of the number of successes in a sequence of dependent trials. *Biometrika*, 46(3/4):454–460, 1959.

[60] H. Gang Wang, C. Xing Pei, and Q. Pan. Test and analysis of passive UHF RFID readability in multipath environments. In *Proc. WiCom*, pages 1–5, 2009.

[61] A. Garg, T. Jayram, S. Vaithyanathan, and H. Zhu. Generalized opinion pooling. In *Proc. Int. Symp. on Artificial Intelligence and Mathematics*, 2004.

[62] A. Gauby, M. Goller, and C. Peter. Verfahren zum Zuordnen von mindestens einem kontaktlos auslesbaren Datenträger zu mindestens einem sich räumlich bewegenden Bereich und Vorrichtung zum Ausführen des Verfahrens. German Patent No. DE102009016557A1, October 2010.

[63] C. Genest and J. V. Zidek. Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, 1(1):114–135, 1986.

[64] T. Germa, F. Lerasle, N. Ouadah, and V. Cadenat. Vision and RFID data fusion for tracking people in crowds by a mobile robot. *Computer Vision and Image Understanding*, 114(6):641–651, 2010.

[65] D. Girbau, J. Lorenzo, A. Lázaro, and R. Villarino. Effects of antenna detuning and gain penalty on the read range of UHF RFID. In *Proc. RFID SysTech*, pages 1–7, 2010.

[66] M. Goller and M. Brandner. Experimental evaluation of RFID gate concepts. In *Proc. IEEE RFID*, pages 26–31, 2011.

[67] M. Goller and M. Brandner. Improving classification performance of RFID gates using Hidden Markov Models. In *Proc. IEEE I2MTC*, pages 1–5, 2011.

[68] M. Goller and M. Brandner. Increasing the robustness of RFID systems using a probabilistic business process model. Poster Presentation, IEEE RFID, 2011.

[69] M. Goller and M. Brandner. Probabilistic modeling of RFID business processes. In *Proc. IEEE RFID-TA*, pages 432–436, 2011.

[70] M. Goller and M. Brandner. Evaluation of feature attributes for an RFID conveyor belt application using probabilistic models. In *Proc. SmartSystech*, 2012.

[71] M. Goller and M. Brandner. Process-level localization of RFID tags using probabilistic models. In *Proc. IEEE ICWITS*, 2012.

[72] M. Goller, M. Brandner, and G. Brasseur. A system model for cooperative RFID readpoints. *Submitted to the IEEE Trans. Instrumentation and Measurement*, 2013.

[73] J. D. Griffin, G. Durgin, A. Haldi, and B. Kippelen. Radio link budgets for 915 MHz RFID antennas placed on various objects. In *Proc. Texas Wireless Symposium*, pages 22–26, 2005.

[74] J. D. Griffin and G. D. Durgin. Multipath fading measurements for multi-antenna backscatter RFID at 5.8 GHz. In *Proc. IEEE RFID*, pages 322–329, 2009.

[75] J. D. Griffin, G. D. Durgin, A. Haldi, and B. Kippelen. RF tag antenna performance on various materials using radio link budgets. *Antennas and Wireless Propagation Letters*, 5(1):247–250, 2006.

[76] Y. Gu, G. Yu, Y. Chen, and B. C. Ooi. Efficient RFID data imputation by analyzing the correlations of monitored objects. In *Database Systems for Advanced Applications*, pages 186–200. Springer, 2009.

[77] F. Gustafsson. Particle filter theory and practice with positioning applications. *IEEE Aerospace and Electronic Systems Magazine*, 25(7):53–82, 2010.

[78] D. Hähnel, W. Burgard, D. Fox, K. Fishkin, and M. Philipose. Mapping and localization with RFID technology. In *Proc. IEEE ICRA 2004*, volume 1, pages 1015 – 1020, April 2004.

[79] Y. Han, Q. Li, and H. Min. System modeling and simulation of RFID. In *Auto-ID Labs Research Workshop*, pages 24–24, 2004.

[80] H. Hashemi. The indoor radio propagation channel. *Proc. of the IEEE*, 81(7):943–968, 1993.

[81] M. Hisakado, K. Kitsukawa, and S. Mori. Correlated Binomial models and correlation structures. *Journal of Physics A: Mathematical and General*, 39(50):15365, 2006.

[82] M. Holweg and J. Bicheno. Supply chain simulation–a tool for education, enhancement and endeavour. *Int. Journal of production economics*, 78(2):163–175, 2002.

[83] A. Ilic, T. Andersen, and F. Michahelles. EPCIS-based supply chain visualization tool. *Auto-ID Labs White Paper*, 2009.

[84] A. Ilic, T. Andersen, and F. Michahelles. Increasing supply-chain visibility with rule-based RFID data analysis. *IEEE Internet Computing*, 13(1):31–38, 2009.

[85] Impinj, Inc. Monza/ID datasheet, 2007.

[86] Impinj, Inc. Monza 5 datasheet, 2012.

[87] A. Isasi, S. Rodriguez, J. L. D. Armentia, and A. Villodas. Location, tracking and identification with RFID and vision data fusion. In *Proc. RFID SysTech*, pages 1–6, 2010.

[88] R. Jacobsen, K. F. Nielsen, P. Popovski, and T. Larsen. Reliable identification of RFID tags using multiple independent reader sessions. In *Proc. IEEE RFID*, pages 64–71, 2009.

[89] R. Jacobsen, K. F. Nielsen, A. Scaglione, P. Popovski, and T. Larsen. Probabilistic RFID tag detector model. In *Proc. IEEE SPAWC*, pages 1–5, 2010.

[90] S. R. Jeffery, M. Garofalakis, and M. J. Franklin. Adaptive cleaning for RFID data streams. In *Proc. VLDB*, pages 163–174, 2006.

[91] H. Jeung, M. L. Yiu, X. Zhou, C. S. Jensen, and H. T. Shen. Discovery of convoys in trajectory databases. *Proc. VLDB*, 1(1):1068–1080, 2008.

[92] M. Jo, H. Y. Youn, S.-H. Cha, and H. Choo. Mobile RFID tag detection influence factors and prediction of tag detectability. *IEEE Sensors Journal*, 9(2):112–119, 2009.

[93] M. Jo, H. Y. Youn, and H.-H. Chen. Intelligent RFID tag detection using support vector machine. *Trans. Wireless Communications*, 8(10):5050–5059, 2009.

[94] D. Joho, C. Plagemann, and W. Burgard. Modeling RFID signal strength and tag detection for localization and mapping. In *Proc. IEEE ICRA*, pages 3160–3165, 2009.

[95] O. Kabadurmus, S. Kilinc, H. Behret, and G. Uygun. Performance evaluation of operational parameters on RFID controlled conveyor system. In *Proc. RFID Eurasia*, pages 1–6, 2007.

[96] M. Kärkkäinen. Increasing efficiency in the supply chain for short shelf life goods using RFID tagging. *Int. Journal of Retail & Distribution Management*, 31(10):529–536, 2003.

[97] M. Kärkkäinen, T. Ala-Risku, and K. Främling. Efficient tracking for short-term multi-company networks. *Int. Journal of Physical Distribution & Logistics Management*, 34(7):545–564, 2004.

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

[99] T. Keller, F. Thiesse, A. Ilic, and E. Fleisch. Decreasing false-positive RFID tag reads by improved portal antenna setups. In *Proc. IEEE IOT*, pages 99–106, 2012.

[100] T. Keller, F. Thiesse, J. Kungl, and E. Fleisch. Using low-level reader data to detect false-positive RFID tag reads. In *Proc. IEEE IOT*, pages 1–8, 2010.

[101] N. Khoussainova, M. Balazinska, and D. Suciu. Probabilistic event extraction from RFID data. In *Proc. IEEE ICDE*, pages 1480–1482, 2008.

[102] W. Khreich, E. Granger, A. Miri, and R. Sabourin. A comparison of techniques for on-line incremental learning of HMM parameters in anomaly detection. In *Proc. IEEE CISDA*, pages 1–8, 2009.

[103] D.-Y. Kim, B.-J. Jang, H.-G. Yoon, J.-S. Park, and J.-G. Yook. Effects of reader interference on the RFID interrogation range. In *Proc. IEEE EuMC*, pages 728–731, 2007.

[104] D.-Y. Kim, H.-G. Yoon, B.-J. Jang, and J.-G. Yook. Interference analysis of UHF RFID systems. *Progress In Electromagnetics Research B*, 4:115–126, 2008.

[105] J. G. Kim, W. J. Shin, and J. H. Yoo. Performance analysis of EPC class-1 generation-2 RFID anti-collision protocol. In *Computational Science and Its Applications – ICCSA 2007*, pages 1017–1026. Springer, 2007.

[106] D. K. Klair, K.-W. Chin, and R. Raad. A survey and tutorial of RFID anti-collision protocols. *Communications Surveys & Tutorials*, 12(3):400–421, 2010.

[107] J. Klotz. Statistical inference in Bernoulli trials with dependence. *The Annals of Statistics*, 1(2):373–379, 1973.

[108] M. Kodialam and T. Nandagopal. Fast and reliable estimation schemes in RFID systems. In *Proc. ACM MobiCom*, pages 322–333, 2006.

[109] R. Krigslund, P. Popovski, and G. F. Pedersen. Orientation sensing using multiple passive RFID tags. *Antennas and Wireless Propagation Letters*, 11:176–179, 2012.

[110] R. Krigslund, P. Popovski, G. F. Pedersen, and K. Bank. Potential of RFID systems to detect object orientation. In *Proc. IEEE ICC*, pages 1–5, 2011.

[111] W. Ku, H. Chen, H. Wang, and M. Sun. A Bayesian inference-based framework for RFID data cleansing. *Trans. Knowledge and Data Engineering*, 1(99):1, 2012.

[112] L. Kupper and J. K. Haseman. The use of a correlated Binomial model for the analysis of certain toxicological experiments. *Biometrics*, 34:69–76, 1978.

[113] T. F. La Porta, G. Maselli, and C. Petrioli. Anticollision protocols for single-reader RFID systems: Temporal analysis and optimization. *Trans. Mobile Computing*, 10(2):267–279, 2011.

[114] D. W. Ladd. An algorithm for the Binomial distribution with dependent trials. *Journal of the American Statistical Association*, 70(350):333–340, 1975.

[115] Y.-C. Lai and C.-C. Lin. A blocking RFID anti-collision protocol for quick tag identification. In *Proc. IEEE WOCN*, pages 1–6, 2009.

[116] A. Lázaro, D. Girbau, and D. Salinas. Radio link budgets for UHF RFID on multipath environments. *Trans. Antennas and Propagation*, 57(4):1241–1251, 2009.

[117] K.-C. Lee, A. Oka, E. Pollakis, and L. Lampe. A comparison between Unscented Kalman Filtering and Particle Filtering for RSSI-based tracking. In *Proc. WPNC*, pages 157–163, 2010.

[118] M. Lehtonen, F. Michahelles, and E. Fleisch. Probabilistic approach for location-based authentication. In *Proc. IWSSI*, 2007.

[119] M. Lehtonen, F. Michahelles, and E. Fleisch. How to detect cloned tags in a reliable way from incomplete RFID traces. In *Proc. IEEE RFID*, pages 257–264. IEEE, 2009.

[120] M. Lehtonen, A. Ruhanen, F. Michahelles, and E. Fleisch. Serialized TID numbers - a headache or a blessing for RFID crackers? In *Proc. IEEE RFID*, pages 233–240, 2009.

[121] K. S. Leong, M. L. Ng, and P. H. Cole. The reader collision problem in RFID systems. In *Proc. IEEE MAPE*, volume 1, pages 658–661, 2005.

[122] K. S. Leong, M. L. Ng, and P. H. Cole. Positioning analysis of multiple antennas in a dense RFID reader environment. In *Proc. IEEE SAINTW*, 2006.

[123] G. Li, D. Arnitz, R. Ebelt, U. Mühlmann, K. Witrisal, and M. Vossiek. Bandwidth dependence of CW ranging to UHF RFID tags in severe multipath environments. In *Proc. IEEE RFID*, pages 19–25, 2011.

[124] H.-J. Li, H.-H. Lin, and H.-H. Wu. Effect of antenna mutual coupling on the UHF passive RFID tag detection. In *Proc. IEEE AP-S*, pages 1–4, 2008.

[125] Z. Li, B. Ding, J. Han, and R. Kays. Swarm: Mining relaxed temporal moving object clusters. In *Proc. VLDB*, volume 3, pages 723–734, 2010.

[126] Z.-M. Liu and R. R. Hillegass. A 3 patch near field antenna for conveyor bottom read in RFID sortation application. In *Proc. IEEE AP-S*, pages 1043–1046, 2006.

[127] J. Llinas, D. L. Hall, and M. E. Liggins. *Handbook of Multisensor Data Fusion: Theory and Practice*. CRC Press, 2009.

[128] C.-H. Loo, K. Elmahgoub, F. Yang, A. Z. Elsherbeni, D. Kajfez, A. A. Kishk, T. Elsherbeni, L. Ukkonen, L. Sydanheimo, M. Kivikoski, et al. Chip impedance matching for UHF RFID tag antenna design. *Progress In Electromagnetics Research*, 81:359–370, 2008.

[129] F. Lu, X. Chen, and T. Ye. Performance analysis of stacked RFID tags. In *Proc. IEEE RFID*, pages 330–337, 2009.

[130] G. Marrocco. RFID grids: Part I–electromagnetic theory. *Trans. Antennas and Propagation*, 59(3):1019–1026, 2011.

[131] G. Marrocco, E. Di Giampaolo, and R. Aliberti. Estimation of UHF RFID reading regions in real environments. *Antennas and Propagation Magazine*, 51(6):44–57, 2009.

[132] L. V. Massawe, H. Vermaak, and J. D. Kinyua. An adaptive data cleaning scheme for reducing false negative reads in RFID data streams. In *Proc. IEEE RFID*, pages 157–164, 2012.

[133] L. W. Mayer and A. L. Scholtz. Sensitivity and impedance measurements on UHF RFID transponder chips. In *Proc. EURASIP RFID*, 2008.

[134] C. R. Medeiros, J. R. Costa, and C. A. Fernandes. UHF RFID smart conveyor belt with confined detection range. In *Proc. IEEE APS/URSI*, pages 1–4, 2009.

[135] C. R. Medeiros, J. R. Costa, and C. A. Fernandes. RFID reader antennas for tag detection in self-confined volumes at UHF. *Antennas and Propagation Magazine*, 53(2):39–50, 2011.

[136] R. Miesen, R. Ebelt, F. Kirsch, T. Schafer, G. Li, H. Wang, and M. Vossiek. Where is the tag? *Microwave Magazine*, 12(7):S49–S63, 2011.

[137] R. Miesen, F. Kirsch, and M. Vossiek. Holographic localization of passive UHF RFID transponders. In *Proc. IEEE RFID*, pages 32–37, 2011.

[138] L. Mirowski, J. Hartnett, and R. Williams. Tyrell: A RFID simulation platform. In *Proc. IEEE ISSNIP*, pages 325–330, 2009.

[139] U. Mühlmann, G. Manzi, G. Wiednig, and M. Buchmann. Modeling and performance characterization of UHF RFID portal applications. *Trans. Microwave Theory and Techniques*, 57(7):1700–1706, 2009.

[140] J. C. Nascimento, M. Figueiredo, and J. S. Marques. Trajectory classification using switched dynamical Hidden Markov Models. *Trans. Image Processing*, 19(5):1338–1348, 2010.

[141] R. B. Nelsen. Consequences of the memoryless property for random variables. *The American Mathematical Monthly*, 94(10):981–984, 1987.

[142] C. T. Nguyen, K. Hayashi, M. Kaneko, P. Popovski, and H. Sakai. Maximum likelihood approach for RFID tag set cardinality estimation with detection errors. *Wireless Personal Communications*, 67:1–17, 2012.

[143] T. Nick, S. Cordes, J. Götze, and W. John. Camera-assisted localization of passive RFID labels. In *Proc. IEEE IPIN*, pages 1–8, 2012.

[144] T. Nick, J. Götze, W. John, and G. Stönner. Localization of passive UHF RFID labels using an Unscented Kalman Filter with relative position information. In *Proc. RFID SysTech*, pages 1–7, 2011.

[145] T. Nick, J. Götze, W. John, and G. Stönner. Localization of UHF RFID labels with reference tags and Unscented Kalman Filter. In *Proc. IEEE RFID-TA*, pages 168–173, 2011.

[146] Y. Nie, Z. Li, S. Peng, and Q. Chen. Probabilistic modeling of streaming RFID data by using correlated variable-duration HMMs. In *Proc. ACIS SERA*, pages 72–77, 2009.

[147] P. Nikitin, S. Ramamurthy, R. Martinez, and K. V. S. Rao. Passive tag-to-tag communication. In *Proc. IEEE RFID*, pages 177–184, 2012.

[148] P. Nikitin and K. S. V. Rao. Effect of Gen2 protocol parameters on RFID tag performance. In *Proc. IEEE RFID*, pages 117–122. IEEE, 2009.

[149] P. Nikitin and K. V. S. Rao. LabVIEW-based UHF RFID tag test and measurement system. *Trans. Industrial Electronics*, 56(7):2374–2381, 2009.

[150] P. V. Nikitin, R. Martinez, S. Ramamurthy, H. Leland, G. Spiess, and K. Rao. Phase based spatial identification of UHF RFID tags. In *Proc. IEEE RFID*, pages 102–109, 2010.

[151] P. V. Nikitin, K. Rao, and R. Martinez. Differential RCS of RFID tag. *Electronics Letters*, 43(8):431–432, 2007.

[152] P. V. Nikitin and K. V. S. Rao. Performance limitations of passive UHF RFID systems. In *Proc. AP-S*, volume 1011, 2006.

[153] P. V. Nikitin and K. V. S. Rao. Theory and measurement of backscattering from RFID tags. *Antennas and Propagation Magazine*, 48(6):212–218, 2006.

[154] P. V. Nikitin and K. V. S. Rao. Antennas and propagation in UHF RFID systems. In *Proc. IEEE RFID*, pages 277–288, 2008.

[155] P. V. Nikitin, K. V. S. Rao, and S. Lazar. An overview of near field UHF RFID. In *Proc. IEEE RFID*, pages 167–174, 2007.

[156] P. V. Nikitin, K. V. S. Rao, R. Martinez, and S. F. Lam. Sensitivity and impedance measurements of UHF RFID chips. *Trans. Microwave Theory and Techniques*, 57(5):1297–1302, 2009.

[157] M. H. Ofner, B. Otto, and H. Österle. Integrating a data quality perspective into business process management. *Business Process Management Journal*, 18(6):1036–1067, 2012.

[158] D. Osherson and M. Y. Vardi. Aggregating disparate estimates of chance. *Games and Economic Behavior*, 56(1):148–173, 2006.

[159] A. Parr, R. Miesen, F. Kirsch, and M. Vossiek. A novel method for UHF RFID tag tracking based on acceleration data. In *Proc. IEEE RFID*, pages 110–115, 2012.

[160] E. Perret, S. Tedjini, and R. Nair. Design of antennas for UHF RFID tags. *Proc. of the IEEE*, 100(7):2330–2340, 2012.

[161] P. Popovski, K. Fyhn, R. M. Jacobsen, and T. Larsen. Robust statistical methods for detection of missing RFID tags. *Wireless Communications*, 18(4):74–80, 2011.

[162] A. Prati, I. Mikic, M. Trivedi, and R. Cucchiara. Detecting Moving Shadows: Algorithms and Evaluation. *Tran. on Pattern Analysis and Machine Intelligence*, 25(7):918–923, 2003.

[163] L. R. Rabiner. A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proc. of the IEEE*, 77(2):257–286, 1989.

[164] R. Ranjan and T. Gneiting. Combining probability forecasts. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(1):71–91, 2010.

[165] K. V. S. Rao, P. V. Nikitin, and S. F. Lam. Impedance matching concepts in RFID transponder design. In *Proc. Workshop on Automatic Identification Advanced Technologies*, pages 39–42, 2005.

[166] C. Rasmussen and G. D. Hager. Probabilistic data association methods for tracking complex visual objects. *Tran. Pattern Analysis and Machine Intelligence*, 23(6):560–576, 2001.

[167] C. Ré, J. Letchner, M. Balazinksa, and D. Suciu. Event queries on correlated probabilistic streams. In *Proc. ACM SIGMOD*, pages 715–728, 2008.

[168] A. Ren, C. Wu, T. Wang, and B. Yao. A novel design for UHF near-field RFID reader antenna based on traveling wave. In *Proc. IEEE ICCT*, pages 239–242, 2010.

[169] S. M. Ross. *Introduction to Probability Models*. Academic Press, Inc. Orlando, FL, USA, 2006.

[170] S. Roy, V. Jandhyala, J. R. Smith, D. J. Wetherall, B. P. Otis, R. Chakraborty, M. Buettner, D. J. Yeager, Y.-C. Ko, and A. P. Sample. RFID: From supply chains to sensor nets. *Proc. of the IEEE*, 98(9):1583–1592, 2010.

[171] A. Rozinat, M. Veloso, and W. van der Aalst. Using Hidden Markov Models to evaluate the quality of discovered process models. *BPM Center Report*, 2008.

[172] J. H. Rudander, P.-S. Kildal, C. Orlenius, et al. Measurements of RFID tag sensitivity in reverberation chamber. *Antennas and Wireless Propagation Letters*, 10:1345–1348, 2011.

[173] S. Shao and R. J. Burkholder. Passive UHF RFID tag localization using reader antenna spatial diversity. In *Proc. IEEE ICWITS*, 2012.

[174] S. Shekhar and Y. Huang. Discovering spatial co-location patterns: A summary of results. In *Advances in Spatial and Temporal Databases*, pages 236–256. Springer, 2001.

[175] P. J. Sherman. Conditional probability and correlation coefficient estimators for a first order Bernoulli process. In *Proc. IEEE SSP*, pages 725–728, 2011.

[176] W. Shi and V. W. Wong. MDS-based localization algorithm for RFID systems. In *Proc. IEEE ICC*, pages 1–6, 2011.

[177] W. J. Shin and J. G. Kim. Partitioning of tags for near-optimum RFID anti-collision performance. In *Proc. IEEE WCNC*, pages 1673–1678, 2007.

[178] L. Simon, P. Saengudomlert, and U. Ketprom. Speed adjustment algorithm for an RFID reader and conveyor belt system performing dynamic framed slotted aloha. In *Proc. IEEE RFID*, pages 199–206, 2008.

[179] A. Surana, S. Kumara, M. Greaves, and U. N. Raghavan. Supply-chain networks: A complex adaptive systems perspective. *International Journal of Production Research*, 43(20):4235–4265, 2005.

[180] Y. Tanaka, Y. Umeda, O. Takyu, M. Nakayama, and K. Kodama. Change of read range for UHF passive RFID tags in close proximity. In *Proc. IEEE RFID*, pages 338–345, 2009.

[181] T. Tran, C. Sutton, R. Cocci, Y. Nie, Y. Diao, and P. Shenoy. Probabilistic inference over RFID streams in mobile environments. In *Proc. IEEE ICDE*, pages 1096–1107, 2009.

[182] Y.-J. Tu and S. Piramuthu. Reducing false reads in RFID-embedded supply chains. *Journal of Theoretical and Applied Electronic Commerce Research*, 3(2), 2008.

[183] N. Vaidya and S. R. Das. RFID-based networks: Exploiting diversity and redundancy. *ACM SIGMOBILE Mobile Computing and Communications Review*, 12(1):2–14, 2008.

[184] H. Vogt. Efficient object identification with passive RFID tags. In *Pervasive Computing*, pages 98–113. Springer, 2002.

[185] P. Vorst and A. Zell. Semi-autonomous learning of an RFID sensor model for mobile robot self-localization. In *Proc. European Robotics Symposium*, pages 273–282, 2008.

[186] P. Vorst and A. Zell. Particle Filter-based trajectory estimation with passive UHF RFID fingerprints in unknown environments. In *Proc. IEEE IROS*, pages 395–401, 2009.

[187] J. Waldrop, D. W. Engels, and S. E. Sarma. Colorwave: An anticollision algorithm for the reader collision problem. In *Proc. IEEE ICC*, volume 2, pages 1206–1210, 2003.

[188] Y. Wang. On the limit of the Markov Binomial distribution. *Journal of Applied Probability*, 18:937–942, 1981.

[189] S.-Y. Wei, T.-J. Huang, and H.-T. Hsu. On the performance degradation of RFID system due to curving in tag antenna: Assessment and solutions. In *Proc. IEEE ICWITS*, 2012.

[190] L. Weiss Ferreira Chaves, E. Buchmann, and K. Böhm. TagMark: Reliable estimations of RFID tags for business processes. In *Proc. ACM SIGKDD*, pages 999–1007, 2008.

[191] H. Woellik. A simulation tool for RFID anti-collision algorithms based on ALOHA. In *Proc. IEEE IWSSIP*, pages 1–4, 2009.

[192] V. K. Wu and N. H. Vaidya. Exploiting space-time correlations in an RFID tag field for localization and tracking. In *Proc. IEEE GLOBECOM*, pages 1–5, 2010.

[193] L. Yang, J. Cao, W. Zhu, and S. Tang. A hybrid method for achieving high accuracy and efficiency in object tracking using passive RFID. In *Proc. IEEE PerCom*, pages 109–115, 2012.

[194] H. Yihua, L. Zongyuan, and L. Guojun. An improved Bayesian-based RFID indoor location algorithm. In *Proc. Computer Science and Software Engineering*, volume 3, pages 511–514, 2008.

[195] H. Yojima, Y. Tanaka, Y. Umeda, O. Takyu, M. Nakayama, and K. Kodama. Dynamic impedance measurement of UHF passive RFID tags for sensitivity estimation. In *Proc. IEEE ISCIT*, pages 344–349, 2010.

[196] A. Zaigraev and S. Kaniovski. A note on the probability of at least k successes in n correlated binary trials. *Operations Research Letters*, 41:116–120, 2012.

[197] T. Zhang, Y. Yin, D. Yue, Q. Ma, and G. Yu. A simulation platform for RFID application deployment supporting multiple scenarios. In *Proc. IEEE CIS*, pages 563–567, 2012.

[198] Z. Zhao and W. Ng. A model-based approach for RFID data stream cleansing. In *Proc. ACM CIKM*, pages 862–871, 2012.

[199] L. Zhu and T.-S. Yum. The optimal reading strategy for EPC Gen-2 RFID anti-collision systems. *Trans. Communications*, 58(9):2725–2733, 2010.

[200] Y. Zhu, W. Howard, and K. Q. Pu. Spatial inference using networks of RFID receiver: A Bayesian approach. In *Proc. IEEE GLOBECOM*, pages 1–6, 2009.

[201] Z. Zivkovic. Improved adaptive Gaussian mixture model for background subtraction. In *Proc. IEEE ICPR*, volume 2, pages 28–31, 2004.

[202] Z. Zivkovic and F. van der Heijden. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern recognition letters*, 27(7):773–780, 2006.

GOTOVO![1]

[1] Croatian for: Finished!