Smart Store Applications in Fashion Retail

Matthias Hauser

Dissertation an der Julius-Maximilians-Universität Würzburg
zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)

vorgelegt von
Matthias Hauser

Würzburg, 2019

Erstgutachter: Prof. Dr. Frédéric Thiesse
Zweitgutachter: Prof. Dr. Christoph M. Flath
Datum der Abgabe: 6. August 2019
Datum der mündlichen Prüfung: 14. November 2019

Danksagung

Die vorliegende Arbeit ist nicht das Werk eines isoliert arbeitenden Forschers, sondern Ergebnis zahlreicher Diskussionen und gemeinsamer Arbeit an Projekten. Ich möchte mich an dieser Stelle bei allen bedanken, die zur Fertigstellung dieser Arbeit einen Beitrag geleistet haben. An erster Stelle möchte ich mich bei meinen Betreuern (und Koautoren) Prof. Dr. Frédéric Thiesse und Prof. Dr. Christoph M. Flath bedanken, die mich einerseits bei jedem Schritt meines Dissertationsprojekts bestmöglich unterstützt haben und mir immer mit Rat und Tat zur Seite standen, mir aber gleichzeitig die notwendige Freiheit zur Weiterentwicklung neuer Ideen gegeben haben. Mein Dank gilt weiterhin Prof. Dr. Daniel Müller, der trotz Forschungssemester die Leitung meiner Disputation übernommen hat.

Danken möchte ich darüber hinaus meinen Projektpartnerinnen und Projektpartnern des Forschungsprojekts SERAMIS sowie meinen Kolleginnen und Kollegen am Lehrstuhl für Wirtschaftsinformatik und Systementwicklung an der Universität Würzburg. Mein ganz besonderer Dank gilt (neben meinen beiden Betreuern) Prof. Dr. Cleopatra Bardaki, Prof. Dr. Massimo Bertolini, Dr. Patrick Föll, Matthias Griebel, Sebastian A. Günther, Dr. Jannis Hanke, Dr. Simon Kloker, Roland Leitz, Prof. Dr. Antonio Rizzi, Prof. Dr. Giovanni Romagnoli, Martin Saal, Dr. Nikolai Stein, Giacomo Welsch und Daniel Zügner, die gemeinsam mit mir an den Veröffentlichungen gearbeitet haben, die die Grundlage dieser Arbeit darstellen, oder einen wichtigen Beitrag zu diesen Veröffentlichungen geleistet haben. Darüber hinaus möchte ich mich bei Christiane Kleespies und Justine Grein für die organisatorische Unterstützung und die Verpflegung bei meiner Disputation bedanken.

Während meines Dissertationsprojekts habe ich viel Unterstützung durch meine Familie und meine Freunde erfahren und ich möchte mich hierfür von ganzem Herzen bedanken. Mein ganz besonderer Dank gilt meinen Eltern Christina und Bernhard, ohne deren Unterstützung diese Arbeit nicht möglich gewesen wäre.

Abstract

Traditional fashion retailers are increasingly hard-pressed to keep up with their digital competitors. In this context, the re-invention of brick-and-mortar stores as smart retail environments is being touted as a crucial step towards regaining a competitive edge. This thesis describes a design-oriented research project that deals with automated product tracking on the sales floor and presents three smart fashion store applications that are tied to such localization information: (i) an electronic article surveillance system (EAS) that distinguishes between theft and non-theft events, (ii) an automated checkout system that detects customers' purchases when they are leaving the store and associates them with individual shopping baskets to automatically initiate payment processes, and (iii) a smart fitting room that detects the items customers bring into individual cabins and identifies the items they are currently most interested in to offer additional customer services (e.g., product recommendations or omnichannel services). The implementation of such cyberphysical systems in established retail environments is challenging, as architectural constraints, well-established customer processes, and customer expectations regarding privacy and convenience pose challenges to system design. To overcome these challenges, this thesis leverages Radio Frequency Identification (RFID) technology and machine learning techniques to address the different detection tasks. To optimally configure the systems and draw robust conclusions regarding their economic value contribution, beyond technological performance criteria, this thesis furthermore introduces a service operations model that allows mapping the systems' technical detection characteristics to business relevant metrics such as service quality and profitability. This analytical model reveals that the same system component for the detection of object transitions is well suited for the EAS application but does not have the necessary high detection accuracy to be used as a component of an automated checkout system.

Kurzzusammenfassung

Das fortschreitende Wachstum des Online-Handels setzt traditionelle Modehändler zunehmend unter Druck. Als entscheidender Schritt zur Rückgewinnung von Kunden wird die Transformation traditioneller Ladengeschäfte hin zu intelligenten Ladenumgebungen gesehen. Die vorliegende gestaltungsorientierte Arbeit beschäftigt sich mit der automatischen Verfolgung von Produkten auf der Verkaufsfläche und stellt drei intelligente Anwendungen vor, die auf derartige Informationen angewiesen sind: (i) ein Diebstahlsicherungssystem, (ii) ein System zur Automatisierung des Kassiervorgangs und (iii) eine intelligente Umkleidekabine. Das erste System erkennt Produkte mit denen Kunden die Verkaufsfläche verlassen; das zweite System ordnet diese zusätzlich den richtigen Warenkörben zu. Das dritte System erkennt die Produkte, die ein Kunde in eine Umkleidekabine bringt und identifiziert, basierend auf der Interaktion des Kunden mit den Produkten, an welchem Produkt er aktuell am meisten Interesse hat. Zu diesem sollen anschließend maßgeschneiderte Dienste angeboten werden (z.B. Produktempfehlungen). Die Einbettung derartiger cyberphysischer Systeme in bestehende Einzelhandelsumgebungen ist aufgrund architektonischer Einschränkungen, etablierten Kundenprozessen und Kundenerwartungen hinsichtlich Datenschutz und Einkaufskomfort mit zahlreichen Herausforderungen verbunden. Zur Lösung der einzelnen Erkennungsaufgaben untersucht die Arbeit den Einsatz von RFID-Technologie und maschinellen Lernverfahren. Um die Systeme zudem optimal zu konfigurieren und belastbare Aussagen über den Wertbeitrag dieser zu treffen, wird zudem ein analytisches Modell vorgestellt, welches es ermöglicht die technischen Erkennungsmerkmale der Systeme auf geschäftsrelevante Kennzahlen wie Servicequalität und Rentabilität abzubilden. Die Bewertung der Systeme mit diesem Modell zeigt, dass die gleiche Systemkomponente zur Erkennung von Objektübergängen als Komponente eines Diebstahlsicherungssystems geeignet ist, jedoch nicht die erforderliche Erkennungsgenauigkeit aufweist, um als Komponente eines Systems zu Automatisierung des Kassiervorgangs verwendet werden zu können.

Contents

List of Figures

List of Tables

Abbreviations

ANN	Artificial Neural Network
AUC	Area Under the ROC Curve
BLE	Bluetooth Low Energy
CRISP-DM	Cross-Industry Standard Process for Data Mining
DET	Detection Error Trade-off
DSR	Design Science Research
EAS	Electronic Article Surveillance
EPC	Electronic Product Code
HF	High Frequency
HMM	Hidden Markov Model
IoT	Internet of Things
IS	Information Systems
IT	Information Technology
LF	Low Frequency
LogReg	Logistic Regression
NFC	Near Field Communication
OEM	Original Equipment Manufacturer

PAM	Partitioning Around Medoids
RF	Radio Frequency
RFID	Radio Frequency Identification
ROC	Receiver Operating Characteristic
RSSI	Received Signal Strength Indicator
SD	Standard Deviation
SVM	Support Vector Machine
UHF	Ultra High Frequency
XGBoost	Gradient Tree Boosting

1 Introduction

"When we talk about the Internet of Things, it's not just putting RFID tags on some dumb thing so we smart people know where that dumb thing is. It's about embedding intelligence so things become smarter and do more than they were proposed to do."

— Nicholas Negroponte

1.1 Digital Transformation in Fashion Retail

The proliferation of information technology is fueling service innovation across different domains (Bitner, Zeithaml, and Gremler 2010; Böhmann, Leimeister, and Möslein 2014, 2018; Medina-Borja 2015; Ostrom et al. 2015). Rust and Huang (2014) describe the service revolution and the information revolution as "two sides of the same coin" and argue that neither can be understood without the other. One very promising avenue for research in this context is the digitization of the physical world—usually discussed using the notion of the Internet of Things (IoT) (Perera et al. 2014; Wortmann and Flüchter 2015)—which promises new ways of creating value for service providers (Manyika et al. 2015; Peters et al. 2016). As an important area of application for IoT-based service innovation, the retailing industry forms the subject of this thesis and is one of the "hotbeds of digital services that thrive on advances in information technology" (Böhmann, Leimeister, and Möslein 2018).

Recently, the retailing industry has been undergoing a series of profound structural changes. For traditional brick-and-mortar retailers, the key strategic challenge has been the rapid growth of online competitors (e.g., Amazon, ASOS, Zalando). Competitive pressure arises not only from lower prices, but also from new digital service offerings altering customer relationships, customer behavior, and customers' expectations regarding service quality (Grewal, Roggeveen, and Nordfält 2017; Ingilizian et al. 2017; Kalish and Eng 2018;

PricewaterhouseCoopers 2015). The measures and initiatives which traditional retailers implement to meet these challenges can be subsumed under the umbrella term 'omnichannel retailing' (Brynjolfsson, Hu, and Rahman 2013; Cao and Li 2015; Gallino and Moreno 2014; Kwon and Lennon 2009; Piotrowicz and Cuthbertson 2014; Rigby 2011; Verhoef, Kannan, and Inman 2015). The rationale behind this concept is not to merely copy the strategies of pure online retailers, but rather to systematically integrate online and offline channels to provide a seamless customer experience across existing channels. In this context, Herhausen et al. (2015) distinguish between (i) online-offline and (ii) offline-online channel integration. The first strategy aims at providing access to and knowledge about physical stores online (e.g., providing customers with the option of picking up or returning products ordered online at a nearby store); the second seeks to provide access to and knowledge about the web store in physical stores (e.g., providing customers with the option of ordering products from the web store while in the physical store). In this context, many scholars presume that the role of traditional stores could change and that they could become the platform for the integration of digital and physical channels (Cao 2014; Piotrowicz and Cuthbertson 2014). Channel integration is considered particularly promising for retailers with both physical stores and an online channel given that they have the opportunity to integrate already existing channels (in contrast to pure play retailers) (Herhausen et al. 2015).[1]

A key element of many channel integration strategies is the provision of digital instore-services aimed at the transformation of traditional retail stores into smart stores. Examples of such services include individual pricing, targeted advertisements, automated product recommendations, customer self-services, automated checkout, article security, out-of-stock prevention, and workforce optimization (Betzing, Hoang, and Becker 2018; Manyika et al. 2015). The estimated economic potential of smart stores is huge with projections exceeding \$410 billion annually by 2025 (Manyika et al. 2015). An important indication of their future potential may also be seen in the fact that it is not only traditional retailers who are now concerned with such ideas, but also e-commerce giants like Amazon with its 'Amazon Go' store (Grewal, Roggeveen, and Nordfält 2017). This new store format is based on image recognition techniques, which provide the technological foundation for an automated checkout system. The system promises to automatically detect products taken

[1]Despite the overall optimistic assessment, scholars also presents arguments against channel integration. The main argument is that channel integration may increase research shopping, which describes consumers' propensity to research a product in one channel (e.g., a brick-and-mortar store) and then purchase it through another channel (Herhausen et al. 2015; Verhoef, Neslin, and Vroomen 2007).

from or returned to shelves, keep track of the products chosen by customers, and charge the customers' Amazon accounts when they are leaving the store (Amazon 2018).

Being able to continuously track products and customers represents the cornerstone of many in-store services in retail environments (Manyika et al. 2015). In terms of practical implementation, Radio Frequency Identification (RFID) is the technology of choice for gathering such information for many retail companies (Donaldson 2015). The technology is particularly widespread in fashion retail, with major companies (e.g., Macy's, Marks & Spencer, Zara) already using it (Roberti 2016). The focus of most of the first RFID roll-outs in this area was on the automatic detection of logistical units in upstream and backroom processes (Hardgrave, Aloysius, and Goyal 2013). In contrast, RFID applications on the retail sales floor are still in their infancy (Blázquez 2014). In general, the impact of Information Technology (IT) on operational and management processes can be categorized into (i) automational, (ii) informational, and (iii) transformational effects (Mooney, Gurbaxani, and Kraemer 1996). While the application of RFID in upstream processes of the supply chain mainly yields automational effects (e.g., automation of inventory management processes) and informational effects (e.g., inventory visibility along the retail supply chain), the ability to track garments on the retail sales floor provides various opportunities with transformative potential (Herhausen et al. 2015; Thiesse and Buckel 2015; Verhoef, Kannan, and Inman 2015). The corresponding economic value has already been recognized by retailers with RFID in active use who are eager to leverage their experience with the technology and their existing IT infrastructure to establish smart store environments and ultimately to improve their customers' shopping experience (Donaldson 2015).

1.2 Research Gaps

This thesis examines the use of RFID as a technological enabler for different smart fashion store applications. The starting point of this thesis was a three-year research project on data-driven innovations in retail environments.[2] The consortium included two European fashion retailers, an RFID system integrator, and multiple research institutes. Object

[2]The project was entitled 'Sensor-Enabled Real-World Awareness for Management Information Systems' and received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grand agreement 612 052. Besides the University of Würzburg, the consortium included Adler Modemärkte, Diffusione Tessile, ID-Solutions, Athens University of Economics and Business, the University of Parma, and the Vienna University of Economics and Business.

tracking in retail store environments and the transformative potential of the tracking data was one of the central research topics in the project. This thesis was motivated by the observation that even in supposedly simple applications (e.g., electronic article surveillance), the base technologies available today are not suitable for providing in-store services with the necessary data quality. As will be shown in the following, retailers are confronted with two interrelated challenges in the practical implementation of such applications. The first challenge refers to the dependencies between digital services and events in the real world; the second challenge to the optimal configuration and evaluation of such applications. These two challenges delineate the research gap addressed in this thesis.

The collection of data regarding the physical environment through the use of sensor systems is a key component of smart systems that are based on the locations of physical objects and the detection of interactions with them (Borgia 2014; Manyika et al. 2015). Smart shelves require, for example, information about the garments that are on them (e.g., to display their prices), smart fitting rooms about the garments in their cabins (e.g., to offer helpful product recommendations), and automated checkout systems about the garments customers want to purchase (to initiate correct payment processes). The application of RFID on the retail sales floor is especially error-prone and challenging (Bottani et al. 2012). This is because, in contrast to controlled processes in upstream supply chain processes, the number and variety of simultaneously moving objects is very high. As a consequence, tracking errors may occur in the form of objects passing through a transition area and being accidentally categorized as not having passed (and vice versa). In the case of tagged objects passing through the transition area and not being registered as having done so, I speak of false-negative events. False-positive events, on the other hand, denote situations in which tagged objects that have not passed through the transition area are classified as having done so. Complexity is further increased by the way objects are transported (e.g., stacked, in bags), unpredictable customer behavior, suboptimal store layouts, and lack of space. When using RFID for product tracking in upstream processes, companies usually instruct their employees on how to behave in the proximity of RFID readers (e.g., instructions for holding objects or crossing an RFID gate) (see Figure 1.1 which displays a real example of a signboard in a warehouse). Clearly, such instructions cannot be imposed on customers in a retail store. Tracking systems must nonetheless be able to reliably distinguish between objects moving from one area to another (e.g., objects carried out of a store) and others (e.g., static objects within range of the Radio Frequency (RF) field).

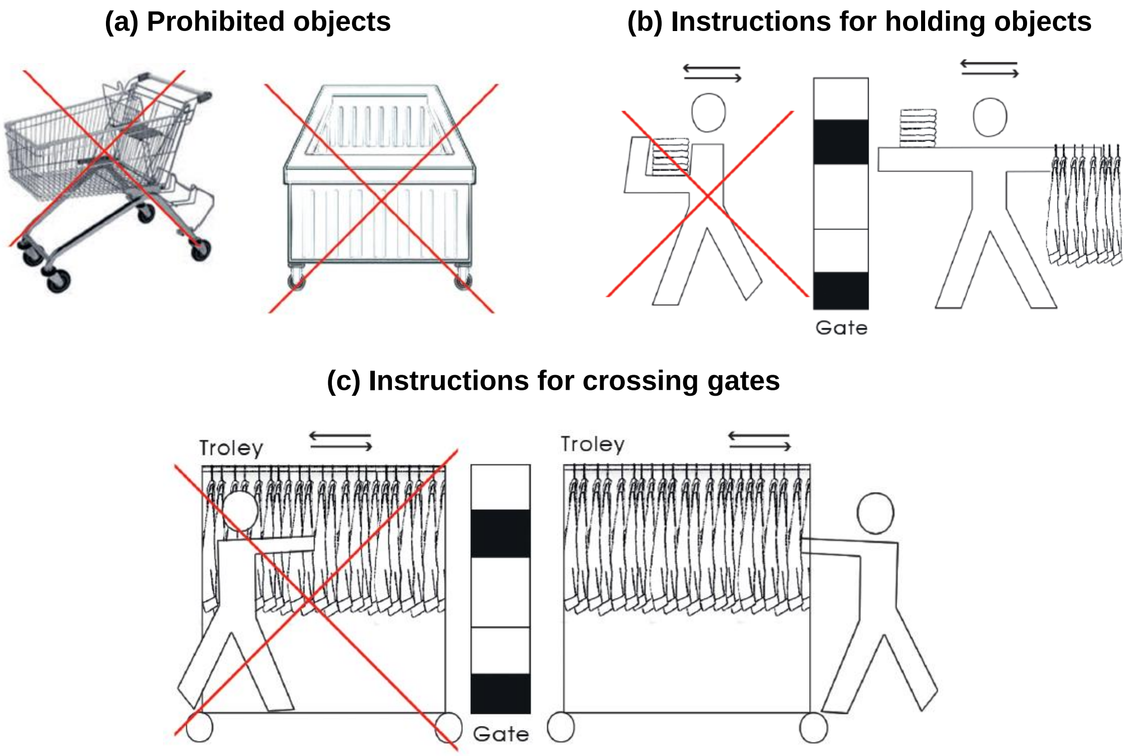

Figure 1.1: Real example of a signboard that prohibits the use of certain logistical equipment and specifies behaviors when passing through the RF field

While in-store detection systems may meet certain technological performance criteria (e.g., the percentage of objects being detected), from an economic perspective it remains unclear how to optimally configure them to minimize the costs incurred (i.e., costs from false-negative and false-positive events). This is particularly important, as misclassifications cannot be ruled out due to the limited process control on the retail sales floor. In addition, technological performance criteria do not allow for conclusions to be drawn concerning the economic value contribution of the detection systems. Typical evaluation procedures for classification systems assume constant cost factors for misclassified entities (e.g., Elkan 2001; Fan et al. 1999; Pazzani et al. 1994). This is a reasonable approach for static settings in which the costs of rework or penalties are fixed (e.g., due to contractual arrangements or internal costing systems). In the case of an in-store detection system, however, this evaluation is particularly difficult because error costs depend not only on individual events (e.g., an item being moved across the sales floor area) but also on the state of the service system in which the detection system is embedded (e.g., time of day, number of customers on the retail sales floor). The complexity of optimizing such a system is therefore substantially higher than optimizing a similar system in the upstream supply chain. The connection between the technological performance of the detection system and its optimal configuration from a store operations perspective renders an evaluation particularly complex.

1.3 Research Objectives

This thesis seeks to design three different smart fashion store applications that are tied to the locations of objects (i.e., garments). The services provided by the applications aim to transform traditional stores into smart stores and help retailers to save costs or effort and increase the attractiveness of their physical stores. The three applications are (i) Electronic Article Surveillance (EAS), (ii) automated checkout, and (iii) smart fitting rooms:

- *EAS systems* are usually located at the exits of retail stores and trigger an alarm if a customer leaves the store with unpaid items.
- *Automated checkout systems* promise to reduce cashier staff requirements and eliminate waiting times at the checkout and must therefore be able to detect customers' purchases and initiate payment processes.
- *Smart fitting rooms* detect the product selections of customers and offer additional services based on these selections (e.g., product recommendations).

In order to guarantee the necessary functionalities of these smart applications, they must be able to reliably detect items that customers carry out of stores (to trigger alarms or initiate payment processes) or into fitting room cabins (to offer additional services based on these items).[3] When developing a model for the reliable detection of item transitions, the two challenges described above must be addressed. To this end, this thesis follows a two-pronged approach (see Figure 1.2). I first seek to improve the accuracy of models for item transition tracking beyond the state of the art ($Classifier_1$ vs. $Classifier_2$). Subsequently, I fine-tune the detection models' configuration to optimally internalize the trade-off between the misclassification events ($Configuration_A$ vs. $Configuration_B$).

[3]While reliable detection of item transitions is of utmost importance for the proposed smart applications, there is more information that can be extracted from sensor data to improve the provided services. Automated checkout systems should not only detect all the products customers want to purchase, but also assign them to individual shopping baskets. Similarly, smart fitting rooms should not only detect the products customers bring into them but also those they are currently interacting with in order to improve service quality (e.g., to highlight recommendations for those items). Besides models for item transition detection, this thesis also develops models that enable the extraction of such additional information.

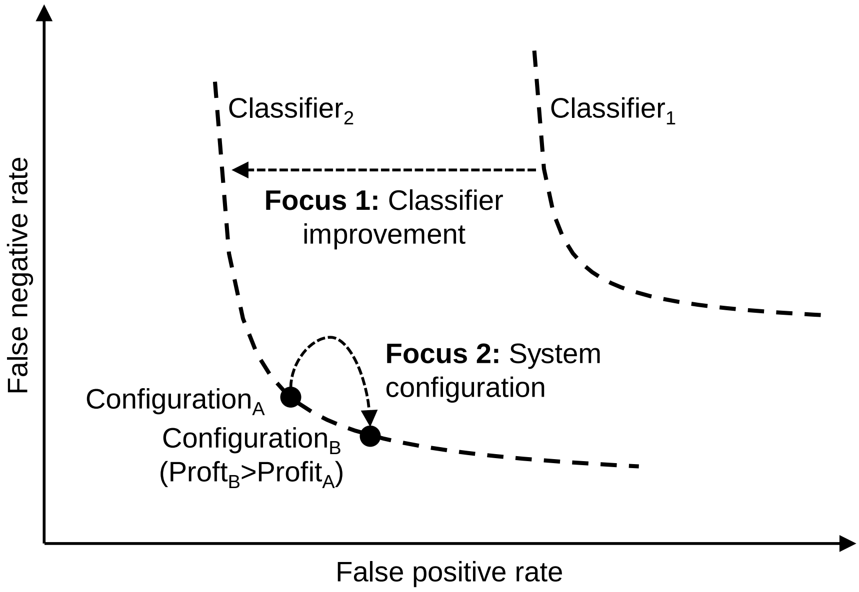

Figure 1.2: Interrelated challenges in the context of RFID-based tracking systems

This thesis first focuses on the question of how to improve the accuracy of classification models for object transition detection in environments with limited process control. To approach this question, I investigate the applicability of machine learning techniques to minimize the occurrence of incorrect classifications in terms of false-negative and false-positive events. The second focus is on the question of how to optimally configure the classification models. False-positive and false-negative classification errors are interdependent, that is, configuring a model for fewer false-positive events typically increases the occurrence of false-negative events and vice versa. To make these performance characteristics more tractable, this thesis proposes a mathematical approximation of a classification model's detection error trade-off curve, which describes the ratio between false-positive and false-negative events as a function of the model configuration. In a second step, this function is integrated into an optimization model on the foundation of prior service operations research to augment the classification model with a retail service operations model. While this thesis presents transition detection models for all three of the above mentioned smart fashion store applications, it focuses on two of the systems, namely EAS and automated checkout, to showcase the applicability of the retail service operations model. The optimization model reflects the costs associated with different types of false classifications (customer dissatisfaction and unpaid merchandise), thus allowing for the identification of an optimal configuration of the two smart fashion store applications.

1.4 Structure of the Thesis

The remainder of the thesis is structured as follows. Chapter 2 continues with a review of literature on (i) cyberphysical systems and service systems, (ii) smart fashion store applications, (iii) RFID-based tracking systems, and (iv) service management and retail operations. Chapter 3 positions this thesis as design-oriented IS research and describes the design science research methodology and evaluation method followed in the course of the research. Chapters 4–6 describe the development of the three above mentioned smart fashion store applications. As outlined above, the focus here is on the development of classification models for the reliable detection of object transitions. Chapter 7 continues with the development of the service operations model. While the assessment of the classification models developed in Chapters 4–6 covers the accuracy of the models in isolation, this chapter evaluates the performance of two of the smart fashion store applications embedded in retail service environments. The thesis closes with a summary, a discussion of the contributions to research and practice, limitations, and an outlook on future research opportunities.

1.5 Previously Published Work

This thesis incorporates research activities conducted over a time span of five years and large parts of the research have already been published in peer-reviewed conference proceedings or journals. This section relates the content of the thesis to these research activities.

An article describing the transition detection classification model presented in Chapter 4 and the service operations model presented in Chapter 7 is currently under review with the *European Journal of Operational Research* (Hauser, Flath, and Thiesse 2019). The article is based on a conference article I presented at the 36th International Conference on Information Systems in Fort Worth, United States (Hauser et al. 2015). The automated checkout artifact presented in Chapter 5 has been published in the journal *Business & Information Systems Engineering* (Hauser et al. 2019). The journal article itself is based on a conference article I presented together with Sebastian A. Günther at the 38th International Conference on Information Systems in Seoul, South Korea (Hauser et al. 2017a). Finally, the smart fitting room artifact presented in Chapter 6 is based on a conference article I presented with Matthias Griebel at the 13th International Conference on Wirtschaftsinformatik in St. Gallen, Switzerland (Hauser et al. 2017b).

2 Background

This thesis draws upon prior research on (i) cyberphysical systems, service systems, and smart service systems; (ii) smart fashion store applications; (ii) RFID-based systems; and (iv) service management and retail operations. The first section focuses on the concept of cyberphysical systems, distinguishes them from service systems and smart service systems, and presents associated design challenges discussed in the literature. The second section presents various smart fashion store applications that are tied to the locations of physical objects. The third section provides background information on RFID technology and summarizes the available design knowledge concerning RFID-based tracking and interaction detection systems. Finally, the last section reviews literature on service management and retail operations and thus forms the basis for the analytical model for the economic evaluation of the smart fashion store applications proposed in Chapter 7.

2.1 Cyberphysical Systems

The term 'cyberphysical system' refers to an intelligent system that connects the physical and digital world using sensors (e.g., RFID, Near Field Communication (NFC), Bluetooth Low Energy (BLE), camera systems, GPS information) and actuators (Borgia 2014). Such systems have progressed beyond speculative visions and early pilot implementations and create previously infeasible processes and establish new business models across various economic sectors (Borgia 2014; Stankovic 2014). In manufacturing, for example, industrial internet applications are increasingly turning shopfloors into smart factories (Lasi et al. 2014; Lee, Bagheri, and Kao 2015; Stein, Meller, and Flath 2018). Smart home applications use information they learn about user behavior to automate energy management and household chores (Manyika et al. 2015). In the automotive sector, ride-hailing platforms (e.g., Uber, Lyft) and recently founded car makers (e.g., Tesla, Waymo) are giving established

Original Equipment Manufacturers (OEMs) a run for their money by replacing individually owned conventional cars with fleets of shared, autonomous vehicles (The Economist 2016). Healthcare innovations (e.g., wearables, augmented surgical tools) promise to improve the well-being and health outcomes of future generations (Lee and Sokolsky 2010). New retail solutions are engendering a fundamental transformation of traditional retail stores into smart stores "that are able to accommodate [customer] needs and wants when desired" (Kourouthanassis and Roussos 2003).

2.1.1 Cyberphysical Systems and Service Systems

Martin, Hirt, and Kühl (2019) find that the term 'cyberphysical system' is often used interchangeably with the terms 'service system' and 'smart service system.' However, while the term 'cyberphysical system' is frequently used in the computer science literature, the latter terms play a dominant role in the Information Systems (IS) community. A closer look at the literature reveals that service systems and smart service systems are usually considered *socio-technical* systems, that is, systems that involve complex interactions between humans, machines, and the environment (Baxter and Sommerville 2011). Cyberphysical systems, on the other hand, are usually characterized as *technical* systems that can be part of a socio-technical system and thus of a smart system or smart service system.

Service systems describe the organizational setting in which services are created, analogous to production systems in manufacturing companies (Maglio et al. 2009). The National Science Foundation (2014) describes them as human-centered, with interactions on the physical or virtual level constituting an essential part of a service that ultimately aims to generate direct or indirect benefits for the parties involved. While the traditional service system construct does not make any statements about the use of technology because of its level of abstraction, a growing number of authors have recently discussed the concept of smart service systems and their incorporation of information technologies (e.g., Beverungen et al. 2017; Frost and Lyons 2017; Medina-Borja 2015). In this context, *smartness* refers to a system's capability for learning, dynamic adaptation, and decision-making, all of which are made possible through the incorporation of technologies for sensing, actuation, coordination, communication, and control (National Science Foundation 2014). Smart service systems can thus be regarded as a special kind of service system. Figure 2.1 depicts the discussed interrelations among the three concepts and their connections to socio-technical systems.

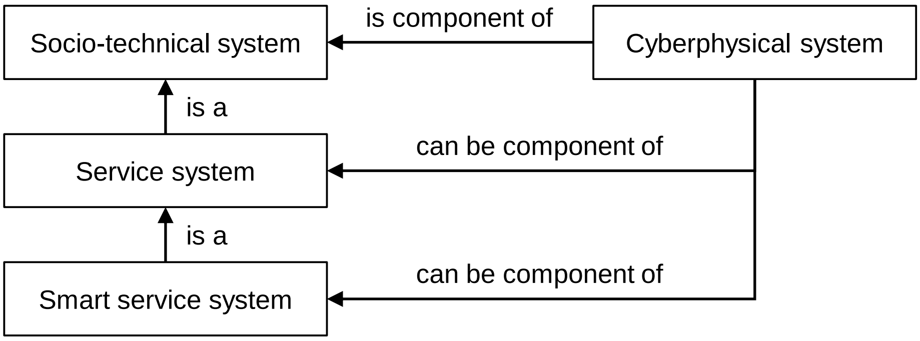

Figure 2.1: Interrelations among service systems, smart service systems, and cyberphysical systems (based on Martin, Hirt, and Kühl (2019))

Beverungen et al. (2017) describe smart service systems as the entirety of service providers, service consumers, and smart products (see Figure 2.2). The latter act as 'boundary objects' at the interface between service providers and consumers and facilitate the transfer of cross-boundary information and knowledge. Smart service systems are organized around one or more products whose smartness may be attributed to various technological features (e.g., unique identification, real-time location tracking, sensor technology).[1] Based on the definition of smart products put forth by Porter and Heppelmann (2014), these features enable four different functions to be performed by or in relation to the smart product, all of which differentiate it from traditional products: (i) monitoring of its environment at the front stage; (ii) remote optimization of the service system using the collected data; (iii) remote control of the smart product; and (iv) the ability of the product to make autonomous decisions. In this context, the term 'front stage' denotes the set of possible interactions between the product and the service consumers, whereas the 'back stage' encompasses all information flows between the product and the service provider.

2.1.2 Design Challenges

Questions regarding the design of (i) cyberphysical systems and (ii) service systems have sparked numerous discussions in the research community (e.g., Böhmann, Leimeister, and Möslein 2014; Ostrom et al. 2010, 2015). The design of cyberphysical systems is considered

[1]Beverungen et al. (2017) argue that smart products need not necessarily fulfill all properties described in the literature (see Beverungen et al. (2017) for a complete list of technological features). Smart garments (e.g., RFID-tagged garments), for example, do not possess actuators or computational capabilities but can still function as smart products.

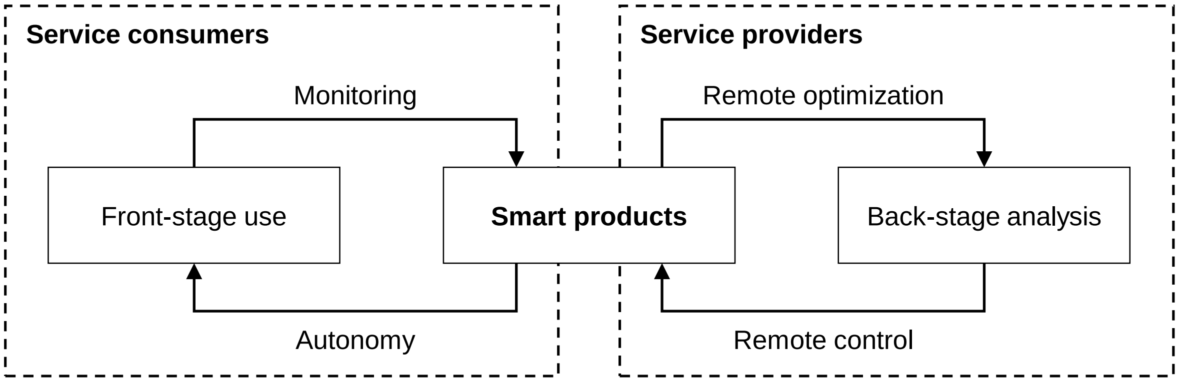

Figure 2.2: Conceptualization of smart service systems (based on Beverungen et al. (2017))

challenging because they have to bridge the boundaries of tangible and intangible resources (Brandt, Feuerriegel, and Neumann 2017; Böhmann, Leimeister, and Möslein 2014) and need to be embedded seamlessly into physical environments (Weiser 1999). The integration of these systems into existing retail environments is considered particularly challenging because these infrastructures are characterized by a high prevalence of immutable components, both physical (e.g., limited store space, architectural constraints) and non-physical (i.e., customers have a clear expectation of how a retail store functions and are unlikely to accept drastic changes) (Kourouthanassis and Roussos 2003).

The design of service systems, on the other hand, is considered challenging because (i) system design must account for the uncertainty that arises from the unpredictability of human behavior (Medina-Borja 2015) and (ii) new technologies have to be leveraged to improve service systems (Ostrom et al. 2015). In this context, Ostrom et al. (2015) call for more research on the issue of how the "IoT and smart services can enhance the customer experience and influence relationships between customers and service providers." Similarly, Medina-Borja (2015) finds that the study of services "has been constrained by existing services enabled by the information technology that is, rather than by the information technology that could be" and that especially advances in the fields of sensing, actuating, and computational and communication technologies could provide ground-breaking contributions to the ongoing development of service systems. An understanding of technology and consideration of the relationship between technology and human behavior must therefore be at the center of the corresponding design activities.

2.2 Smart Fashion Store Applications

Brick-and-mortar retailers are facing increasing competition from their online counterparts (Grewal, Roggeveen, and Nordfält 2017). A recent survey showed that 92 % of retail businesses consider digital innovation as vital or very important with participants referring to it as “something retailers can’t afford not to do” or “one of the most powerful tools [they] have in being able to learn about what [their] customers need” (Morrell 2017). Examples of such digital retail innovations include personal shopping assistants, smart kiosks, automated checkout systems, and smart fitting rooms (Blázquez 2014; Gregory 2015; Herhausen et al. 2015; Manyika et al. 2015; Parada et al. 2015; Senecal and Nantel 2004; Wong et al. 2012). They allow retailers to increase efficiency (e.g., better process control, improved inventory transparency), offer services usually associated with online retailers (e.g., recommendation services, contextualized information) and provide extensive opportunities for the integration of retail channels (e.g., purchasing products that are currently unavailable from the online store while in the smart fitting room). Consequently, such systems can increase the attractiveness of retail stores and at the same time increase their cost efficiency (Gregory 2015; Manyika et al. 2015). While the technological medium for the delivery of digital services in physical stores has so far been primarily the customer’s smartphone (Venkatesh et al. 2017), many of the aforementioned services also require the store to be equipped with hardware and software components that support both the collection of data and new forms of interaction with customers. The implementation of a smart fitting room, for example, requires not only easy-to-use touchscreens but also an infrastructure for the automatic identification of garments in real time.

Some fashion retailers have recently started deploying such cyberphysical systems in their physical stores. The systems can roughly be categorized as (i) applications that offer utilitarian benefits and (ii) applications that save costs or effort (Willems et al. 2017). Applications that fall into the first category are smart kiosks, the aforementioned smart fitting rooms, and smart shelves. J. C. Penney and Louis Vuitton, for example, have installed smart kiosks that allow customers to browse product offerings or order products that are not available in the store (Herhausen et al. 2015; Shankar et al. 2011). Similar systems are also integral components of Amazon’s recently opened ‘Amazon Books’ stores. In these stores, customers can obtain book prices, additional information (e.g., online reviews from the Amazon online store) and access to additional services (e.g., home delivery of products) by scanning books' barcodes with their smartphone or at an in-store terminal (Amazon 2018; Thottam 2016). Rebecca Minkoff, Nordstrom, Ralph Lauren, and Bloomingdale's, on the other hand, have installed smart fitting rooms in their retail stores. Such fitting rooms are not just cabins for trying on selected garments. Instead, they offer customers additional services on a screen within the cabin based on their product selection. One example of a smart fitting room service is product recommendations, which facilitate cross- and up-selling and can lead to substantial sales increases for retailers (Senecal and Nantel 2004; Wong et al. 2012). More importantly, smart fitting rooms enable retailers to provide customers with a seamless shopping experience as they offer various possibilities to bridge the gap between the different retail channels by, for example, offering customers the option of purchasing products that are currently unavailable in the store from the online store while in the smart fitting room. Finally, Parada et al. (2015) introduced an interaction detection system that leverages RFID technology for the detection of RFID-tagged books customers remove from shelves. Such systems offer similar opportunities for the development of novel services as the applications discussed above. Retailers could, for example, use information indicating which books customers remove from shelves (presumably because they are interested in them) to provide them with additional information on these books.

Automated checkout systems are a promising example of an application that falls into the second category of cyberphysical systems in fashion retail (i.e., applications that save costs or effort). Kourouthanassis and Roussos (2003) present an automated checkout system that relies on shopping carts equipped with RFID readers that automatically detect objects placed in the carts. As customers have their own RFID-equipped shopping carts during a shopping trip, the assignment of products to customers is a somewhat trivial task; customers are charged for the products that the RFID reader of their shopping cart has detected. A system which has recently received enormous attention in the media is the so-called 'Amazon Go' store (Grewal, Roggeveen, and Nordfält 2017). The system leverages camera installations and image recognition techniques and promises to automatically detect products taken from or returned to shelves, keep track of the products chosen by customers in virtual shopping carts, and charge the customers' Amazon accounts after they leave the store. In addition, Amazon promises that all customers need to use their system is an Amazon account, a supported smartphone, and the Amazon Go app to register their entrance into the store (Amazon 2018). The corresponding benefits include both, personnel cost reductions for the retailer as well as the complete elimination of waiting times.

2.3 RFID-based Systems

In terms of practical implementation of cyberphysical systems in fashion retail environments, RFID is the technology of choice for many retail companies (Donaldson 2015).[2] Fashion retailers and suppliers first adopted RFID at product case-level mainly for inventory management purposes (Hardgrave, Aloysius, and Goyal 2013). Item-level tagging has, however, moved out of the research environment and into mainstream commerce (Barthel, Hudson-Smith, and de Jode 2014). Today, major fashion retailers such as Kohl's, Macy's, Marks & Spencer, and Zara have already implemented item-level RFID tagging of products. The main reason retailers implemented item-level RFID tagging in the first place was to improve their inventory management (Hardgrave, Aloysius, and Goyal 2013). Item-level tagging enables complete supply chain visibility, due to automatic identification, and the seamless tracking of goods as they move from the suppliers to the customers. This allows one to, for example, detect the causes of shrinkage, monitor the performance of logistical processes, and analyze the movements of individual items in stores. Retailers were thus mainly interested in automational effects (i.e., process automation) and informational effects (i.e., an improvement in at least one data quality aspect). However, fine-granular information stemming from RFID reads also provides opportunities for various data-driven applications to support management decisions and enable novel customer services.

In contrast to barcode scanning, RFID tags not only enable the automatic detection of the number of items belonging to a specific product category but also permit the identification of each specific item (Finkenzeller 2015; Want 2006). Moreover, RFID-based object identification does not require a direct line of sight between the tag and the reader device, allows for the simultaneous bulk detection of multiple objects, and is very robust even under harsh industrial conditions. In addition, certain RFID systems allow for optional data storage within the tag or sensor-based monitoring of various environmental parameters. Camera systems are a possible alternative to RFID technology (Parlak and Marsic 2013).

[2] RFID technology is not only used in retail, but also in many other industries (Zhu, Mukhopadhyay, and Kurata 2012). Examples include the food industry, manufacturing, and healthcare: In the food industry, the technology is used to identify and track animals and to trace the history and location of products to guarantee their quality (Kumar et al. 2009; Ruiz-Garcia and Lunadei 2011). In manufacturing, RFID enables the tracking of materials and components in order to detect disturbances and improve decision-making (Zhong et al. 2017). In healthcare, the technology allows users (e.g., hospital staff) to monitor patients, increase asset utilization through real-time tracking, and improve supply-chain efficiency (Zhu, Mukhopadhyay, and Kurata 2012).

However, they cannot be used in fashion stores because (i) it is difficult (and in some cases even impossible) for cameras to distinguish between garments of different sizes and similar garments from different brands and (ii) such systems raise privacy concerns and their usage is thus problematic in key areas of fashion stores (Litfin and Wolfram 2006).

2.3.1 RFID Technology

RFID systems store data on electronic data carriers commonly referred to as RFID tags. These tags are attached to the objects to be identified. In addition to the tags, an RFID system comprises (i) RFID readers (including antennas) that can both read and write data to the tags and (ii) a data processing system that supports reading and writing functions and processes the sensor data (see Figure 2.3) (Lampe, Flörkemeier, and Haller 2005).

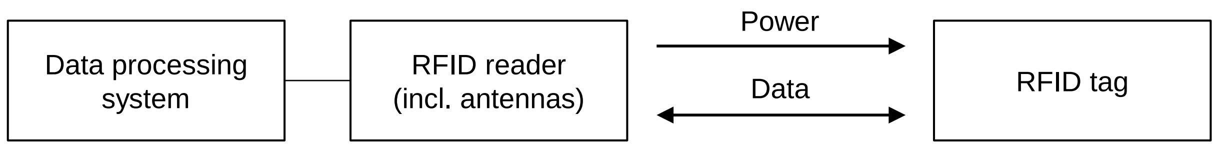

Figure 2.3: Components of an RFID system

RFID tags can be either passive or active (Want 2006). While passive tags use the radio energy transmitted by the reader, active tags have an on-board battery to power their internal circuits (Finkenzeller 2015). Passive tags are generally cheaper than active tags and the component of choice for companies that want to track products along the supply chain (Zhu, Mukhopadhyay, and Kurata 2012). The components of a passive RFID tag are (i) an integrated circuit (also referred to as a chip or a microchip), (ii) antennas that absorb energy propagated by a reader antenna's RF field, and (iii) the encasement (paper or synthetic label or hard case) (see Figure 2.4). The number that uniquely identifies an RFID-tagged object is the Electronic Product Code (EPC), which is encoded into the chip of every RFID tag. An EPC comprises the company code, the product code, and the unique serial number of each object (Finkenzeller 2015).

Passive tags can operate in Low Frequency (LF) (30-300 kHz), High Frequency (HF) (3-30 MHz), and Ultra High Frequency (UHF) (0.3-3 GHz) bands (Finkenzeller 2015). The LF band offers a short read range and a slower read speed than the higher frequencies. The higher the frequency, however, the more similar the behaviour of electromagnetic waves becomes to visible light (Kern 2007). This means that (i) reflections increase (which may

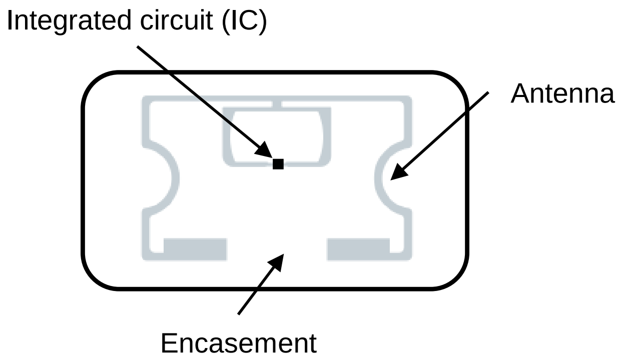

Figure 2.4: Components of an RFID tag

significantly extend the read range of RFID systems) and (ii) losses occur when certain media (e.g., water) are penetrated. Attenuation effects can already be observed in the HF band range and even more so in the UHF band range.

In the fashion retail industry, passive UHF tags are used with frequencies varying from region to region due to different regulations. The main frequencies are 865-868 MHz (Europe) and 902-928 MHz (USA) (Finkenzeller 2015). Passive UHF tags are easier to manufacture than passive LF and HF tags (and therefore cheaper) and typically offer ranges of up to three meters (Finkenzeller 2015). However, the read range achieved in practice depends on many aspects such as the size and quality of the RFID antennas and the sensitivity and transmission power of the RFID reader. In fashion stores, (i) spatial conditions such as the presence of walls and other obstacles and (ii) metal foils or metal ink in goods or packages significantly influences the maximum read range between RFID readers and tags (Keller, Thiesse, and Fleisch 2014a; Lampe, Flörkemeier, and Haller 2005).

2.3.2 RFID Data Analytics

To transform traditional stores into smart stores, retailers must be enabled to gather information about sales floor processes in real time. RFID technology is a prime candidate to gather this information given its unique features for the identification of physical goods. The information needed for the transformation of traditional stores can be categorized as (i) information about the movements of products and (ii) information about customers' interactions with products. While information in the first category answers the question of where RFID-tagged products currently are, information of the second category is concerned with the question of what is currently being done with an RFID-tagged product.

RFID-based Tracking Systems

RFID-based tracking systems may be categorized as (i) systems that aim at determining physical item coordinates (i.e., indoor localization) and (ii) systems that aim at detecting item transitions between areas of interest (e.g., front- and backstore). Indoor environments exhibit severe multi-path effects and low probability of line-of-sight between the tagged objects and the RFID antennas (Motamedi, Soltani, and Hammad 2013). Despite these challenges, techniques for the indoor localization of tagged objects have attracted considerable attention in recent years (Papapostolou and Chaouchi 2011). RFID-based indoor localization typically relies on three techniques: (i) triangulation, (ii) proximity estimation, and (iii) scene analysis (Liu et al. 2007). The first technique uses distance measurements between reference points; the second relies on the measurement of the nearness of a set of neighboring points with known positions. The third technique consists of an offline training and an online phase. The objective of the offline phase is to analyze relationships between signal strength measurements and positions within the environment where the system is deployed. Then, during the online phase, locations of tagged objects are estimated based on the previously collected data. While many authors compare new measurements to the closest a priori measurement in a database (offline phase) (e.g., Hoang et al. 2013; Yuanfeng et al. 2016), machine learning techniques may be applied as well. Such models allow for the formulation of the localization problem as (i) a regression or (ii) a classification problem (Brunato and Battiti 2005). In the first case, the physical coordinates of tagged objects are learned during the offline and estimated during the online phase. The formulation of the classification problem, on the other hand, requires dividing the scene into selected areas. During the offline phase, the machine learning model learns the radio signal behavior within these areas. Then, during the subsequent online phase, the raw data streams of tagged objects are matched to the areas with the closest radio signal characteristics.

The prediction of physical coordinates is often not necessary for the development of in-store applications (Goller and Brandner 2011a; Uckelmann and Romagnoli 2016). Instead, many in-store applications depend on the ability of a system to (i) reliably distinguish between RFID-tagged objects within adjacent areas and (ii) detect transitions between these areas in a timely fashion. Such RFID-based systems are, for example, gates in stores that detect transitions between the sales floor and backroom, EAS systems that detect items carried out of a store, and smart fitting rooms that detect items customers bring into individual cabins. Such systems must reliably distinguish between tagged objects that pass through a transition area and others (e.g., static objects near the RFID reader). False-negative events denote situations in which tagged objects passing through the transition area are not registered as having done so; false-positive events are situations in which tagged objects that do not pass through the transition area are classified as having done so. In practice, the decisive factor in distinguishing between objects that pass through a transition area and others is the Received Signal Strength Indicator (RSSI), a measure of the strength of a signal received from a tagged object. Signals with RSSI values above a certain threshold lead to tagged objects being classified as having moved through the transition area. A typical countermeasure to avoid false-positive events is to reduce the RSSI threshold value (Bottani et al. 2012). However, this usually leads to an increase of false-negative events, which makes the determination of the threshold very difficult.

Approaches addressing this problem can be roughly categorized into (i) hardware-based and (ii) software-based solutions. The first group comprises, for example, shielding measures, antenna design improvements, and additional hardware—like multiple RFID tags per object or additional RFID antennas—and is usually associated with high costs (Ma, Wang, and Wang 2018). Approaches of the second variety apply data analytics techniques to distinguish between objects that pass through a transition area and others. Early contributions in this area considered threshold-based algorithms that use the frequency of tag detections in a sliding-window heuristic (Bai, Wang, and Liu 2006; Brusey et al. 2003; Massawe, Kinyua, and Vermaak 2012). Here, the number of times a particular tag is detected by the reader within a fixed time interval determines whether it is classified as a valid tag detection. The underlying assumption is that undesired reads occur only sporadically, whereas tags that correctly pass the RF field are detected several times. Accordingly, a threshold value regarding the number of detections per time unit must be determined for each RFID installation. Extensions of such algorithms were presented by Fishkin et al. (2004) and Ju Tu and Piramuthu (2008), who propose the use of more than one antenna. The authors argue that valid passages through the RF field have a higher probability of being detected by more than one antenna. The total number of detections per tag should thus be complemented by additional information regarding the number of antennas that detect a particular tag. Keller et al. (2010) include RSSI measurements in the data analysis. The authors utilize RFID data gathered in a distribution center equipped with more than 40 RFID portals in the context of pallets being loaded sequentially into a truck and compare the value of thresholds based on several timestamp-, antenna-, and RSSI-based indicators for the distinction of static and moving RFID-tagged objects. The results indicate that RSSI information provides the best means for distinguishing between RFID-tagged objects.

An alternative approach is the application of classification models from data mining research to distinguish between items that are moved through an RFID gate and others (see Table 2.1). To this end, RFID data streams need to be aggregated to so-called features (e.g., the average signal strength measured during a gathering cycle). These features encode information regarding observed real-world events. In a second step, these features are used for the training of classification models (e.g., logistic regression, decision trees, support vector machines, or artificial neural networks). These models facilitate the automatic mapping of sequences of RFID data streams to classification events.

Table 2.1: Overview of prior research studies leveraging data mining techniques to detect transitions of RFID-tagged items

Study	Environment	Objective
Keller et al. (2012)	Distribution center	Classification between static and moving tags
Keller, Thiesse, and Fleisch (2014a)	Distribution center	Classification between static and moving tags
Ma, Wang, and Wang (2018)	Production environment	Classification between static and moving tags
Buffi et al. (2017)	Office building	Classification between different moving tags

Keller et al. (2012) use decision trees and an empirical data set that was again collected in a distribution center and investigate the impact of different RFID portal configurations (i.e., portals with different numbers of antennas and different antenna orientations) on classification performance. In another study, Keller, Thiesse, and Fleisch (2014a) investigate the applicability of different classification models to distinguish between pallets that are loaded onto trucks and those that are not. In a more recent study, Ma, Wang, and Wang (2018) consider a production environment and aim at detecting tagged keychains carried through an RFID gate by factory workers. In contrast to earlier studies, they consider phase values in addition to RSSI measurements in the development of their predictors and show that this information is indeed useful when distinguishing between static and moved RFID-tagged objects. While the studies presented so far describe the distinction between static and moved RFID-tagged objects as the main problem to solve in regard to the respective classification problem, Buffi et al. (2017) investigate the applicability of classification models to the task of distinguishing between different moving RFID-tagged objects. The authors consider a transition area in an office building and aim at distinguishing between RFID-tagged objects that are moved through an RFID gate—considering incoming as well as outgoing tag events—and RFID-tagged objects that are carried in close proximity to the same gate without passing through it. Interestingly, the authors find that these cases cannot be distinguished based on the average signal strength measured during a gathering cycle which represents a feature with high predictive power in the previously mentioned studies. Instead, Buffi et al. (2017) propose partitioning the sequences of the data stream into a certain number of windows, computing the mean values of the signal strengths for each of these windows, and then using these values as input to the classification models.

Additional approaches presented in the literature leverage (i) dynamic time warping and (ii) hidden Markov models for the distinction between items that are moved through an RFID gate and those that are not. Keller, Thiesse, and Fleisch (2014b) use dynamic time warping, a technique used for speech recognition, for the analysis of RFID time series. The empirical dataset used for evaluation purposes was again collected in a distribution center (see Keller, Thiesse, and Fleisch 2014a; Keller et al. 2010, 2012). The model achieves a detection accuracy that surpasses prior results for the specific case of data sets with exactly one valid tag (i.e., all other simultaneous tag detections are misreads). In contrast, Goller and Brandner (2011a,b, 2012) present probabilistic approaches based on hidden Markov models to detect objects moved through RFID gates. The resulting tag detection algorithms show high classification accuracy for an automated transportation process using conveyor belts under laboratory conditions.

Prior research on RFID-based tracking systems usually assesses tracking performance using standard performance metrics for predictive power in terms of accuracy, that is, the number of correct classifications relative to the total size of the dataset. However, the focus on accuracy neglects the economic impact of misclassifications and the inherent trade-off between different misclassification events. Regarding predictive models for transition detection, this gap also implies that a major degree-of-freedom of these models, the freedom of fine-tuning detection sensitivity, is not used. Unlike hardware-based solutions with hard-wired detection sensitivity, data-driven approaches can be dynamically adjusted to favor either false-positive or false-negative events.

RFID-based Interaction Detection Systems

Several scholars have focused on systems that are based on the detection of interactions with RFID-tagged items. While most research focuses on the detection of object motion (e.g., Parada et al. 2015; Parlak and Marsic 2013), some articles propose systems that are able to distinguish between several interaction types (e.g., Li, Ye, and Sample 2015; Yao et al. 2015). In line with Parlak and Marsic (2013), this thesis defines object motion as any human interaction that causes a change in an object's orientation and location, as well as occlusions with hand or body. In their study, Parlak and Marsic (2013) extract features from low-level RFID data to detect objects that are being used during trauma resuscitations in hospitals. Li, Ye, and Sample (2015), on the other hand, demonstrate that RFID data can be used to differentiate between four interaction types: 'translation' (i.e., movements of RFID-tagged objects of more than ten centimeters), 'rotation' (i.e., rotation around one of the RFID-tagged objects' axes), 'swipe touch' (i.e., swiping a finger across the RFID tag), and 'cover touch' (i.e., touching more than half of an RFID tag).

2.4 Retail Service Management

A proper understanding and modeling of the retail system dynamics is key to the successful design and configuration of smart retail solutions. Such models can help prioritize certain design options and provide a natural approach for progressing beyond merely technological evaluation scenarios. The importance of comprehensive economic evaluation of new technological solutions has been underlined in prior research. Ostrom et al. (2015), for example, find that "measuring and optimizing service performance and impact" represents one of the most important research priorities, with one participant in a roundtable discussion describing the tools used today as "simply too blunt." Lee and Özer (2007), on the other hand, note that industry reports on the value of RFID-based systems are often vague in describing how the promised benefits can be achieved. The authors ascribe the resulting "credibility gap" primarily to a general lack of models and techniques for the assessment of quantifiable economic benefits. Furthermore, they argue that models and techniques from operations management research lend themselves to the quantification of RFID benefits and provide a way to show and understand what RFID can actually achieve in the future.

One important theme in the academic literature is the relationship between service productivity and quality (Parasuraman 2010). Exploring factors that influence quality and productivity performance in service capacity management, Armistead and Clark (1994) conclude that capacity and quality management are fundamentally intertwined and note that “operations managers in a service organization will either succeed or fail in the process of balancing quality of service and resource management [. . .] depending on their skill in managing capacity to match demand.” Furthermore, they identify “coping capabilities” as a key asset of flexible and successful service providers. Similarly, Oliva and Sterman (2001) find that temporary imbalances between service capacity and demand can lead to the permanent erosion of service standards and revenues. A case in point is that of congestion effects as exemplified by queuing. Davis and Vollmann (1990) explore this fundamental question in service quality and conceptualize an integrative framework. They highlight the value of queuing models from operations management to quantify waiting time effects. This line of thought is adapted by Ho and Zheng (2004), who leverage a queuing model with endogenous customer choice to formally analyze trade-offs in congestion-sensitive service environments. Their model simultaneously captures the impact of process variability on quality and the impact of congestion on customer choice. They show that service quality management needs to account for both congestion effects and customer sensitivity. Lu et al. (2013) empirically investigate the effect of in-store queues using information on people waiting at a deli counter and point-of-sale data. They find that waiting in a queue has a non-linear effect on purchases and that queue length has a greater impact than expected wait time on purchase decisions. Building on a model motivated by queuing theory and using a large empirical data set from multiple retail stores, Mani, Kesavan, and Swaminathan (2015) show that reducing understaffing may lead to significant increases in sales and profits. In particular, they underline the importance of acting upon store traffic information to guide staffing decisions. Considering service capacity as a given, Kesavan, Deshpande, and Lee (2014) also rely on different queuing models and use point-of-sale and staffing data to show that congestion management in fashion stores is critical to store performance. Specifically, they identify and explain an inverted U-shape relationship between service system traffic and sales. Based on this observation, the authors make recommendations for active management of staff assignments, with a particular focus on conversion-relevant areas such as fitting rooms.

3 Methodology

Questions relating to the design of information systems have always been an important focus in IS research (Baskerville et al. 2018; Peffers, Tuunanen, and Niehaves 2018). Nevertheless, for a long time researchers struggled to publish designed artifacts in leading IS journals (in particular the Senior Scholars' Basket of Eight journals) (Peffers, Tuunanen, and Niehaves 2018).[1] The emergence of Design Science Research (DSR) as a mainstream research paradigm in IS research is often associated with the Hevner et al. (2004) *MIS Quarterly* article in which the authors provide guidelines for understanding, executing, and evaluating design science research (Gregor and Hevner 2013). Since then many design science articles have been published in leading IS journals. However, most of them have not focused on the actual design of artifacts but rather on "conceptual, theoretical, and guidance contributions to help researchers conduct, present, and publish design science endeavours" (Peffers, Tuunanen, and Niehaves 2018). Against this backdrop, leading scholars in the IS community have called for more research on the actual design of novel and useful artifacts (e.g., Baskerville et al. 2018; Peffers, Tuunanen, and Niehaves 2018). The authors emphasize that the design of such artifacts is an important contribution to the design science knowledge base and that articles focusing on artifact design do not necessarily also have to present a fully developed design theory. In this context, Baskerville et al. (2018) refer to the seminal article by Gregor and Hevner (2013), arguing that design science contributions "could be justified in terms of advances in knowledge in either a problem or a solution domain [and that] design theory development may occur over time and multiple projects, with small steps and revisions on an ongoing basis."

[1] Senior Scholars' Basket of Eight journals are *MIS Quarterly*, *Information Systems Research*, the *Journal of Management Information Systems*, the *European Journal of Information Systems*, the *Information Systems Journal*, the *Journal of the Association for Information Systems*, the *Journal of Information Technology*, and the *Journal of Strategic Information Systems*. These journals are endorsed by the Association for Information Systems as high quality journals within the IS discipline (Levy and Ellis 2006).

3.1 Design Science Research Genres

Peffers, Tuunanen, and Niehaves (2018) find that the great number of guidelines, rules, and frameworks put forward by IS scholars also poses challenges for design science researchers because they make it difficult to conduct DSR projects. The authors consider part of the problem that DSR is still an undifferentiated research paradigm that "leaves reviewers and editors without concepts for how to differentiate among submissions." As a result, many submissions are rejected because reviewers follow one of the many published guidelines and use them indiscriminately as the basis for criticism of DSR submissions (Peffers, Tuunanen, and Niehaves 2018). To address this problem, the authors propose five DSR genres that are grounded in one or more founding papers, define their contributions differently (and evaluate them accordingly), have their own expectations for methodology, and their own presentation style. The objective here is that "researchers who work within a genre will, to some extent, avoid facing reviewer criticism based on premises that deny the legitimacy of work from the genre." The five genres put forward by the authors are (i) IS design theory, (ii) design science research methodology, (iii) design-oriented IS research, (iv) explanatory design theory, and (v) action design research.

The present thesis is positioned as design-oriented IS research, a form of DSR which puts special emphasis on utility for practice and is particularly popular within the German-speaking IS community. Important founding papers for this form of DSR are Winter (2008) and Österle et al. (2011). The latter article is a "memorandum on design-oriented information systems research" which was undersigned by 111 full professors from the community who declared that they "fully agree with this memorandum and make efforts to effectively promote the viewpoints and principles stated therein." The memorandum defines four basic principles that design-oriented IS research must comply with: (i) abstraction, (ii) originality, (iii) justification, and (iv) benefit. The first principle states that each artifact must be applicable to a class of problems, the second that each artifact must substantially contribute to the advancement of the body of knowledge, the third that each artifact must be justified in a comprehensive manner and must allow for its validation, and the fourth that each artifact must yield benefit for the respective stakeholder group (e.g., companies, managers, employees, taxpayers, students) (Österle et al. 2011).

3.2 Research Process and Evaluation Method

As regards the procedure for designing the artifacts, this thesis follows a well-established process model. Since the proposed solutions are based on machine learning methods, I rely on the Cross-Industry Standard Process for Data Mining (CRISP-DM) (Chapman et al. 2000), the most popular process model for data mining projects. The model describes six phases of a data mining project, their respective tasks, and the relationships between these tasks. The phases are (i) business understanding, (ii) data understanding, (iii) data preparation, (iv) modeling, (v) evaluation, and (vi) deployment. The focus of the first phase is to understand the project objectives and requirements from a business perspective. Here, the most important goal is to determine the data mining objectives. The tasks of the second phase are to collect data, to familiarize oneself with available data, to identify possible problems with regard to data quality, and to gain first insights into the data. The third phase covers all the activities needed to construct the final data set that will then be fed into the data mining models (e.g., data selection, feature engineering). The objectives of the modeling phase are to select different modeling techniques, generate a procedure to test model quality (e.g., separate the dataset into train and test sets), build the models (in particular to determine optimal hyperparameters), and assess model quality. While the modeling phase deals with performance indicators such as the accuracy of the models, the evaluation phase assesses the degree to which they meet the business objectives. Finally, the last phase takes the evaluation results and determines a strategy for model deployment.

Though the origins of CRISP-DM lie outside the IS community, the similarities between the model and the DSR approach are remarkably close. Figure 3.1 compares the CRISP-DM phases with the DSR methodology process model introduced by Peffers et al. (2007), an oft-cited example of a methodology for carrying out DSR research. The methodology is based on DSR principles—in particular the principles put forward by Hevner et al. (2004)—and comprises six phases (in Peffers et al. (2007) referred to as 'activities') of a design science project, their respective tasks, and the relationships between these tasks. The objectives of the six phases are to (i) identify and motivate the research problem, (ii) define the objectives for a solution, (iii) design and develop the artifact, (iv) demonstrate its use or to solve one or more instances of the problem, (v) measure how well the artifact contributes to the solution of the research problem, and (vi) communicate the problem and the solution to appropriate audiences.

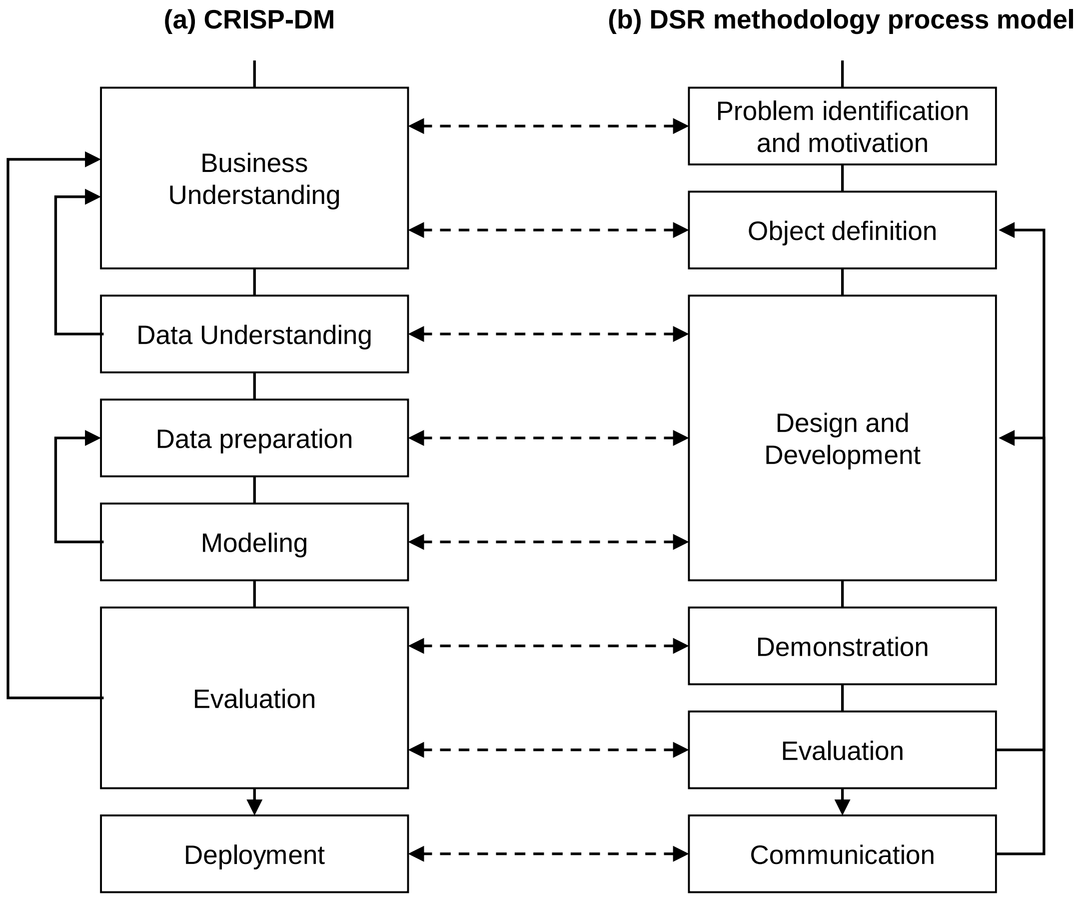

Figure 3.1: Comparison of the project phases in CRISP-DM (Chapman et al. 2000) and the DSR methodology process model put forward by Peffers et al. (2007)

As the graphical depiction of the two process models indicates, some project activities are elaborated in more detail in the one model than in the other. For example, due to its focus on data mining, CRISP-DM puts more emphasis on the design and development phase, which is divided into the three distinct sub-phases of data understanding, data preparation, and modeling. On the other hand, while CRISP-DM only provides for a single evaluation phase, the model of Peffers et al. (2007) distinguishes between demonstration and evaluation. Whereas the former activity demonstrates the use of the artifact in solving a problem as such, the latter evaluates “how well the artifact supports a solution to a problem” which “involves comparing the objectives of a solution to actual observed results from use of the artifact in the demonstration” (Peffers et al. 2007). Furthermore, there are also minor differences regarding the possible process iterations. In particular, CRISP-DM explicitly provides for a jump back from the design and development phase into the business understanding phase.[2] Overall, however, I consider the congruence between the two process models to be so high that CRISP-DM appears acceptable as a more specialized, yet complete, design science research methodology. The only substantial difference may be seen in the fact that, as a research method, the model by Peffers et al. (2007) ends with a communication phase (i.e., scholarly and professional publications), while the CRISP-DM cycle terminates with the deployment of a solution in practice.

Following the framework for evaluation in DSR proposed by Venable, Pries-Heje, and Baskerville (2016), the evaluation method followed in this thesis belongs to the class of artificial ex post evaluations. The ex post evaluation serves a summative purpose (Sein et al. 2011), that is, the thesis evaluates the artifact components after they have been created to support the decision of selecting them for an application. This objective may be achieved either by artificial evaluation (e.g., experiments, simulation, mathematical proofs) or by evaluation in a naturalistic setting. While at first glance a naturalistic evaluation seems clearly more appropriate to assess the effectiveness of an IT artifact, it is also associated with high costs and the risk of misinterpreting results due to confounding variables (Venable, Pries-Heje, and Baskerville 2012). In the specific case at hand, a complete naturalistic evaluation even seems practically impossible, since the involvement of real shoplifters in the EAS scenario, for example, is not feasible for obvious reasons. The thesis therefore relies on an artificial evaluation method using experimental data collected in the lab under real-world conditions in combination with an analytical model, which in turn offers advantages in the form of better repeatability and falsifiability (Gummesson 1998).

[2]Chapman et al. (2000) note that moving back and forth between the six different phases is always possible and that the arrows displayed in the CRISP-DM reference model only indicate the most important and frequent dependencies between the process phases.

4 Design of Electronic Article Surveillance Systems[1]

The first smart fashion store application this thesis is concerned with is an electronic article surveillance system that leverages RFID technology and machine learning techniques to reliably detect RFID-tagged items leaving the shopping floor area so that an alarm can be triggered in case a customer leaves the store with items that have not been paid for.

Inventory shrinkage is a major issue worldwide (Bamfield 2011; Bottani et al. 2012; Fan et al. 2014). In 2010, total inventory shrinkage in retail amounted to more than 100 billion USD globally, which is about 1.36 % of global retail sales (Bamfield 2011). The largest cause of shrinkage is customer theft with 42.4 % of total shrinkage, followed by employee theft (35.3 %), administrative failure (16.9 %) and supplier/vendor fraud (5.4 %) (Bamfield 2011). The negative consequences of theft are not limited to the value of stolen items but rather include store security expenses, higher prices for consumer goods, and lost sales taxes (Deyle 2015). Moreover, like any other form of shrinkage, theft may lead to inventory inaccuracies and thus to stock-outs and inefficient store replenishment processes (Kang and Gershwin 2005). The main approach to combat theft in retail stores is the introduction of EAS. These systems trigger an alarm if a customer leaves the store with unpaid items. Yet while protecting from theft, EAS systems often give rise to false alarms which "create collateral damage such as store disruption and customer irritation" (Hayes and Blackwood 2006). According to Dawson (1993), they may result in lowered goodwill, negative word-of-mouth, the entire loss of a future stream of revenue, and possibly costs from legal actions brought by patrons. The same author summarizes that EAS "provides a good example of the unintended pitfalls of poorly implementing a new technology" as retailers often do not account for the direct and indirect costs of false alarms.

[1] The chapter is adapted from Hauser, Flath, and Thiesse (2019) (see Section 1.5).

4.1 Practical Background

EAS systems are usually located at the exits of stores and trigger an alarm if a customer leaves the store with items that have not been paid for. Traditional EAS systems are so-called 1-bit systems, which means that they can only detect whether objects are in the surveillance zone but cannot uniquely identify them (Finkenzeller 2015). In contrast, RFID tags permit the identification of each specific item in the surveillance zone (Want 2006).[2]

4.1.1 Traditional EAS Systems

Traditional EAS systems are (i) RF systems, (ii) electromagnetic systems, and (iii) acousto-magnetic systems (Bottani et al. 2012; Herzer 2003). The tags of RF systems have a metal antenna with a small diode that allows them to emit a radio signal in response to a radio signal received from an antenna of the system (Bottani et al. 2012). Loop antennas are used to generate the required alternating magnetic field in the detection area of the article surveillance system and the technology allows for gate widths of up to two meters (Finkenzeller 2015). In contrast, the tags of electromagnetic and acousto-magnetic systems consist of soft magnetic strips, which act as sensors, and a semi-hard bias magnet for activation and deactivation (Finkenzeller 2015; Herzer 2003). Electromagnetic systems are commonly used in libraries because the tags can be deactivated when books are borrowed and reactivated upon return (Bottani et al. 2012; Finkenzeller 2015). In addition, the systems are well suited to securing low value goods in retail stores because the tags are comparatively cheap (Bottani et al. 2012; Finkenzeller 2015). Acousto-magnetic systems, on the other hand, are considered more sophisticated than electromagnetic systems (Herzer 2003). In general, these tags are thicker than electromagnetic tags and the systems have a higher sensitivity, which enables, for example, much wider EAS gates (Herzer 2003).

4.1.2 RFID-based EAS Systems

In contrast to 1-bit systems, RFID-based systems are able to send more complex signals that uniquely identify the objects they are attached to. RFID-based EAS systems are particularly interesting for retailers that already use RFID technology for the automatic

[2]Despite the differences between RF systems and RFID systems, the term RFID is often loosely used in the literature to describe both systems (Finkenzeller 2015).

detection of garments in upstream and backroom processes because the same tags can also be used for article surveillance. Moreover, the cost-efficient substitution of traditional EAS solutions by RFID has become an important foundation for the business case underlying many RFID implementations in retail (GS1 Germany 2007).

RFID-based EAS systems must reliably distinguish between tagged objects that pass through an EAS gate and others (e.g., static tagged objects near the gate). If tagged objects passing through the transition area are not registered as doing so, we speak of false-negative events. False-negative events result in losses for retailers and lead to inventory inaccuracies (Kang and Gershwin 2005). False-positive events, on the other hand, denote situations in which tagged objects that do not pass through the transition area are classified as having done so. In the case of EAS, false-positives not only lead to incorrect inventory data but also trigger false alarms, which impair customer satisfaction. False-positives may occur, for example, if customers with unsold items walk in close proximity to but not through the gate or products are displayed in close proximity to the antennas. The presence of products with metal applications or any other metallic object within range of the antennas may also influence the readability of RFID tags. In practice, the decisive factor in distinguishing between objects that pass through a transition area and others is the RSSI, a measure of the strength of a signal received from an RFID-tagged object. Signals with RSSI values above a certain threshold lead to tagged objects being classified as having moved through the transition area. A typical countermeasure to avoid false-positive events is to reduce the RSSI threshold value (Bottani et al. 2012). However, this usually leads to an increase of false-negative events, which makes the determination of the threshold very difficult (see Section 2.3).

To gain a better understanding of the challenges associated with RFID-based EAS systems, we conducted a pre-study in an RFID laboratory equipped with an EAS gate with a gate-mounted RFID reader from Impinj (Impinj Inc. 2017a) and four far-field antennas, a typical setup for RFID-based transition detection systems (Bottani et al. 2012). In this environment, we considered (i) two movement paths and (ii) two test scenarios. The two movement paths are walking straight through the gate and walking straight by the gate with one metre distance to the gate. In the first scenario, tagged garments were held in front of the carrier's body; in the second scenario, the tagged garments were put in a booster bag, that is, a standard shopping bag with a single layer of alumina foil. We chose these two typical scenarios in consultation with the industry partners of the three-year research project introduced in the Introduction (see Section 1.2) and discussions with Eleonora Bottani who had conducted a comprehensive study on the potentials of RFID for EAS considering similar scenarios (Bottani et al. 2012).

We repeated each of the four tests multiple times. Figure 4.1 shows exemplary antenna traces from the test runs. The graphs show that with non-concealed tags, 'by the gate' and 'through the gate' events can reliably be distinguished by means of their RSSI measurements. When the tagged garment is held in front of the carrier's body, the signal strength of the RFID reads increases steeply when the person with the garment approaches the gate. Moreover, the maximum signal strength is considerably higher than it is in the 'by the gate' setting and is reached when the person with the tag is closest to the antennas, that is, when the person is standing right in the middle of the gate. However, during in-store situations with concealed tags (i.e., clothes in a bag), antenna readings become so strongly distorted that accurate separation of the two classes of events using thresholds becomes unfeasible. The results show that the booster bag scenario is particularly challenging and therefore seems well-suited for the present study.

4.2 System Design

Instead of adjusting the RSSI threshold value, we propose keeping antenna power levels at 100% and using data mining models to distinguish between items that are carried through an EAS gate and others. Given the broad scope of our study, we approach the classification problem using a set of standard algorithms: logistic regression, decision trees, Artificial Neural Networks (ANNs), and Support Vector Machines (SVMs). Other methods or ensembles over multiple classifiers should not qualitatively change the results. This approach is in line with prior research on RFID data analytics (Buffi et al. 2017; Keller, Thiesse, and Fleisch 2014a; Ma, Wang, and Wang 2018). To train these models, so-called 'features' must be extracted from the raw data, which contain information regarding observed real-world events. Although the classifiers used in the thesis are generic, the considered features are appropriately specific to RFID and must be developed based on knowledge of the particular business process. Several authors stress the fact that feature generation is a key phase of any data mining project (Domingos 2012; Flath and Stein 2018; Halevy, Norvig, and Pereira 2009; Stein and Flath 2017).

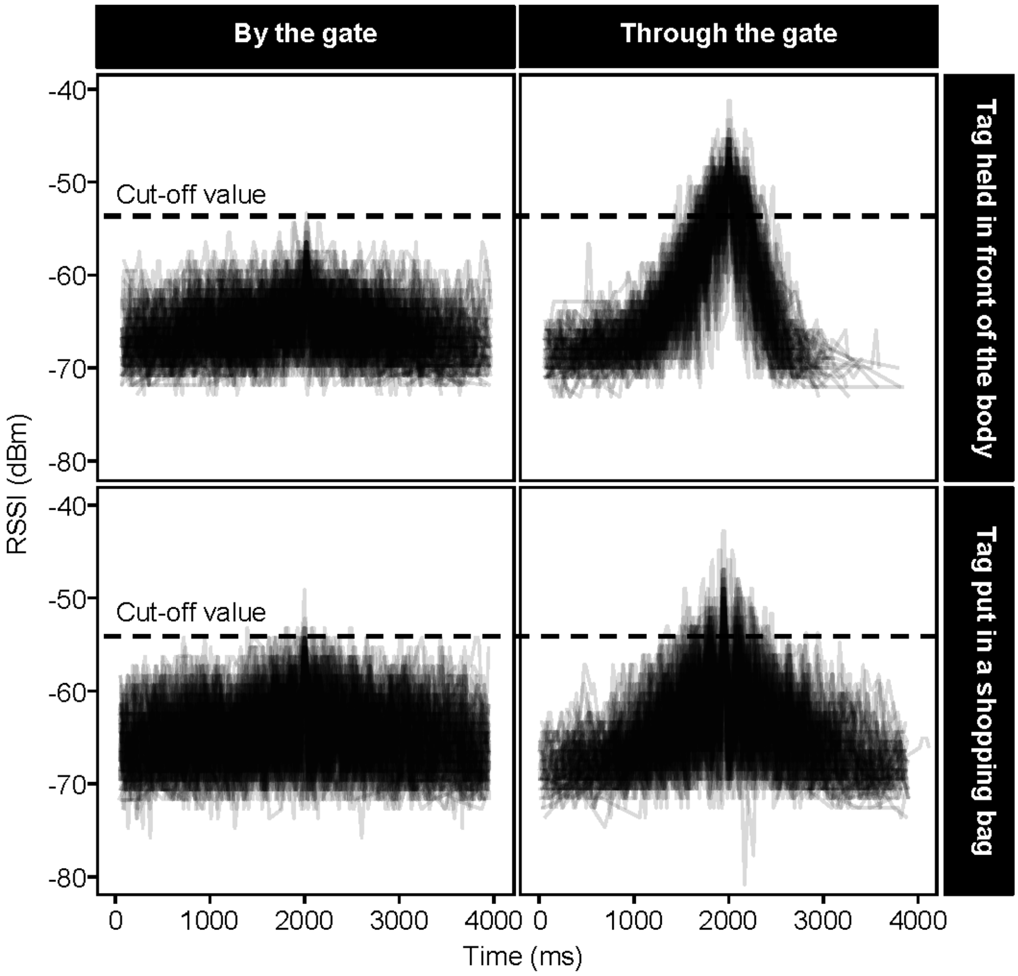

Figure 4.1: Exemplary RFID detection time series

4.2.1 Data Understanding

Table 4.1 provides a representative excerpt from the low-level RFID data gathered with the RFID infrastructure described in the last section (i.e., an EAS gate with a gate-mounted Impinj RFID reader (Impinj Inc. 2017a) and four far-field antennas). Each row reflects a single RFID tag read event triggered by one of the four far-field antennas. EPC is the unique identifier of the RFID tag, Timestamp is the Unix timestamp of when the tag was read, RSSI is the radio signal's power measured in dBm, and Antenna is the unique ID of the far-field antenna that read the RFID tag.

Table 4.1: Representative low-level RFID data excerpt

EPC	Timestamp	RSSI	Antenna
0x3551c704600c35b000018bcb	1416221066.269	-63	4
0x3551c704600c35b000018bcb	1416221066.354	-67	2
0x3032d58e474bf440000000b8	1416221066.396	-71	3
0x3032d58e474bf440000000b8	1416221166.861	-74	1

4.2.2 Feature Engineering

Features are generated by applying various aggregation functions that correspond to specific characteristics of the data (e.g., maximum RSSI value of a tag during a test run). We seek to identify features with high predictive power to achieve robust classification for all scenarios, including the scenario where the tag is located in a booster bag.

Table 4.2 lists typical features considered for the classification of RFID events used in previous studies (Buffi et al. 2017; Keller, Thiesse, and Fleisch 2014a; Ma, Wang, and Wang 2018). These features are very useful for distinguishing RFID tags in controlled environments, but their ability to distinguish moving objects from other moving objects in uncontrolled environments seems limited. For this reason, we came up with additional features that are specifically tailored towards our application (see Table 4.3). These features put information from different attributes into relation to one another (e.g., temporal relation of reads depending on antenna or signal strength information) which should help distinguishing moving from other moving objects.

Table 4.2: Features used in previous studies on RFID data analytics

Feature	Description
F1	Average signal strength of all reads
F2	Difference between the highest and lowest signal strength
F3	Maximum signal strength of all reads
F4	Standard Deviation (SD) of RSSI measurements

To illustrate our reasoning, Figure 4.2 shows the signal strength recorded by the two antennas on the left and right side of the gate over the course of two representative test runs. In both cases, the tagged garment was held in front of the carrier's body. In the left plot, the RFID-tagged garment was moved by the gate and in the right one, the garment

Table 4.3: Additional features used in this study

Feature	Description
F5	Maximum signal strength measured by the antennas on the bottom
F6	Maximum signal strength measured by the antennas on the top
F7	Mean RSSI measurement of the antennas on the left side of the gate
F8	Mean RSSI measurement of the antennas on the right side of the gate
F9	Mean temporal shift between the signals' timestamps of the antennas on the top and the bottom
F10	Mean temporal shift between the signals' timestamps of the antennas on the right and the left side of the gate
F11	Number of tag reads in the time interval 500ms before and after the maximum RSSI measurement
F12	Proportion of the tag reads before the maximum RSSI measurement over all tag reads
F13	Proportion of the tag reads in the time interval 500ms before and after the maximum RSSI measurement over all tag reads
F14	Temporal shift between maximum RSSI values of the antennas on the bottom
F15	Temporal shift between maximum RSSI values of the antennas on the top
F16	SD of the timestamps of the antennas' maximum RSSI measurements
F17	Temporal shift between maximum RSSI value of the antenna with earliest maximum value and the antenna with the latest
F18	Temporal shift between maximum RSSI value of the antennas on the one side of the gate and the antennas on the other side of the gate
F19	Proportion of maximum timestamp minus timestamp of the maximum RSSI measurement over maximum minus minimum timestamp
F20	Regression coefficient of linear regression model with dependent variable signal strength and explanatory variable timestamp
F21	Regression coefficient of linear regression model based on the signals measured after the maximum signal strength measurement with dependent variable signal strength and explanatory variable timestamp
F22	Regression coefficient of linear regression model based on the signals measured before the maximum signal strength measurement with dependent variable signal strength and explanatory variable timestamp
F23	Linear regression coefficient of quadratic regression model with dependent variable signal strength and explanatory variable timestamp
F24	Quadratic regression coefficient of quadratic regression model with dependent variable signal strength and explanatory variable timestamp

was moved through the gate. This basic example highlights how both types of features help separate the two classes: mean and maximum of the signal strength measurements over the interval as well as a spatio-temporal relationship. In the example on the right, the RSSI values measured by all antennas increase steeply until the maximum is met. All antennas exhibit their maximum roughly at the same time—this is the point when the carrier with the RFID-tagged item is in the middle of the gate. After passing the antennas, the RSSI values decrease even more steeply, given that the body shields the tag while moving away from the gate. Conversely, in the left-hand panel the two sides' RSSI maxima are shifted in time. This can be explained by the fact that the tag moves by the gate in a straight line. Hence, the tag is first closer to one side of the gate and then moves towards the other side. The combination of many different features allows us to augment pure signal strength readings with spatial (i.e., the proximity of the tag to the antenna) as well as temporal (i.e., the time between sequential read events) information.

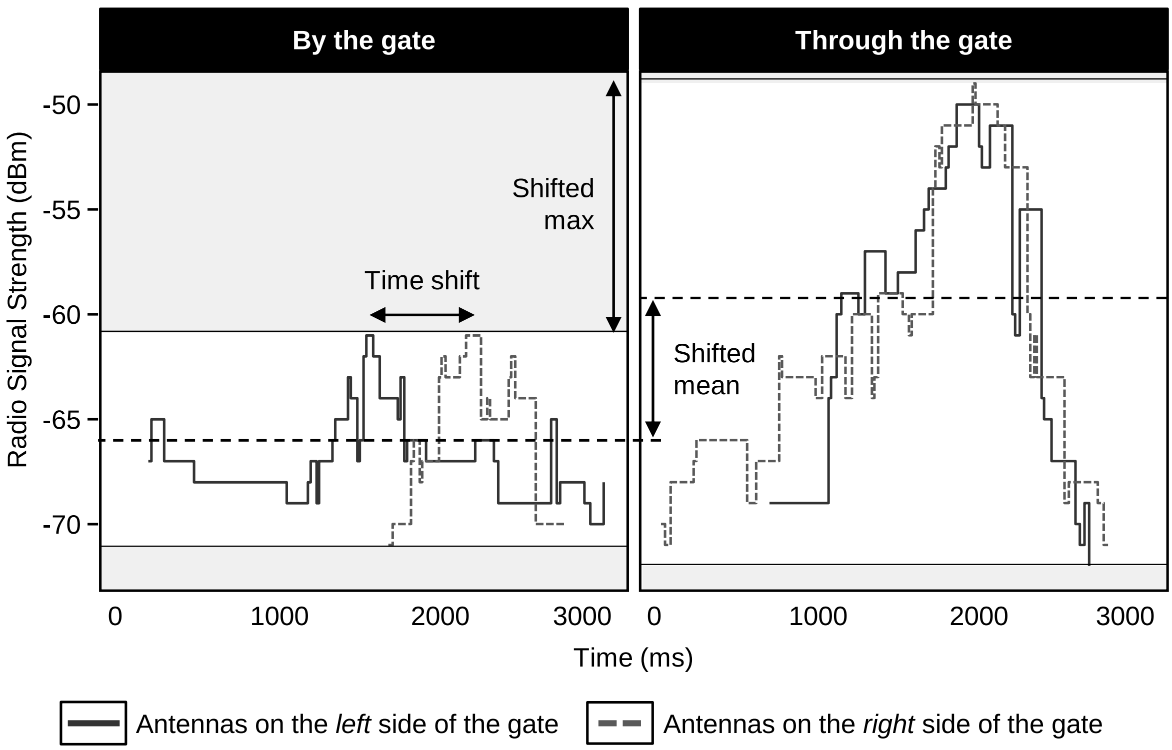

Figure 4.2: Side-by-side comparison of two representative test runs

4.3 Evaluation

We collected large data sets in two laboratories under real-world conditions for instantiation and evaluation of the EAS artifact. In the following, the evaluation setting (Section 4.3.1) and the evaluation results (Section 4.3.2) are described. The evaluation focuses particularly on comparing the results of our classification models with (i) classification models based only on features used in previous studies and (ii) the threshold approach. The threshold approach classifies RFID-tagged objects with signals above an optimal RSSI threshold as having moved through the transition area (see Section 4.1.2).

4.3.1 Evaluation Setting

The detection system prototypes were set up in two university research laboratories, one at the University of Würzburg in Germany and one at the University of Parma in Italy. Figure 4.3 shows the dimensions of the gates used in the laboratories (which were identical in both cases) and a picture of the evaluation setting in Italy.

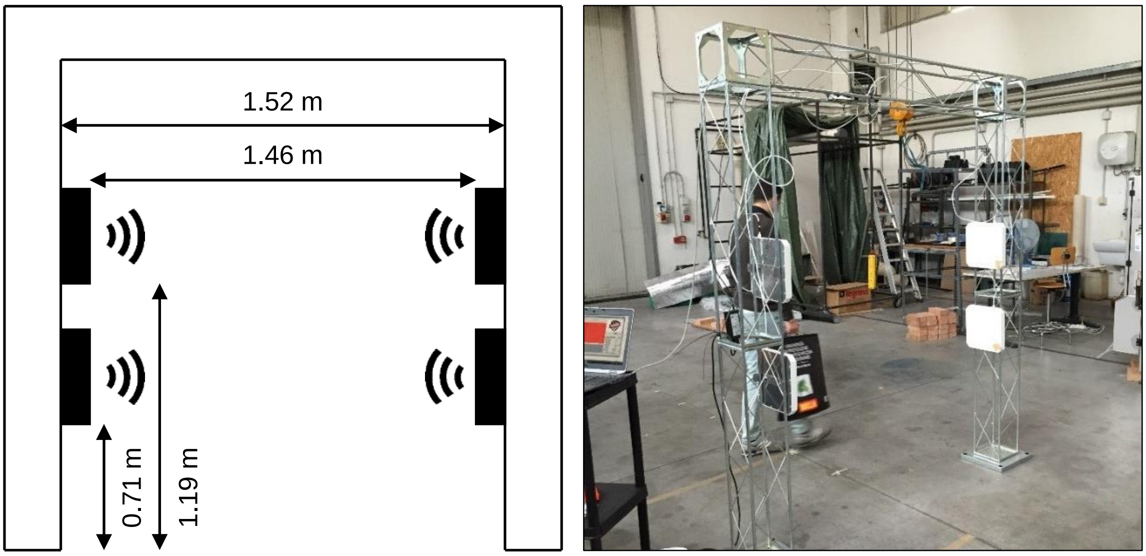

Figure 4.3: Dimensions of the RFID gates and picture of the laboratory

Naturally, the experimental setup had to appropriately reflect the intricacies of a real-word store (Bottani et al. 2012). To achieve this, we adopted two different test scenarios, multiple walking paths, and different movement speeds (running and walking). In the first test scenario, tagged garments were held in front of the carrier's body; in the second scenario, the tagged garments were put in a booster bag. Figure 4.4 illustrates the customer movement paths we considered in our experiments. We included complex settings with combinations of movement paths that require the transition detection system to be able to distinguish between different moving objects.

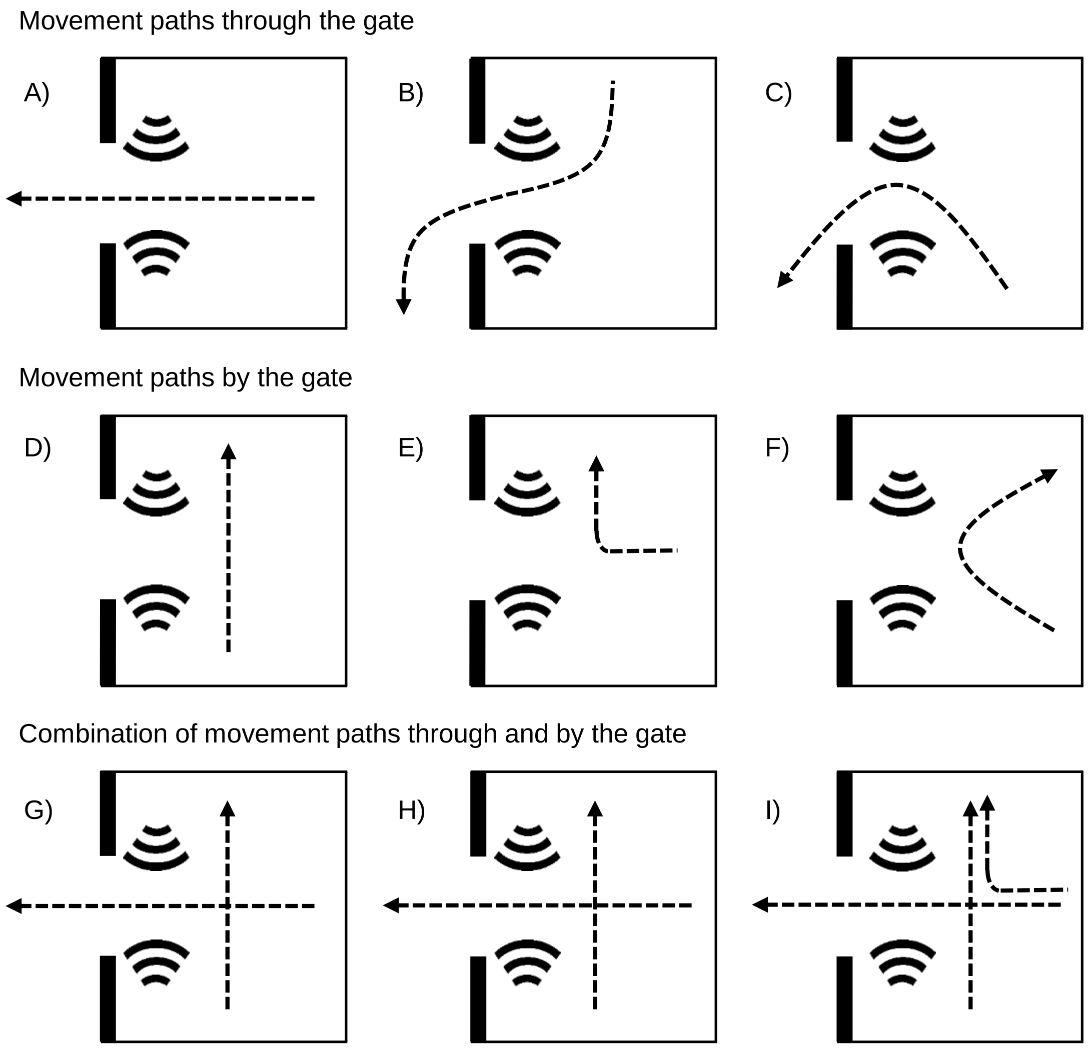

Figure 4.4: Test setting with typical movement paths in retail stores

In total, our experimental design includes 20 test settings, which are combinations of different movement paths, numbers of people, movement speeds, and test scenarios (e.g., two people per path are running straight through the gate with one RFID-tagged garment held in front of the body) (see Table 4.4). The resulting data set comprises 4554 detection time series, of which 2640 represent valid passages through the gate. The total number of individual tag read events across these time series is 488 006.

Table 4.4: Experimental design

Test description	Movement patterns A	B	C	D	E	F	G	H	I
One person per path walking with one garment held in front of the body	150	150	150	150	150	170	120	120	120
One person per path walking with one garment put in a shopping bag	150	150	150	150	150	150	-	-	-
One person per path running with one garment held in front of the body	150	-	-	-	-	-	-	-	-
One people per path running with one garment put in a shopping bag	150	-	-	-	-	-	-	-	-
Two people per path walking with one garment held in front of the body	240	-	-	-	-	-	-	-	-
Two people per path running, with one garment held in front of the body	120	-	-	-	-	-	-	-	-
Three people per path walking with one garment held in front of the body	120	-	-	-	-	-	-	-	-

4.3.2 Evaluation Results

We followed best practices in machine learning to ensure that the results were representative and performed 10-fold cross validation: In each round, we used 90 % of the data for the training of our classification models and the remaining 10 % for the evaluation of the artifact. In addition, we performed hyper-parameter optimization of the models by considering, for example, different maximum numbers of constructed decision trees for the random forest classifier to ensure robust classification results (Witten et al. 2016).

Following Keller, Thiesse, and Fleisch (2014a) and Ma, Wang, and Wang (2018), we assess the performance of the classification models considering the performance measures *Accuracy*, *Precision*, *Recall* and *Area Under the ROC Curve (AUC)*. Accuracy is defined as the number of correct classifications relative to the total size of the data set. Precision is the share of instances classified as 'moved through the gate' that actually were moved through the gate. In our application, if tags that were not moved through the gate are erroneously classified as 'moved through the gate' (false alarms), precision is diminished. Recall, on the other hand, measures the proportion of correctly classified 'through the gate' instances. For very conservative classifiers, which tend to classify uncertain cases as 'not moved through the gate,' recall will be low. The Area Under the ROC Curve (AUC) provides an aggregate measure of performance across all possible discrimination thresholds (Fawcett 2006). Receiver Operating Characteristic (ROC) curves plot the true-positive rate against the false-positive rate at various threshold settings (see Figure 4.5). When evaluating a data mining model with the ROC curve, the farther the curve shifts towards the top-left corner (i.e., the greater the area under the ROC curve) the better the model is.

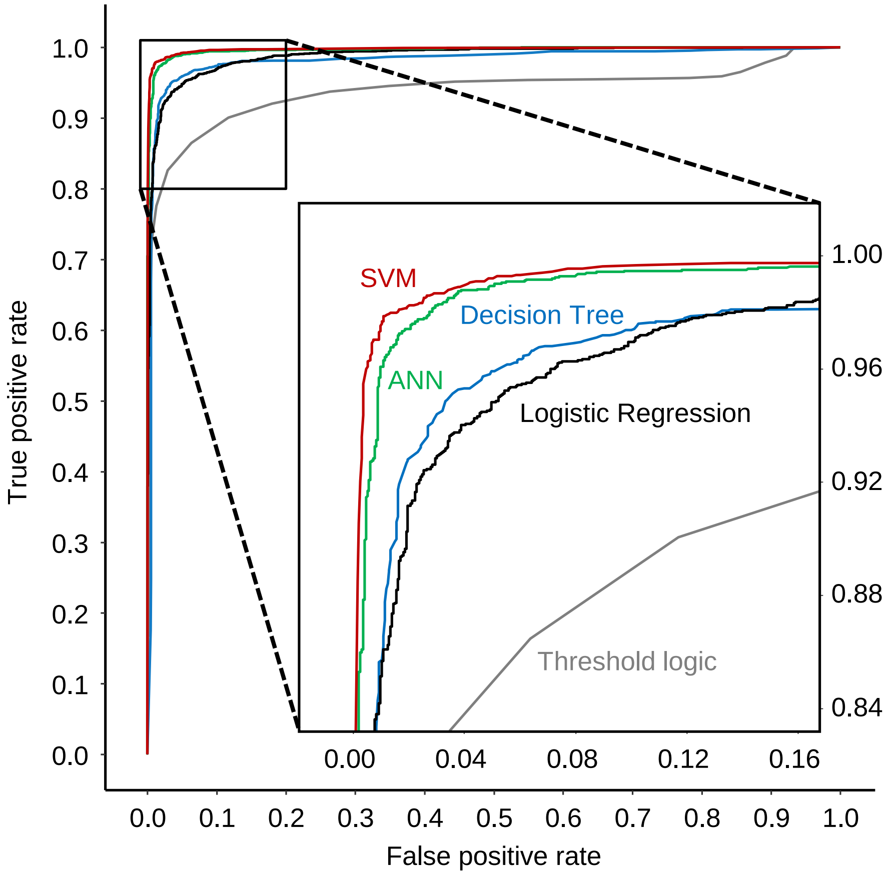

Figure 4.5: ROC curves for the different classification approaches

Table 4.5 shows the classification results for (i) the classification models and (ii) the threshold approach (see Section 4.1).[3] The classification results in Table 4.5 and Figure 4.5 show the superiority of (i) the classification models in comparison to the threshold approach and (ii) the ANN and the SVM in comparison to the decision tree and the logistic regression. The latter result is in line with the results described in Keller, Thiesse, and Fleisch (2014a) and Ma, Wang, and Wang (2018). With greater than 97 % accuracy, greater than 97 %

[3]We obtained different performance measures for the different possible threshold values. Following Keller, Thiesse, and Fleisch (2014a) who consider accuracy the most important indicator, we chose to report the performance measures obtained with the threshold value that led to the best best accuracy value.

precision, greater than 98 % recall, and an AUC value of 0.9975, the SVM arguably achieves the best result. In addition, Table 4.6 shows the results for models that leverage only features used in previous studies (i.e., features F1-F4). A comparison of the results with the results from Table 4.5 clearly shows the importance of the additional features (i.e., features that augment pure signal strength readings with spatial and temporal information) for the reliable differentiation between ’through-the-gate’ and ’by-the-gate’ events.

Table 4.5: Classification results

Classifier	Balanced Accuracy (%)	Precision (%)	Recall (%)	AUC
Decision Tree	94.77	94.44	96.59	0.9817
Logistic Regression	94.34	96.12	93.94	0.9881
SVM	97.60	97.03	98.86	0.9975
ANN	96.95	97.35	97.35	0.9966
Threshold Approach	89.50	94.86	86.47	0.9385

Table 4.6: Classification results when training the models only with features F1-F4

Classifier	Balanced Accuracy (%)	Precision (%)	Recall (%)	AUC
Decision Tree	90.85	95.12	88.64	0.9343
Logistic Regression	91.72	95.20	90.15	0.9433
SVM	94.77	95.45	95.45	0.9832
ANN	94.34	94.40	95.83	0.9821

4.4 Discussion

This chapter was concerned with the use of RFID as a technological enabler for EAS in retail store environments. The practical relevance of our research is given not only by the economic extent of losses due to theft (e.g., value of stolen item, inventory inaccuracies) but also by the cost reduction potential of replacing existing proprietary EAS systems such as electromagnetic or acoustomagnetic systems (GS1 Germany 2007). To address the problem of limited process control on the retail sales floor, we collected a large amount of data in a retail research laboratory mimicking real-world conditions featuring various scenarios (e.g., people carrying garments that are put into shopping bags). We then applied machine learning techniques to distinguish between theft and non-theft events. To this end, we developed novel machine learning features which facilitate reliable identification of multiple moving objects. In doing so, we go beyond prior research on RFID analytics which has almost exclusively focused on standardized processes in controlled environments (e.g., production or logistics facilities). Our study confirms the suitability of advanced analytics to extract valuable information from low-level RFID data streams.

Naturally, there are limitations and future research opportunities inherent to the presented research. First, our empirical data was not collected in the field but in a retail research laboratory. This allowed for rapid experimentation and recording of training data while avoiding major interruptions of store operations. While the experimental setup tried to capture as many particularities of retail environments as possible, the vast number of different store layouts and products ultimately limits the level of generalizability. However, given the integral role of EAS systems in retail stores, experimentation in practice is hardly possible. As a next step, data from several real stores needs to be passively collected to validate the applicability of the classification models in a real-world setting. A richer data set will also offer the potential of refining the classifiers by introducing new features. To further boost predictive power ensemble methods and alternative algorithmic approaches (e.g., deep learning) may help create a more reliable detection system. Moreover, we only considered RFID data in the context of our study. However, these data points may be complemented by additional data sources such as surveillance cameras systems or checkout systems. In addition, the integration of the EAS system in business processes may improve the performance as contextual information (e.g., if a specific item was sold at the check-out desk) can be taken into consideration when assessing whether or not a specific tag in the reading field of the antennas is stolen. Finally, the classification approach described performs ex-post classification in the moment a thief leaves the store. This may be too late for actual capturing of the shoplifter which may necessitate some form of in advance alarm. Using non-aggregated data on the single read level (or aggregate data only for short-time intervals) may allow us to detect tags RFID-tagged objects in the very moment (or shortly after) they pass the gate. Our ultimate objective is to ensure feasibility of our system under real-world conditions to facilitate a subsequent rollout in a real retail environment.

5 Design of Automated Checkout Systems[1]

The second detection system this thesis is concerned with is an automated checkout system. These systems have to (i) detect customers' purchases and (ii) initiate payment processes. We focus on the main challenge of automatically detecting customer purchases. To this end, we develop a system that (i) reliably and timely detects RFID-tagged objects when they are leaving the store and (ii) assigns them to customers' shopping baskets. For this purpose, we take a two step approach. We (i) further develop the transition detection model described in the previous chapter to detect RFID-tagged items leaving the store and (ii) develop a model that associates these items with individual customer shopping baskets. With regard to the transition detection model, we take up the idea of aggregating data only for short time intervals discussed in the last chapter to detect RFID-tagged items at the very moment (or shortly after) they are leaving the store.

Traditional checkout systems are labor-intensive and can be a great source of frustration for customers when having to wait in line (Manyika et al. 2015). Automated checkout systems, on the other hand, promise greater sales due to an improved customer experience and cost savings because less store personnel is needed (Kasiri, Sharda, and Hardgrave 2012; Manyika et al. 2015). With an economic potential of more than $150 billion annually by 2025, automated checkout systems have emerged as the most significant opportunity among smart fashion store applications (Manyika et al. 2015). In addition, automated checkout is a particularly suitable application for the present thesis, as it needs to be embedded in an environment with immutable physical system components (e.g., architectural constraints) and immutable non-physical system components (e.g., established customer behavior patterns) (see design challenges described in Section 2.1).

[1] The chapter is adapted from Hauser et al. (2019) (see Section 1.5).

5.1 Practical Background

To reduce costs, retailers have started adopting self-service technologies which enable shoppers to scan, bag and pay for their purchases with little or no help from store personnel (Litfin and Wolfram 2006; Orel and Kara 2014). These systems, however, offer hardly any improvements over the traditional checkout process with respect to the customer experience, potentially creating new challenges as many customers consider the service as frustrating, irritating and alienating (Meuter et al. 2000).[2] Self-service checkout systems can be roughly categorized into (i) centralized systems at store exits and (ii) decentralized systems (e.g., handhelds, mobile phones) that customers carry with them while moving through the store. Both types of system usually rely on linear or matrix barcodes (e.g., QR codes). The first group comprises self-checkout terminals (e.g., NCR self-checkout systems) and tunnel scanners (e.g., Wincor Nixdorf 360-degree scanners). In the former case, customers themselves must scan the items they want to purchase one at a time. Tunnel systems, on the other hand, rely on cameras that scan the barcodes of items on a conveyer belt, thus requiring customers to simply put their items on the belt. In contrast to centralized systems, decentralized systems allow for the continuous scanning of items while customers are walking through the store. Such portable systems can be handhelds that retailers provide to their customers or even customers' own mobile phones (the latter case requiring that customers install an app that provides self-checkout functionality).

Automated checkout systems scan, total, and charge a customer's purchases to a registered payment account as the customer is leaving the store (Manyika et al. 2015). These systems promise greater sales due to an improved customer experience and cost savings because less store personnel is needed. Automated checkout systems have to detect customers' shopping baskets and initiate payment processes. To solve the detection task, these systems must tackle two subtasks: They have to reliably detect purchased products and assign these to individual shoppers. Figure 5.1 presents an overview of the different checkout systems we identified: we first differentiate between clerk-based and unmanned systems (criterion 'staffing'). Unmanned systems can be further broken down into self-service and automated checkout systems (criterion 'process'). Third, we differentiate between systems

[2]Meuter et al. (2000) found that causes of dissatisfaction with self-service technologies were failure of the technology, design problems in regard to both the technological interface and the service that it offered, as well as customer-based failures, e.g., forgetting one's personal identification number.

with a central point of scanning (e.g., at the store exit) and systems with decentralized points of scanning, that is systems that require scanning at the very moment customers select items from shelves or put them into shopping carts (criterion ‘infrastructure’).

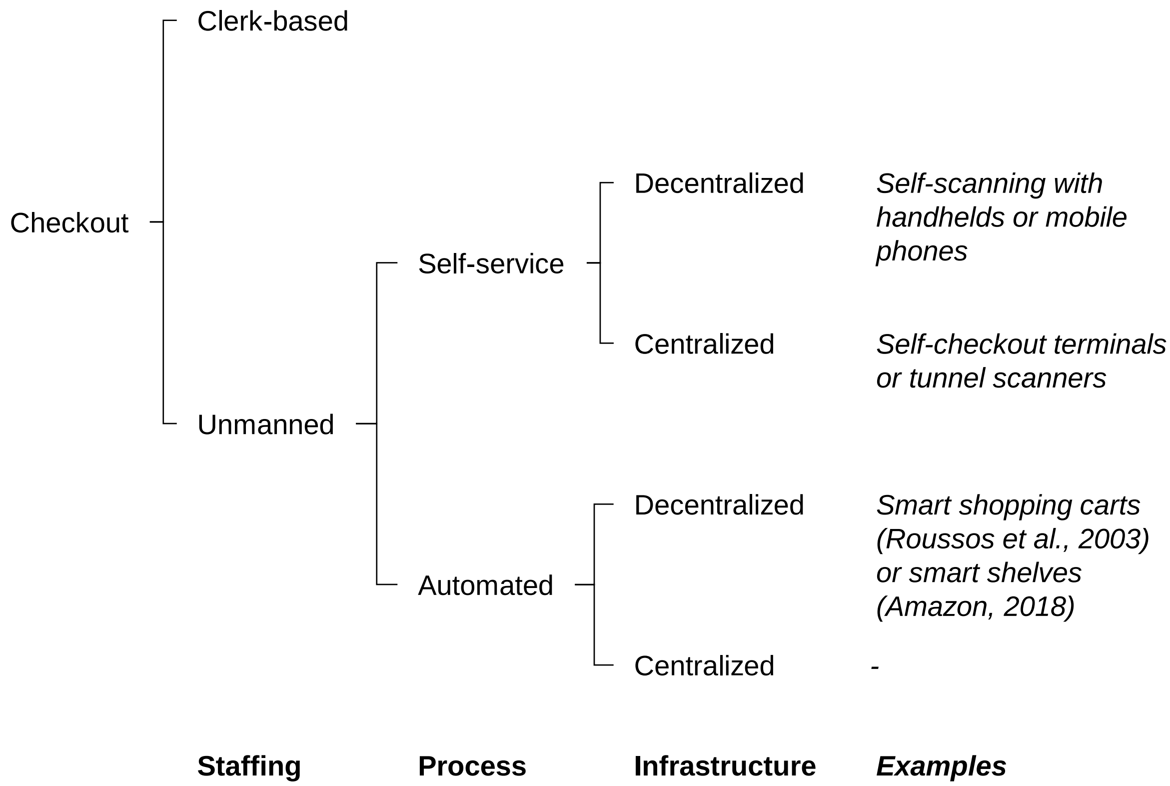

Figure 5.1: Differentiation of checkout systems

The literature on fully automated systems is sparse. To the best of our knowledge, details of only two systems that address the afore mentioned challenges have been published. The first system (MyGrocer) relies on shopping carts equipped with RFID readers that automatically detect objects placed in the carts (Kourouthanassis and Roussos 2003; Roussos et al. 2003). As customers have their own RFID-equipped shopping carts during a shopping trip, the assignment of products to customers is a somewhat trivial task; customers are charged for the products that the RFID reader of their shopping cart has detected. The second system is ‘Amazon Go’, which recently received enormous attention in the media. The system promises to automatically (i) detect products taken from or returned to shelves, (ii) keep track of the products chosen by customers in virtual shopping carts, and (iii) charge the customers’ Amazon accounts after they leave the store. In addition, Amazon promises that all customers need to use their system is an Amazon account, a supported smartphone, and the Amazon Go app to register their entrance into the store (Amazon 2018). Available information suggests that the system stores the inventory locations of all products within Amazon Go stores and mainly relies on cameras to detect products that customers take from or return to particular inventory locations.[3] In addition to the cameras, additional sensors (e.g., pressure sensors, infrared sensors, and RFID readers) as well as customer information (e.g., purchase history) can be utilized to identify and assign purchases.

5.2 System Design

Automated checkout systems must identify customers' shopping baskets and initiate payment processes. We focus on the first task, which entails reliably and instantaneously detecting products and correctly assigning them to shopping baskets. We do not aim at assigning these shopping baskets to individual customers because we consider customer identification as part of the payment initialization process. The main reason for focusing on the identification of shopping baskets is that this task cannot be adequately solved by the automated checkout systems described in the literature. This is because these solutions were developed for supermarket settings which differ significantly from fashion retail environments with respect to in-store processes and the suitability of specific technologies.

5.2.1 Requirements Analysis

The automated checkout solution presented in this chapter was developed in the course of the three-year research project introduced in the Introduction (see Section 1.2). Together with the industry partners within the project, we put forward the following observations and explain how they affected various design decisions:

1. *There are no shopping carts or baskets in fashion retail stores.* We consider this an immutable property of fashion retail, as customers will likely be alienated by fashion stores requiring them to use shopping carts to track their purchases (Litfin and Wolfram 2006). Furthermore, store layouts may not permit carts to navigate the shopping area (i.e., an immutable physical component of fashion store environments).

[3] Although Amazon has not published any technical details about their system, information on the company's website and two patents filed by the company (Kumar et al. 2015; Puerini, Kumar, and Kessel 2015) provide insights into the implementation of this cyberphysical retail system.

In addition, the mental association of bulk shopping with the use of carts and baskets may be detrimental to brand image (i.e., an immutable non-physical component). Another argument against the application of smart shopping carts is that the costs for equipping every shopping cart with RFID sensors are very high (Roussos et al. 2003). In addition, we also expect high operating costs because the individual carts would probably have to be regularly recharged and checked for defects.

2. *Customers in fashion retail stores usually leave unwanted garments in the changing room.* We consider this to be another immutable business process of fashion store environments as some customers might not accept the necessity of going back to search for the shelf from which they picked up a garment.

3. *Usage of cameras is problematic in key areas of fashion stores (i.e., changing rooms).* Several scholars have highlighted the importance of considering the potential intrusiveness of technological innovations in retail stores with regard to customer privacy (e.g., Grewal, Roggeveen, and Nordfält 2017; Litfin and Wolfram 2006).

4. *Major fashion retailers have implemented item-level RFID tagging of products.* Fashion retailers and suppliers first adopted RFID at case-level mainly for inventory management purposes (Hardgrave, Aloysius, and Goyal 2013). Item-level tagging has, however, moved out of the research environment and into mainstream commerce (Barthel, Hudson-Smith, and de Jode 2014). Today, major fashion retailers such as Walmart, J. C. Penney, and Zara have already implemented item-level RFID tagging of products. Leveraging the available sensor infrastructure facilitates a cost-effective and less intrusive integration of checkout systems into existing store environments.

These requirements are violated by the decentralized automated checkout solutions presented in Section 5.1. The first observation rules out automated checkout systems based on smart shopping carts. The second observation rules out automated checkout systems that rely on shelf activity to track purchases. We therefore decided to design an automated system with a central point of detection (i.e., items are detected when customers leave the store). With respect to technology selection, observations 3 and 4 make a very strong case for RFID-based item detection. However, the use of RFID is more challenging than in the MyGrocer project, where carts only need to detect items within them. In our case, the system needs to detect items that leave the store through an exit gate. This requires antennas with a large read range and high power. Unfortunately, this leads to the detection of RFID tags carried near the gate instead of through the gate. Furthermore, assigning items to individual customers is very challenging unless customers wait in line and pass through the gate one at a time. However, prior work has demonstrated that RFID-based solutions can successfully execute diverse and complex processes in retail environments (e.g., Chaves, Buchmann, and Böhm 2010; Li, Ye, and Sample 2015; Parada et al. 2015).

5.2.2 System Architecture and Infrastructure

The architecture of the artifact combines hardware and software components (see Figure 5.2). The hardware consists of two RFID reader installations, a ceiling-mounted system that tracks items in the store, and a gate-mounted system that helps to detect items that are leaving the store. This infrastructure collects low-level RFID data that is then processed by the software components. There are two distinct software functionalities. The first software component uses machine learning techniques to reliably and instantaneously detect items that are leaving the store (item detection approach); the second assigns items leaving the store (identified by the first component) to individual shopping baskets (purchase assignment approach). These shopping baskets are the output of the artifact.

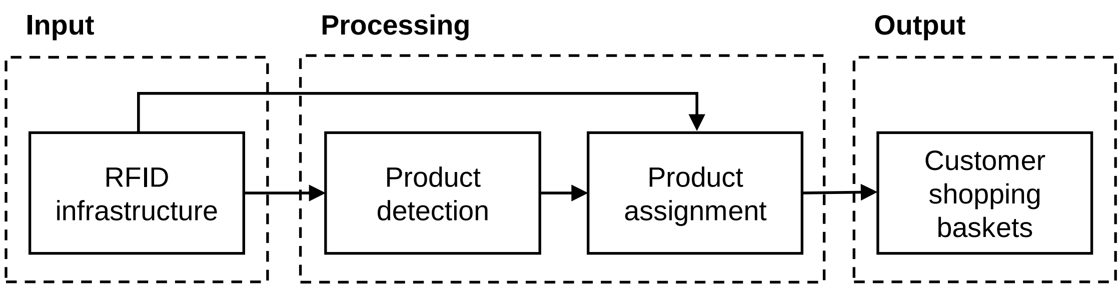

Figure 5.2: Architecture of the automated checkout artifact

Figure 5.3 depicts the infrastructure with two parallel RFID readers from Impinj, a manufacturer of RFID devices and software based in Seattle. The gate-mounted system features four far-field antennas (Impinj Inc. 2017a), the ceiling-mounted system an array with 52 far-field antenna beams mounted in one housing (Impinj Inc. 2017b).

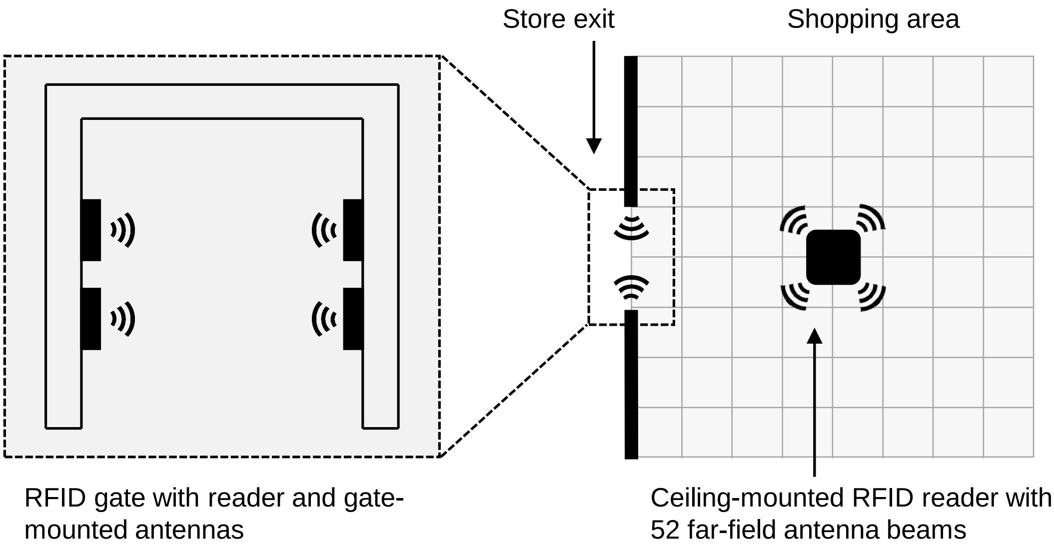

Figure 5.3: Infrastructure with two parallel RFID reader installations

5.2.3 Item Detection Approach

The item detection software component has to reliably detect items that pass through a transition area. If items passing the transition area are not registered, we speak of false-negative events. False-positive events, on the other hand, denote situations in which items that do not pass the transition area are classified as having done so. Similar to prior research (Keller, Thiesse, and Fleisch 2014a; Ma, Wang, and Wang 2018), the approach presented in Chapter 4 considered aggregates for single runs and the classification was thus performed after a tag has moved through a transition area. In contrast, in this chapter we aim to detect products at the very moment they are moved through the gate (i.e., when a person leaving the store is standing right in the middle of the RFID gate). This is important because detecting a shopping basket after a customer has left the store is obviously too late to initiate a payment process. Therefore, to enable continuous evaluation in real time, the RFID data streams first need to be split into chunks. In a second step, these chunks are aggregated to extract predictive features encoding information regarding observed real-world events. These features are then used to train classification models, which automatically map RFID data streams to classification events.

Data Understanding and Preprocessing

Table 5.1 provides a representative excerpt from the raw data gathered by the RFID infrastructure. Each row reflects a single tag read event triggered by one of the RFID readers' antennas. Here, EPC is the unique identifier of the RFID tag, Timestamp is the Unix timestamp of when the tag was read, RSSI indicates the radio signal's power measured in dBm, Phase Angle is the current state of the back-scattered sinusoidal wave, and Antenna is the unique ID of the antenna that read the tag.

Table 5.1: Representative low-level RFID data excerpt

Reader	EPC	Timestamp	Antenna	RSSI	Phase Angle
Ceiling	3032...7D	1453989765.31	15	-59.0	3.50
Ceiling	3032...D1	1453989765.31	15	-56.0	2.91
Gate	3032...7D	1453989765.34	4	-69.0	2.72
Ceiling	3032...7D	1453989765.34	17	-56.0	3.07

We aim to detect products at the very moment they are moved through the RFID gate. Similar to Parlak and Marsic (2013), we therefore first apply a sliding window approach to enable continuous evaluations in real time. A sliding window is a window of a certain size (e.g., detection events of the last two seconds) that is updated at regular time intervals (Jeffery, Garofalakis, and Franklin 2006). Each window contains only detection events from one particular tagged product within reading range of the antennas. Our research determined that window sizes of two seconds offer sufficient information to reliably classify the events. To facilitate real-time evaluation, we apply window shifts every 250 milliseconds.

Feature Engineering

In a second step, we examine the two-second windows and extract features from the raw data stream. These features condense information regarding observed real-world events. The considered features are specific to the RFID analysis task at hand and must be developed based on knowledge of the particular business process in question.

We focus on the development of features that facilitate the reliable identification of multiple moving objects. To this end, we engineered 184 different features for training of the classification models (see Table 5.2). One example of a feature with high predictive power is the maximum RSSI value measured in a series of detections of a particular tag within the two-second windows. Here we first consider the reader level and derive a maximum RSSI value for the gate antenna detections and one for the ceiling antenna detections. In addition, we focus on the individual antenna level and derive values for the detections of the antennas. Maximum signal strength values are standard features considered for the classification of RFID events in previous studies (Keller, Thiesse, and Fleisch 2014a; Ma, Wang, and Wang 2018). These features are very useful in distinguishing static and moving tags, but their ability to distinguish moving objects from other moving objects is limited (see Chapter 4). For this reason, we came up with additional features that put individual readings into temporal relation to one another and augmented them with antenna information. Examples are the parameters of a Gaussian fit of the signal strength values for detections of a particular tag within the two-second windows.

Modelling

Similar to our approach in Chapter 4, we approach the classification problem using a set of standard algorithms: Logistic Regression (LogReg) (Menard 2018), ANN (Bishop 2006), SVM (Chang and Lin 2011), and Gradient Tree Boosting (XGBoost) (Chen and Guestrin 2016). Other methods or ensembles over multiple classifiers should not qualitatively change the results. Similar to our approach in Chapter 4, we again performed hyper-parameter optimization (e.g., different numbers of hidden layers and nodes for the ANN classifier or maximum number of constructed trees for the XGBoost classifier).

Every 250 milliseconds, the data-mining models consider two-second windows of raw data for every tagged item within reading range of the antennas and analyze whether the particular tags have moved through the gate or not. To detect an item that has moved through the gate, the data mining models have to classify at least one of the associated two-second windows as moving through the gate (true-positive event). In this context, associated windows are all the windows containing detection events for a particular item while the item was being moved out of the store. In contrast, to avoid false alarms (false-positive events), the models have to classify none of the two-second windows associated with detections from products that are in vicinity of the gate but have not been moved through it (e.g., products that are carried near the gate or products on shelves close to the gate) as having moved through the gate.

Table 5.2: Item detection model features

Feature	Description
F1-F52	Maximum RSSI measurements of individual xArray antennas
F53-F104	Median RSSI measurements of individual xArray antennas
F105-F156	Number of tag reads of individual xArray antennas
F157-F163	Mean, standard deviation, 0.25 quantile, median, 0.75 quantile, maximum, interquartile range, and median absolute deviation of the RSSI values of the R420 readings
F164-F165	Mean RSSI measurement of the R420 antennas on the right and on the left gate side
F166-F168	Mean temporal shift between the signals' timestamps of the R420 on the right and the left gate side as well as on the top and the bottom
F169-F171	Number of R420 antennas that detected the RFID tag at least once in total, in the first quarter of the time window, and in the last quarter of the time window
F172-F174	Parameters of fitted Gaussian function based on the R420 measurement (height of Gaussian curve peak, position of peak center and parameter that controls its width) of RSSI measurements against timestamps
F175	Regression coefficient of linear regression model based on the R420 signals measured after the maximum signal strength measurement with dependent variable signal strength and explanatory variable timestamp
F176	Quadratic regression coefficients of quadratic regression model based on the R420 measurements with dependent variable signal strength and explanatory variable timestamp
F177	Temporal shift between the mean of the R420 signals' timestamps and the start of the time window
F178	Average deviations of RSSI values of adjacent measurements of the R420 antennas
F179	Sum of the absolute distance values of the R420 measurements (calculated using phase angels of consecutive measurements)
F180	Logical attribute that determines whether all R420 signals have the same signal strength value
F181	Number of Doppler outliers in the R420 measurements (values that are outside of the 1.5 interquartile distance of the second and third quartile)
F182	Mean of standard deviations of the Doppler values of the individual R420 antennas
F183	Number of negative Doppler values in the R420 measurements in the last quarter of the time window
F184	Number of individual RFID tags in reading range of the R420 antennas (unlike the features, this feature does not only take into account the measurements of a particular tag but the measurements of all tags)

5.2.4 Purchase Assignment Approach

The software component for purchase assignments associates items leaving the store (identified by the first component) with individual customers. To this end, we first infer item paths in the shopping area and then apply cluster analysis to group them. The procedure rests on the assumption that the paths of items purchased by one customer are more similar to each other than to paths of other items.

Item Path Determination

We rely on state-of-the-art indoor localization techniques (see Section 2.3.2) to infer item paths. To this end, we first apply the 'Scene Analysis' technique to estimate the position of an object by matching its real-time measurements with the raw data 'fingerprints' at different positions (Liu et al. 2007).[4] We again consider a sliding window approach with window shifts every 250 milliseconds to facilitate continuous evaluation. In contrast to the development of the first software component, we do not, however, rely on window sizes of equal length but split the data such that each chunk contains only detections from one collection cycle covering all 52 successively activated antenna beams of the ceiling-mounted RFID reader. The durations of the physical cycles depend on the number of tags in the antenna field and therefore vary over time. Considering time intervals of equal length would have the drawback that some antenna beams might not yet have been activated. This, in turn, would lead to areas not being covered by the system, thus resulting in undetected items. In the artifact's first software component, we consider time intervals instead of collection cycles because objects that are carried out of the store will definitely be detected by the gate antennas (in contrast to objects that are somewhere within the shopping area in front of the gate). Whereas the data from the ceiling antennas is decisive for the localization of RFID-tagged objects, the gate antennas are more important for the identification of objects that pass through the gate.

[4]The ceiling-mounted RFID system offers a 'Wide Area Monitoring' mode and a 'Location' mode (Impinj Inc. 2017b). The first mode provides information about every read event (e.g., timestamp, signal strength); the second estimates physical coordinates of tags within reading range of the antennas. We cannot use these coordinates for our purchase approach because only the Wide Area Monitoring mode provides the low-level RFID data we need for the item detection approach and the RFID system can only be used in one mode at a time. Another argument against using the coordinates estimated by the system is that in an earlier publication we compared our localization approach with the system's localization mode and achieved better results with our approach (Hauser, Griebel, and Thiesse 2017).

We developed 174 features for the training of the classifiers that help localize tags within reading range of the RFID antennas. Most of them are antenna-based features pertaining to the ceiling-mounted RFID system, but we also leverage the low-level data from the gate-mounted antennas. For instance, a high maximum signal strength from the gate antennas in combination with a low number of reads from the ceiling-mounted reader is a good indicator that an object is very close to the exit. Intuitively, the high maximum signal strength indicates that the person is near the gate, while the low number of readings suggests that the person is facing away from the ceiling-mounted system (i.e., that the person's body is shielding the RSSI signals). A complete list of the features considered in our classification models is provided in Table 5.3.

Table 5.3: Purchase assignment model features

Features	Description
F1-F56	Median RSSI measurements of individual xArray and R420 antennas
F57-F112	Maximum RSSI measurements of individual xArray and R420 antennas
F113-F168	Number of tag reads of individual xArray and R420 antennas
F169	Ratio of the number of xArray measurements to the number of all measurements
F170	Logical attribute that determines whether the xArray measurements cover an entire gathering cycle of the xArray
F171-F172	Number of tag reads of the xArray and the R420 antennas
F173	Number of individual RFID tags in reading range of the two systems' antennas (unlike all other purchase assignment model features, this feature does not only take into account the measurements of a particular tag but the measurements of all tags)
F174	Time difference between the first and the last xArray reading

We apply multiclass classification for solving the localization task, which requires dividing the shopping floor area in front of the gate into grid fields and collecting training data for each of these fields (raw data 'fingerprints'). Here the number of grid fields denotes the number of classes considered in the data-mining model. We consider the same machine learning models as for the first software component (see Section 5.2.3) and again perform hyper-parameter optimization. To determine item paths, we concatenate the most probable locations of individual items over time.

Assignment Process

To assign RFID-tagged items to customers, the automated checkout artifact needs to identify the correct customer for the items that are currently leaving the store. Thus, the task is to group the items within the antennas' reading field such that items in the same group belong to the same customer. We approach the problem by first determining all individual item paths within the antennas' reading range. The procedure for the assignment of items then rests on the assumption that paths of items carried by one customer are more similar to each other than to paths of other items.

Figure 5.4 illustrates the assignment process. The process is triggered every time the first software component detects an item being moved through the gate. The objective then is to determine the other items that also belong to the shopping basket of the item first identified. To this end, we analyze the paths of all items in the antennas' reading field. We first determine whether all the items belong to a single customer by applying a simple threshold rule based on the average Euclidean distance between pairs of items. If all items belong to one customer, we assign them to one customer shopping basket. Otherwise, we use clustering techniques to determine the items that belong with the item that triggered the 'through the gate' event. If the first software component triggers another 'through the gate' event, we repeat the process. This time, however, we exclude items that are already assigned to customer shopping baskets.

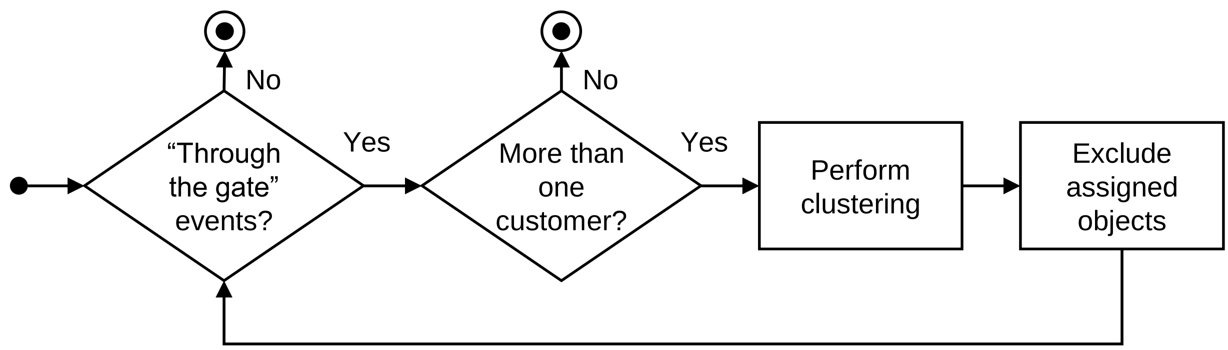

Figure 5.4: Visualization of the process for the assignment of objects to customers

We follow a two-step approach to grouping items. We first determine clusters for every possible number of customer shopping baskets and evaluate each clustering result. Then, in a second step, we choose the best result. To determine the item groups, we use the Partitioning Around Medoids (PAM) clustering algorithm (Reynolds et al. 2006). In order to evaluate the similarity between pairs of tagged items, we again rely on the Euclidean distance. For the evaluation of the goodness of the clustering results, we calculate the average silhouette width for each cluster result, which indicates whether objects are matched well to their own clusters and poorly to neighboring clusters (Rousseeuw 1987).

5.3 Evaluation

We collected large data sets in the laboratory under real-world conditions for instantiation and evaluation of the artifact. The artifact design necessitates, on the one hand, the collection of RFID raw data traces stemming from tests with people carrying RFID-tagged objects and simulating real world customer movements in the experimental shopping area. On the other hand, we need raw data fingerprints at different locations within the shopping area for training of the indoor localization model. Whereas the first data set is used for model training and evaluation of the artifact, the second data set is used only for model training (i.e., training of the indoor localization model).

5.3.1 Evaluation Setting

We set up an experimental shopping area in a retail research laboratory for the evaluation of the automated checkout artifact (see Figure 5.5). The dimensions of our experimental shopping area were 4.8 m by 4.8 m.[5] For the collection of training data for the indoor localization model, we divided this area into 64 grid fields of equal size (0.6 m by 0.6 m).

The artifact design necessitates the collection of (i) RFID raw data fingerprints at different locations within the shopping area for training and testing of the indoor localization model and (ii) RFID raw data traces stemming from tests with people that carry RFID-tagged objects and simulate real-world customer movements in the experimental shopping area. For the collection of the first data set, we collected RFID raw data fingerprints for each of the 64 grid fields within the experimental shopping area. To achieve this, a person carrying garments stood in the shopping area and held the garments such that they were positioned right above one of the fields. During the tests, the garments were moved up and down to reflect real-life shopping situations. We collected approximately two minutes of

[5]The proposed system can be applied in retail environments that are larger than our experimental shopping area because the automated checkout solution we propose requires only observation by RFID systems of the area in front of the store exit and not observation of the entire store.

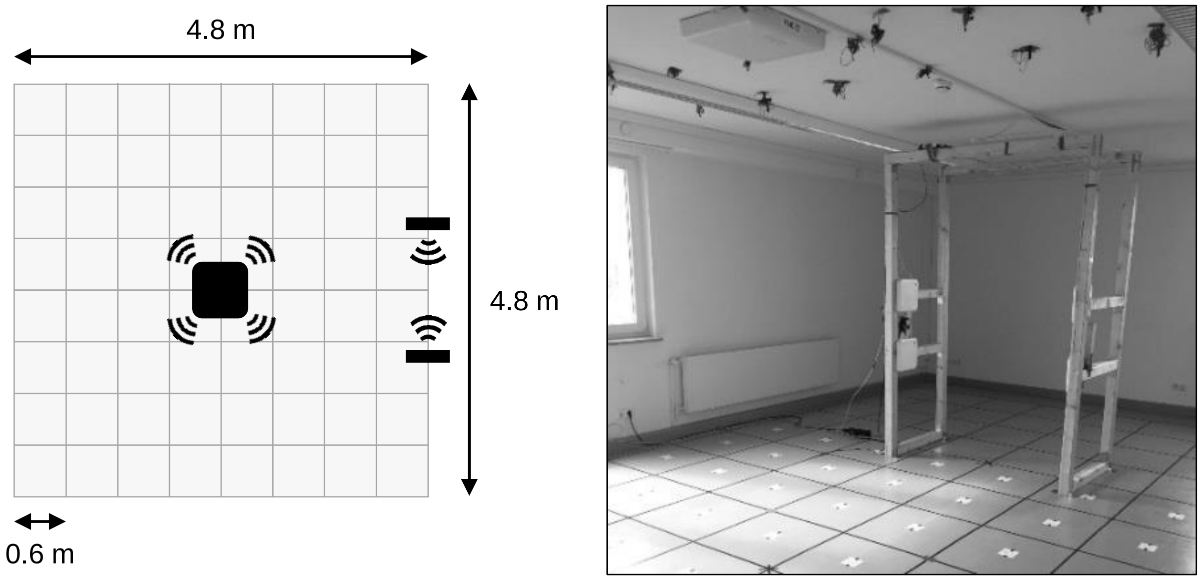

Figure 5.5: Dimensions and picture of the test setting in the laboratory

low-level RFID data for every grid field and two different numbers of tagged items (one and three objects). The resulting RFID data set comprises 1 515 918 individual tag readings.

Similar to our approach in Chapter 4, our experimental setup takes into account the limited process control at store exits by considering multiple walking paths, different numbers of people and RFID-tagged items as well as different movement speeds (i.e., walking and running). Error sources that we identified during our experiments are (i) customers with tagged objects who walk in close proximity to the gate and (ii) customers with tagged objects who leave the store at the same time and on similar movement paths. To account for such settings, we expanded our analysis. Figure 5.6 illustrates the customer movement paths that we consider in our analysis. Training and testing of supervised classification models necessitates labelled data. To obtain precise labels concerning garment position, we additionally installed a light barrier at the gate for the data collection process to identify the exact time a tag was moved through the gate. We did not use the information from the light barrier for the development of our features.[6]

In total, our experimental design includes 18 tests, which are combinations of different movement paths, number of people, number of RFID-tagged items, and movement speeds (e.g., one person with three RFID-tagged items running straight through the RFID gate at

[6]We decided against using the information because (i) our objective is the development of an artifact that facilitates automated checkout with as little hardware investment as possible and (ii) the light barrier requires a direct line of sight and is thus very susceptible to fault in real-world implementations.

the same time or three people carrying three RFID-tagged items each walking on different movement paths). Table 5.4 provides a complete overview of the experimental design. Each of the tests was repeated was repeated 50 times. The data set comprises 1500 runs with a total of 1 431 347 individual tag readings.

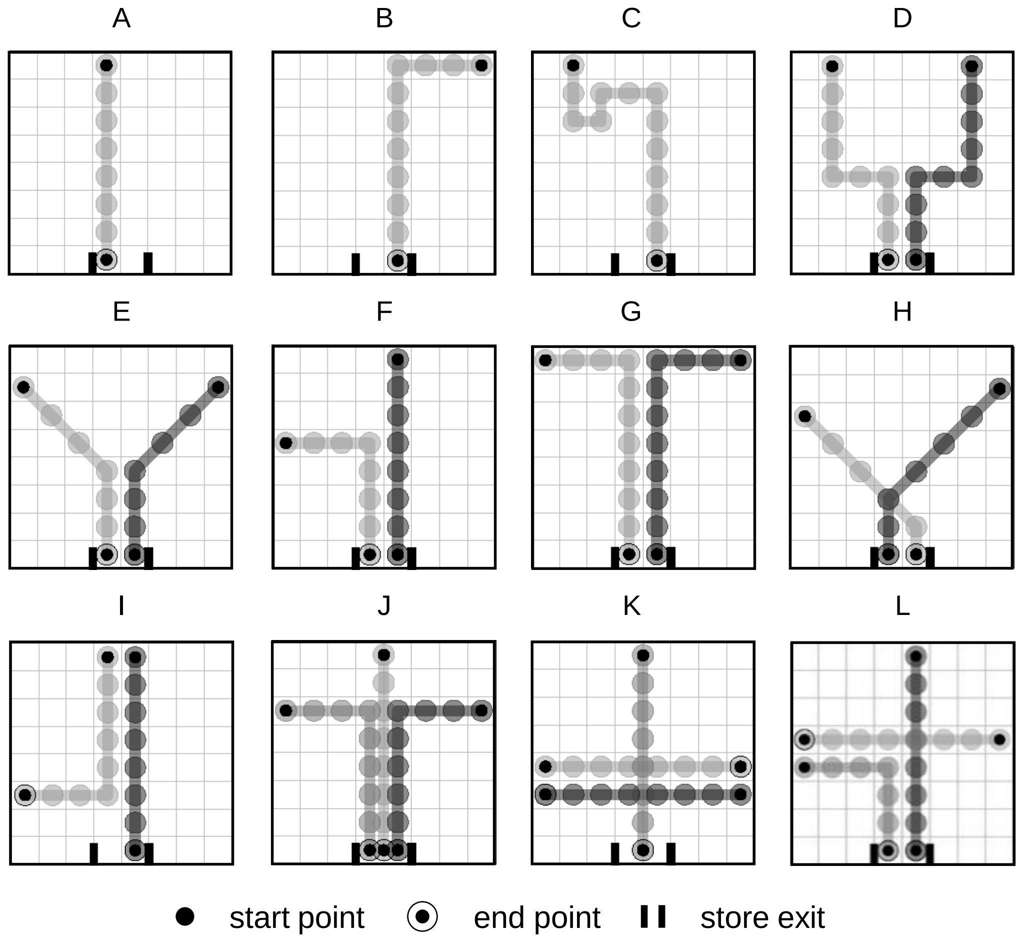

Figure 5.6: Test setting with typical customer movement paths

5.3.2 Evaluation Results

The artifact's item detection and purchase assignment approach rely on supervised machine learning techniques. We thus have to train the models to instantiate the artifact. We use data stemming from the tests dealing with typical movement paths in retail stores (i.e., the second data set) for the training of the item detection component's underlying data mining model. Therefore, we first split the low-level data streams into data chunks (i.e., windows)

Table 5.4: Experimental design

			Movement patterns											
People	Tags	Speed	A	B	C	D	E	F	G	H	I	J	K	L
1	3	Walking	50	50	50	-	-	-	-	-	-	-	-	-
1	3	Running	50	50	50	-	-	-	-	-	-	-	-	-
1	6	Walking	50	50	50	-	-	-	-	-	-	-	-	-
2	6	Walking	-	-	-	50	50	50	50	50	50	-	-	-
3	9	Walking	-	-	-	-	-	-	-	-	-	50	50	50

and calculate the features for each of them (see Section 5.2.3). In addition, we use the RFID fingerprints for every grid field within the experimental setting (i.e., the first data set) for the training of the indoor localization model. To this end, we first split the low-level data streams into collection cycles and calculate the features for each of them (see Section 5.2.4). We then train a classification model with one class for each of the 64 grid cells. We use the collected RFID fingerprints only for the training of the indoor localization model and not for the evaluation of the automated checkout artifact.

The evaluation of the automated checkout artifact is based on the tests dealing with typical movement paths in retail stores (i.e., the second data set). To ensure representative results, we performed 5-fold cross validation: In each round, we used 80 % of the data for the training of the item detection model and the remaining 20 % for the evaluation of the automated checkout artifact. We first evaluate the system's ability to detect, in a reliable and timely fashion, items that are moved through the RFID gate. Subsequently, we evaluate the assignment of purchases to shopping baskets.

Reliability of Detection

In our tests, 4350 items (1300 customer shopping baskets) were carried through the RFID gate and another 600 items (200 customer shopping baskets) were carried around the store but did not leave the shopping floor area (see movement patterns I, K and L in Figure 5.6).

We base our evaluation of the model's reliability on the criteria of *Balanced Accuracy*, *Precision*, and *Recall*. As outlined in Section 5.2.3, items are classified as 'moved through the gate' if the model classifies at least one of the associated two-second windows as moving through the gate. In contrast, to avoid false alarms (false-positive events), the models have to classify none of the two-second windows associated with detections from products that are in vicinity of the gate but have not been moved through it as having moved through the gate. Balanced accuracy is the arithmetic mean of the detection rates of both classes, while precision represents the share of instances classified as ‘moved through the gate’ that were actually moved through the gate. In our application, precision values below 100 % indicate that tags which were not moved through the gate were erroneously classified as ‘moved through the gate.’ Recall measures the proportion of correctly classified ‘through the gate’ instances. For very conservative classifiers that tend to classify instances as ‘not through the gate’ in uncertain cases, recall will be low.

The performance indicators for the four classifiers are summarized in Table 5.5. With the exception of the logistic regression model, all models achieve a high level of classification performance. Recall values of 96.56 % (SVM), 95.47 % (XGBoost), and 98.85 % (ANN) indicate that the models appropriately classified almost all items that were moved through the gate. As outlined in Section 5.2.3, items are classified as ‘moved through the gate’ if the model classifies at least one of the associated two-second windows as moving through the gate. A detailed analysis of the false positive classifications (false alarms) reveals that most errors were caused by false classifications of items that were carried in very close proximity to the gate, but not through it (see movement pattern K in Figure 5.6).

Recall values below 100 % on item level do not necessarily imply that some items might not get assigned to shopping baskets. This is because the item detection component only needs to classify at least one of the items in a shopping basket as ‘through the gate’ in order to trigger the assignment process for the items that are currently within reading range of the antennas. To obtain a more accurate evaluation of the item detection component, we therefore additionally consider classification results at basket level. Table 5.6 presents the evaluation results. A basket is correctly classified as ‘moved through the gate’ if at least one item in that basket was correctly classified as ‘moved through the gate.’ Accordingly, the component correctly identifies shopping baskets that did not leave the shopping floor if it never classifies any of the items in those baskets as ‘moved through the gate.’ With 99.50 % balanced accuracy, 99.85 % precision, and 100 % recall the SVM and the XGBoost achieve the best classification results. The 100 % recall rate indicates that the models detected all the shopping baskets that were moved through the gate.

Table 5.5: Item-level classification results

Classifier	Balanced Accuracy (%)	Precision (%)	Recall (%)
ANN	98.59	99.76	98.85
LogReg	79.23	98.70	64.62
SVM	98.13	99.95	96.56
XGBoost	97.57	99.95	95.47

Table 5.6: Basket-level classification results

Classifier	Balanced Accuracy (%)	Precision (%)	Recall (%)
ANN	97.75	99.31	100.00
LogReg	89.25	97.65	93.00
SVM	99.50	99.85	100.00
XGBoost	99.50	99.85	100.00

Timeliness of Detection

Apart from reliability, the timeliness of detection is important. If the shopping basket of a customer is detected after the customer has already walked through the RFID gate, it may be too late to initiate a payment process. The initiation of a payment process long before the customer actually walks through the gate, on the other hand, could also be a source of potential error because these customers might not yet have made up their mind and, on their way to the exit, decide not to leave the store after all.

Figure 5.7 visualizes the distribution of the detection times (difference between the time at which the item detection component correctly classified the shopping basket as moving through the gate and the time at which the light barrier was triggered by the customer carrying the basket in question). The histograms and boxplots show that the classifiers detected most baskets shortly after the customers walked through the gate. As outlined above, the SVM and the XGBoost classifiers achieved the best classification results at basket level. With the earliest detection occurring at 0.16 s, a 2.5 % percentile value of 0.55 s, a median detection time of 1.03 s, a 97.5 % percentile value of 1.28 s and the latest detection recorded at 1.63 s, the XGBoost classifier arguably detects items faster than the SVM classifier. For this reason, we choose the XGBoost classifier for the item detection component of our automated checkout artifact.

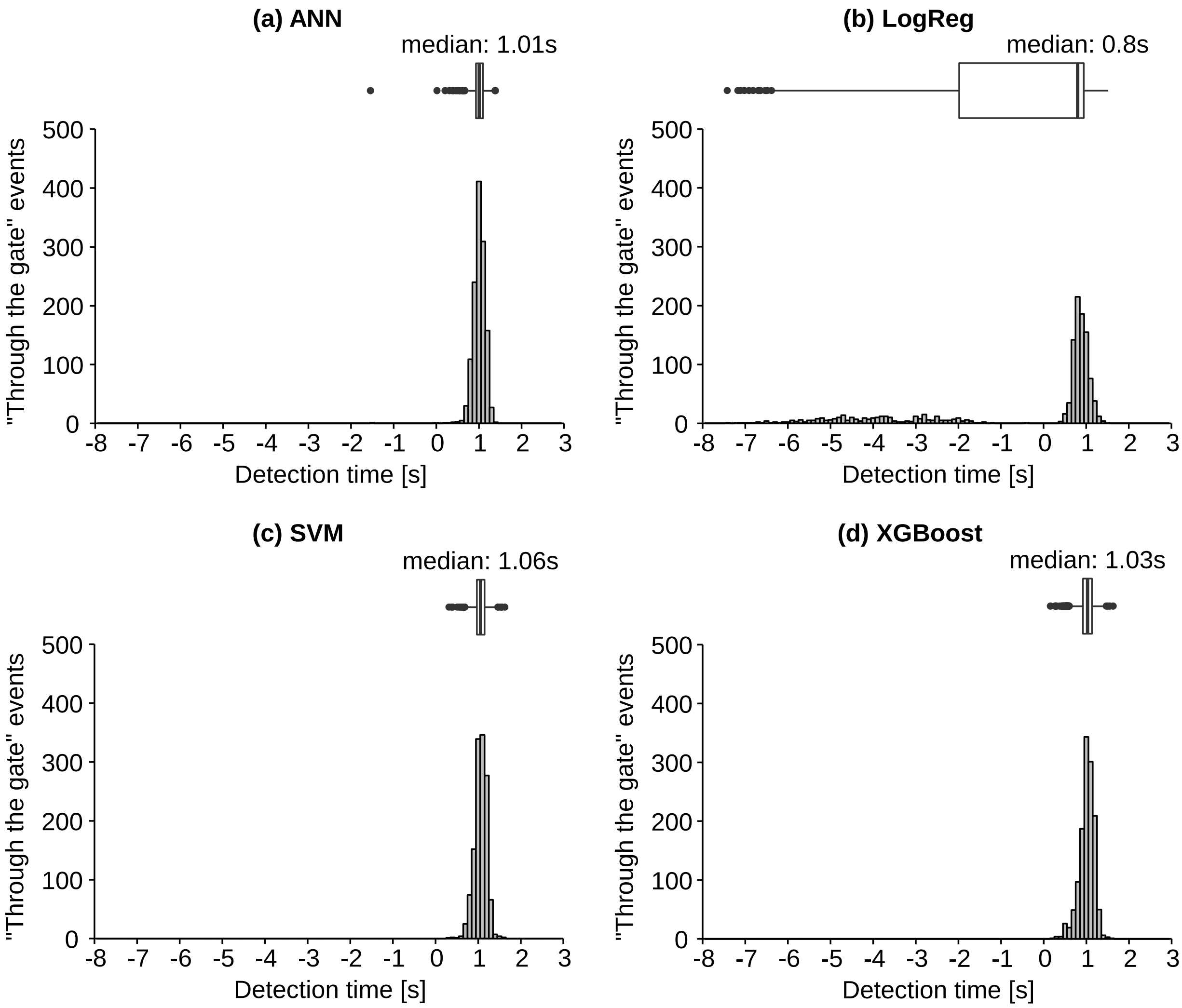

Figure 5.7: Detection time histograms and boxplots with 2.5 and 97.5 percentiles

Purchase Assignment

Every time a basket is detected, the purchase assignment component determines the items that are in the basket by considering the paths of all items within the shopping area in front of the gate. Table 5.7 presents the evaluation results for the different movement patterns in our experiment and the different classifiers that we considered for indoor localization of RFID-tagged items. The results indicate that the component assigns most purchases to customers correctly if we use XGBoost, SVM, or ANN for indoor localization. In all three cases, the misclassifications arise in particularly challenging test scenarios where multiple customers approach the exit gate simultaneously on very similar movement paths. The most difficult movement patterns seem to be I and J. In the first case (movement pattern I), two customers approach the gate next to each other, but one of them turns to the right just before reaching the gate and walks by the gate. Under such circumstances in some of the tests, the component assigns items of the customer not leaving the store to the customer leaving the store. In the second case (movement pattern J), three customers with very similar movement paths leave the store next to each other and at the same time, which results in some items being assigned to the wrong shopping baskets.

Table 5.7: Correctly assigned purchases

Classifier	A-C (%)	D (%)	E (%)	F (%)	G (%)	H (%)	I (%)	J (%)	K (%)	L (%)
ANN	100	100	100	100	100	100	42	50	100	90
LogReg	100	54	62	16	22	66	2	10	84	44
SVM	100	100	100	100	68	100	18	24	100	100
XGBoost	100	100	100	100	100	96	42	70	100	100

5.4 Discussion

The system we proposed in this chapter is the first automated checkout system specifically developed for fashion stores. Existing automated checkout systems were developed for supermarket settings and are not applicable in the fashion retail domain because they either (i) rely on shopping carts or baskets, (ii) use camera systems (which is problematic in key areas of fashion stores), or (iii) require changes to well-established customer processes (e.g., returning a garment tried on in the fitting room to the shelf from which it was picked up). To this end, we conceptualized and implemented an RFID-based system that reliably and instantaneously detects items that are leaving a store and correctly assigns them to individual shopping baskets. In contrast to existing automated checkout solutions, which rely on the continuous scanning of products, we developed a system with a central point of scanning whereby items are detected when customers leave the store.

The proposed artifact is a direct extension of the EAS solution presented in the last chapter. In contrast to the EAS system's item detection approach, we apply a sliding window approach to enable continuous evaluations in real time and detect products at the very moment they are moved through the RFID gate. In addition, the proposed system not only detects the objects that are leaving the store but also assigns them to individual customer shopping baskets to initiate correct payment processes. We find that while most purchases were correctly assigned, our artifact suffered from sub-par performance in more challenging test instances where multiple customers approached the exit gate simultaneously on very similar movement paths. In practice, such situations could easily arise when friends are shopping together, which highlights the limitations of the pilot implementation. Nonetheless, the research demonstrates the fundamental feasibility of RFID-based automated checkout and shows that machine learning techniques can be leveraged to mitigate problems arising from immutable components of the environment in which the system is to be embedded.

Naturally, there are several limitations and future research opportunities inherent to the presented research. First, we see various potential model improvements that might enable us to distinguish between customers, even when their movement paths are very similar. As a next step, we propose the use of probabilistic models to improve the accuracy of item paths (see Hauser, Griebel, and Thiesse 2017).[7] Instead of concatenating the most probable locations of individual items over time, we suggest considering the layout of store areas and characteristics of processes within retail stores to improve path accuracy. The former consideration comprises, for example, information about the location of shelves or walls, while the latter builds on the assumption that particular sequences of item locations within a certain time are more likely than others. To further boost predictive power, ensemble methods and alternative algorithmic approaches (e.g., deep learning techniques) may help create a more reliable system.

Secondly, the integration of additional data sources can improve the assignment process. One possibility is the implementation of additional sensor systems (e.g., camera systems). In cases in which it is not technologically feasible to reliably assign items to customers (even with additional hardware), we propose the inclusion of data from additional information systems in the assignment process (e.g., customer purchase history, sales data, garment characteristics). Product characteristics, for example, could be very helpful in a case in which two customers of different height or of the opposite sex, leave the store at the same time (as they would likely carry very different products with them). This approach is in line with Lee (2008), who suggests that in such cases "the next level of abstraction [...] must

[7] Preliminary results with regard to this idea have been published in Hauser, Griebel, and Thiesse (2017). In this paper, we demonstrate that a hybrid approach based on an ANN and a Hidden Markov Model (HMM) that leverages not only low-level RFID data streams but also information about physical constraints and process knowledge is able to reliably distinguish between RFID-tagged objects within adjusted areas (even in cases where the objects are very close to the borders of the adjacent areas).

compensate with robustness" and notes that "successful designs today follow this principle." Moreover, expanding the monitored area through additional hardware (i.e., installation of more ceiling-based RFID systems) would make it possible to more accurately distinguish item movement paths.

Finally, we did not have access to real-world store data but rather ran experiments in a retail research laboratory. While our experimental setup tried to capture as many particularities of retail environments as possible, the vast number of different store layouts and products ultimately limits the level of generalizability. As a next step, expanding the test setting in the laboratory to scenarios that are more complex (e.g., situations in which customers take objects from shelves that are placed near the exits) could be considered. A richer data set will also offer the potential to refine the classifiers by introducing new features. The ultimate objective is to ensure the feasibility of our system under real-world conditions in order to facilitate a subsequent roll-out in a real store environment.

6 Design of Smart Fitting Rooms[1]

The third detection system this thesis is concerned with is smart fitting rooms. While current implementations of such systems mainly rely on hardware-based solutions (e.g., shielding measures) to reliably detect a customer's product selection, we investigate the applicability of software-based solutions to perform the detection task. In contrast to hardware-based solutions, software-based solutions enable the easy integration of smart fitting room functionality into existing retail store environments. In addition, software-based approaches can also be leveraged to identify not only the garments within individual cabins but also those that are currently most relevant to customers (e.g., garments that they are currently trying on), which allows retailers to improve smart fitting room services. To develop the IT artifact, we again take a two step approach: We (i) further develop the transition detection system described in the previous two chapters and (ii) develop a model that identifies customer interactions with RFID-tagged items. With regard to the transition detection model, the main challenge addressed in this chapter is that the antennas of the proposed RFID infrastructure are not located at the transition areas (i.e., the transitions from outside the cabins into the individual cabins), which necessitates a different approach to enable the reliable and timely detection of RFID-tagged items carried into the cabins.

Smart fitting room applications offer great potential to (i) enhance the customer shopping experience (Melià-Segui et al. 2013; Piotrowicz and Cuthbertson 2014; Walter et al. 2012; Wong et al. 2012) and (ii) support managerial decisions (Thiesse, Al-Kassab, and Fleisch 2009). A popular service is product recommendations, which facilitate cross- and up-selling and can lead to substantial sales increases for retailers (Senecal and Nantel 2004; Walter et al. 2012; Wong et al. 2012). More importantly, smart fitting rooms enable retailers to provide customers with a seamless shopping experience as they offer various possibilities to bridge the gap between the different retail channels by, for example, offering customers the

[1]The chapter is based on Hauser et al. (2017b) (see Section 1.5).

opportunity to purchase products that are currently unavailable in the brick-and-mortar store from the online store while in the smart fitting room (Piotrowicz and Cuthbertson 2014). Additional possible services include (i) the display of additional product information (e.g., available sizes, cuts, or colors of garments), (ii) enhanced customer experience through video projections, lighting effects, and sounds (e.g., playing a surf video when shoppers bring swimwear into the fitting room cabin) (Sjøbakk, Landmark, and Hübert 2016), (iii) the possibility for customers to give feedback about the products they have tried on in the fitting room cabins, and (iv) the possibility for customers to call for assistance in case they need, for example, a different size or color of a particular garment (Melià-Segui et al. 2013). In addition, RFID data gathered by smart fitting room applications can also be used to support managerial decisions. Thiesse, Al-Kassab, and Fleisch (2009), for example, use the information about garments tried on in smart fitting rooms to visualize the number of products that were taken from individual catchment areas of the sales floor, which could help retailers to optimize sales floor layouts, product placements, or the positioning of fitting rooms. The authors also show that combining the RFID data with sales data allows one to analyze merchandise performance by comparing the ratios of try-ons and sales events.

6.1 Practical Background

Retailers like Rebecca Minkoff, Nordstrom, Ralph Lauren, and Bloomingdale's have already started testing fitting rooms that provide additional services to customers (Wahba 2014). While some of these retailers provide customers with barcode scanners for the identification of items (e.g., Nordstrom), others rely on RFID technology (e.g., Rebecca Minkoff). As outlined in Section 2.3, RFID offers several advantages over barcode scanning. In particular, RFID-based object identification (i) does not require a direct line of sight between the tag and the reader device, (ii) allows for simultaneous bulk detection of multiple objects, and (iii) permits the identification of each specific item (Finkenzeller 2015; Want 2006).

Melià-Segui et al. (2013), Thiesse, Al-Kassab, and Fleisch (2009), and Wong et al. (2012) describe pilot projects with smart fitting rooms that rely on RFID technology. While, Thiesse, Al-Kassab, and Fleisch (2009) focus on the value of the RFID data collected by smart fitting room installations to support managerial decisions, Melià-Segui et al. (2013) and Wong et al. (2012) focus on (i) the enhancement of the customer shopping experience and its effect on sales and (ii) customer acceptance of the IT artifacts. Melià-Segui et al.

(2013) describe that sales staff at retail stores with smart fitting rooms reported that the functionalities were well accepted by the customers and led to increased sales. Similarly, Wong et al. (2012) find that recommendations presented to customers in smart fitting rooms in a fashion chain store in Hong Kong led to sales increases of more than 20 %. To cope with the problem of limited process control, smart fitting room solutions described in the literature require the installation of an RFID reader system for each cabin and the use of shielding measures (e.g., shielding paint, thick fitting room walls from floor to ceiling) to ensure that only items within the cabins are detected by the RFID systems (Melià-Segui et al. 2013; Thiesse, Al-Kassab, and Fleisch 2009; Wong et al. 2012). The need for shielding measures obviously complicates the integration of such systems into existing fashion store environments. Another drawback of current implementations is that they cannot distinguish between products that are currently of interest to the customer (e.g., products that they are currently trying on) and others (e.g., products hanging on a coat hook). It is obvious that such information would be valuable for retailers as this would, for example, allow them to highlight recommendations for the most relevant products.

6.2 System Design

We aim to develop a system that leverages RFID technology to (i) detect the garments customers bring into individual cabins and (ii) identify the garments they are currently interacting with. We are again focused on building a system that can be easily integrated into existing (and, over time, constantly changing) fashion store environments. Instead of using (i) shielding measures and (ii) multiple RFID systems, we therefore want to investigate the applicability of machine learning models to achieve these objectives.

6.2.1 System Architecture and Data Understanding

The architecture of the artifact consists of hardware and software components (see Figure 6.1). We propose the use of an off-the-shelf, ceiling-mounted RFID system with multiple far-field antenna beams mounted in one housing. Recently, several such systems have been developed. Examples are Impinj's xArray (Impinj Inc. 2017b), Nedap's ID AR series (Nedap Retail 2018), and Nordic ID's ID TOP (Nordic ID 2018). The sensor infrastructure collects the data, which is then processed by the software components. The first software component uses machine learning techniques to reliably and instantaneously detect items that are carried into fitting room cabins by customers; the second distinguishes between items customers are currently interacting with (e.g., items they are currently trying on) and others (e.g., items hanging on coat hooks in the cabins). The items within the cabins that are currently of the highest relevance to the customers are the output of the artifact.

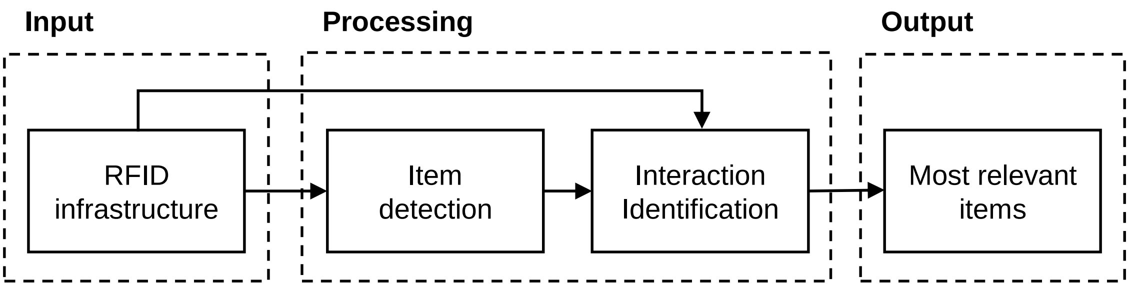

Figure 6.1: Architecture of the smart fitting room artifact

Table 6.1 provides a representative excerpt from raw data gathered with an Impinj xArray (Impinj Inc. 2017b). Each row reflects a single tag read event triggered by one of the reader's antenna beams. Here, EPC is the unique identifier of the RFID tag, Timestamp is the Unix timestamp indicating when the tag was read, RSSI indicates the radio signal's power measured in dBm, Phase Angle is the current state of the back-scattered sinusoidal wave, and Antenna is the unique ID of the antenna that read the tag.

Table 6.1: Representative low-level RFID data excerpt

EPC	Timestamp	Antenna	RSSI	Phase Angle
3032D58E4729FF00000000D1	1453989765.34	17	-58.0	3.21
3032D58E4729FF00000000D1	1453989765.34	17	-58.5	3.18
3032D58E4729FF000000007D	1453989765.35	17	-56.0	3.06
3032D58E4729FF00000000D1	1453989765.36	17	-66.0	5.06

6.2.2 Item Detection Approach

The RFID infrastructure in prior research on RFID-based transition detection systems consists of systems with gate-mounted antennas (see Section 2.3). The antennas are thus directly located at the transition point. Prior research has shown that this infrastructure results in characteristic RSSI profiles for tagged items that are carried through the gate, which can be leveraged in the feature engineering process (e.g., Keller, Thiesse, and Fleisch 2014a). This is because the signal strength increases steeply when a person with a tagged item approaches the gate. Also, the maximum signal strength is usually measured when the person with the tagged item is closest to the antennas, that is, when the person is standing right in the middle of the gate. Prior research (including the studies presented in the previous two chapters) has shown that features that are built based on such observations help to reliably distinguish between tagged items that are moved through an RFID gate and others. In Chapter 4 we consider, for example, the temporal shift between the maximum RSSI values of the antennas on each side of the gate. The rationale here was that carrying tagged items through the gate results in very low temporal shifts, as the antennas should detect the items simultaneously. Similarly, Keller, Thiesse, and Fleisch (2014a) considered so-called logical attributes, which provide a means by which to analyze the order in which items were detected by the antennas of RFID transition portals with two pairs of antennas on each side of the portal, one mounted behind the other.

The antennas of the RFID infrastructure proposed in this chapter are not located at the transition areas (i.e., the transitions from outside the cabins into the individual fitting room cabins). In addition, we want to use only one RFID system for distinguishing between different transitions. As a result, we cannot observe the characteristic RSSI profiles in the data described in the last paragraph. Therefore, many features that have proven useful in previous studies (including the features proposed in the two previous chapters) do not help us to detect the individual cabin transitions. We thus propose a different approach: Instead of focusing on features that help detect the transition of a tagged item, we focus on features that help determine the location of an item and define transition detection time as the time at which the item changed its location from outside the fitting room cabins into one of them. This requires data to be labeled differently than in previous studies. Figure 6.2 illustrates the differences between our approach and the approach presented in the last chapter (considering in both cases exemplary movement paths of customers with one tagged item). To enable real-time classification, the RFID data streams in both cases are split into data chunks—usually referred to in the literature as ‘sliding windows’—before the data is aggregated to extract predictive features. Sliding windows are windows of certain size (e.g., detection events of the last two seconds) that contain only detection events from one particular tagged item within the RF field. In the study presented in Chapter 5, we labeled the window sampled while the item was moved through the gate as ‘moving through the gate' and developed data mining models that aimed at classifying the window accordingly. In this chapter, we aim to predict the current location of an item and regard its transition time as the time at which it gets classified into one of the individual cabins (and label the associated windows accordingly).

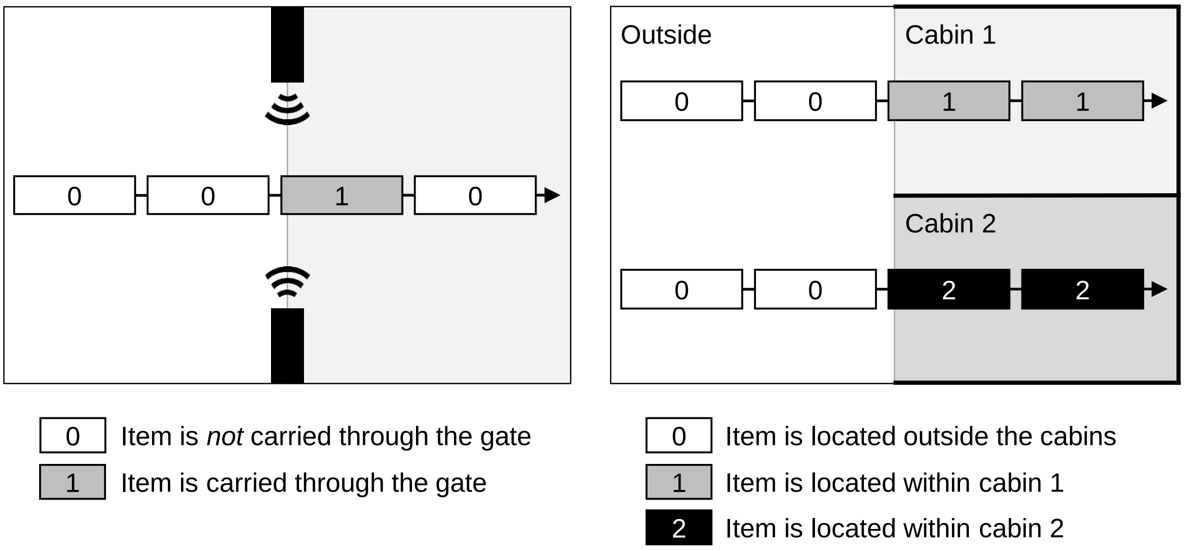

Figure 6.2: Comparison of our data labeling approach (right) with previous studies (left)

Data Preprocessing

We aim to detect the transition of a tagged item at the very moment it is carried into a fitting room cabin. To this end, we first apply a sliding window approach, then aggregate the windows' data to extract predictive features, and finally feed these features into classification models. In contrast to prior research experimenting with fixed windows (e.g., Parlak and Marsic 2013), we propose windows that contain only detection events from one collection cycle covering all successively activated antenna beams of the ceiling-mounted RFID system (similar to the approach presented in Section 5.2.4). Considering time intervals of equal length would have the drawback of some antenna beams not having been activated yet. This would lead to areas not covered by the system, thus resulting in undetected items. The duration of the physical cycles depends on the number of tags in the RF field and therefore varies over time. Despite the varying cycle durations, shifts of arbitrary length between the windows are possible. We choose to evaluate our models every second to ensure regular and frequent item location updates.

Modeling Techniques and Feature Engineering

Similar to previous research on RFID-based transition detection systems (including the studies presented in the previous two chapters), we approach the classification problem using a set of standard algorithms. To train them, features need to be generated from the low-level RFID data that encode the information from observed real-world events. An extensive evaluation of different RSSI-based features (e.g., standard deviation, mean, or median of RSSI measurements) suggests that the maximum measurements of individual antenna beams are well suited to solving the classification task. We attribute this mainly to the fact that maximum values originate from measurements with the most direct path between RFID tags and antennas. In contrast, tag events with low RSSI measurements are often the result of reflections, refractions, diffractions, or absorption of the radio signal (Brusey et al. 2003). Because—in contrast to earlier research on RFID-based transition detection (including the studies presented in the previous two chapters)—we do consider features that take into account the arrangement of the RFID antennas, our approach is based on comparatively simple features and shows a higher level of generalization.

Every second, the data mining models consider data windows that contain detection events from one collection cycle covering all successively activated antenna beams for every RFID-tagged item within reading range of the antennas. We consider the first fitting room that an RFID-tagged item gets classified into to be the fitting room the individual carrying the item has entered. We regard a transition as correctly detected if this location is the cabin that the individual has actually entered (otherwise we regard the transitions as not correctly detected). In case individuals do not enter any of the fitting room cabins, on the other hand, we consider transition detections of items that they are carrying as correct if the localization model never maps the associated data windows into one of the fitting room cabins (i.e., if they are always classified as being outside of the fitting room area).

6.2.3 Interaction Detection Approach

Several scholars have developed systems that detect interactions with RFID-tagged items (see Section 2.3.2). Similar to Parlak and Marsic (2013), we aim to identify items humans are currently interacting with based on low-level RFID data and machine learning techniques. To this end, we rely on the same data windows as for the item detection approach (i.e., windows that contain only detection events from one collection cycle covering all successively activated antenna beams of the RFID system). Similar to Li, Ye, and Sample (2015) and Parada et al. (2015), we furthermore propose using (i) RSSI and (ii) RF phase information to detect interactions with RFID-tagged items. In addition to the maximum RSSI measurement of individual antenna beams, we consider the maximum phase difference (i.e., the difference between two consecutive phase angles of the same RFID tag) of individual antenna beams. As regards modeling, we use the same classification algorithms as for the item detection approach. The training (and test) data must, however, be labelled as ‘moving’ (see definition of object motion) and ‘static’ for training (and evaluation) of the models.

6.3 Evaluation

We collected large datasets in the laboratory under real-world conditions for the instantiation and evaluation of the two software components described above. The first data set comprises raw data traces stemming from tests with people carrying RFID-tagged objects and simulating real world customer movements in the experimental shopping area; the second data from tests reflecting typical behavior and activities in regular fitting rooms.

6.3.1 Evaluation Setting

We set up experimental shopping areas in two research laboratories for the evaluation of the artifact (see Figure 6.3). This allowed for rapid experimentation and recording of training data while avoiding major interruptions of store operations. The experimental shopping areas comprised three fitting room cabins and an RFID system with 52 antenna beams from Impinj (Impinj Inc. 2017b) mounted in the middle of the grid field at a height of 2.5 m. The dimensions of the experimental shopping areas were 2.4 m by 3.6 m and the dimensions of each fitting room 1.2 m by 1.2 m. We chose the cabin layouts and dimensions according to real-world fitting rooms at a leading German fashion retailer with whom we were collaborating in the course of our research (see Section 1.2).

We collected low-level RFID data for training and testing of (i) the transition detection model and (ii) the interaction detection model. For the collection of the first data set, we conducted 30 test cases depicting typical behavior in fitting room areas (four of them are shown in Figure 6.4). The experimental design includes different numbers of people (one and three), different numbers of RFID-tagged objects (one and three), and combinations

Figure 6.3: Three fitting room cabins with one ceiling-mounted RFID system

of seven distinct customer movement patterns.[2] Each test lasted 30 seconds and was performed ten times. The 30 seconds comprise, for example, test cases with three test subjects, each entering one of the three fitting room cabins and trying on selected garments. Garments that were not tried on were put on coat hooks attached to the interior walls of the fitting room cabins. In the course of the data collection process, 1080 RFID-tagged items were carried into individual fitting room cabins. The corresponding data set comprises 1 050 630 individual tag read events. To obtain precise labels concerning garment position, we additionally installed a light barrier to identify the exact time at which a tagged garment was moved in or out of a particular cabin. The lightbarrier information is again neither used nor needed by the software components of the artifact.

We gathered a second data set for training and evaluation of the interaction detection model. The test subjects tried on one of three garments for 20 seconds and then changed to the next garment, which they tried on for another 20 seconds before changing into the last remaining garment. The garments that were not being tried on at any given moment in time were hanging on the coat hooks within the cabins. These experiments were performed with one cabin and one individual, two cabins and two individuals, and three cabins and three individuals with each of these trial runs performed 10 times. As 20 seconds is a short period of time, we only considered shirts as the garments being tried on. To achieve more generalized movement patterns and avoid learning specific body characteristics, there were different individuals executing the tests. As a result, we obtained about 35 minutes of low-level RFID data with 300 276 unique tag read events.

[2]The seven movement patterns comprise walking paths into the three cabins from the right and left side of the shopping area, as well as a walking path leading past the cabins without entering any of them.

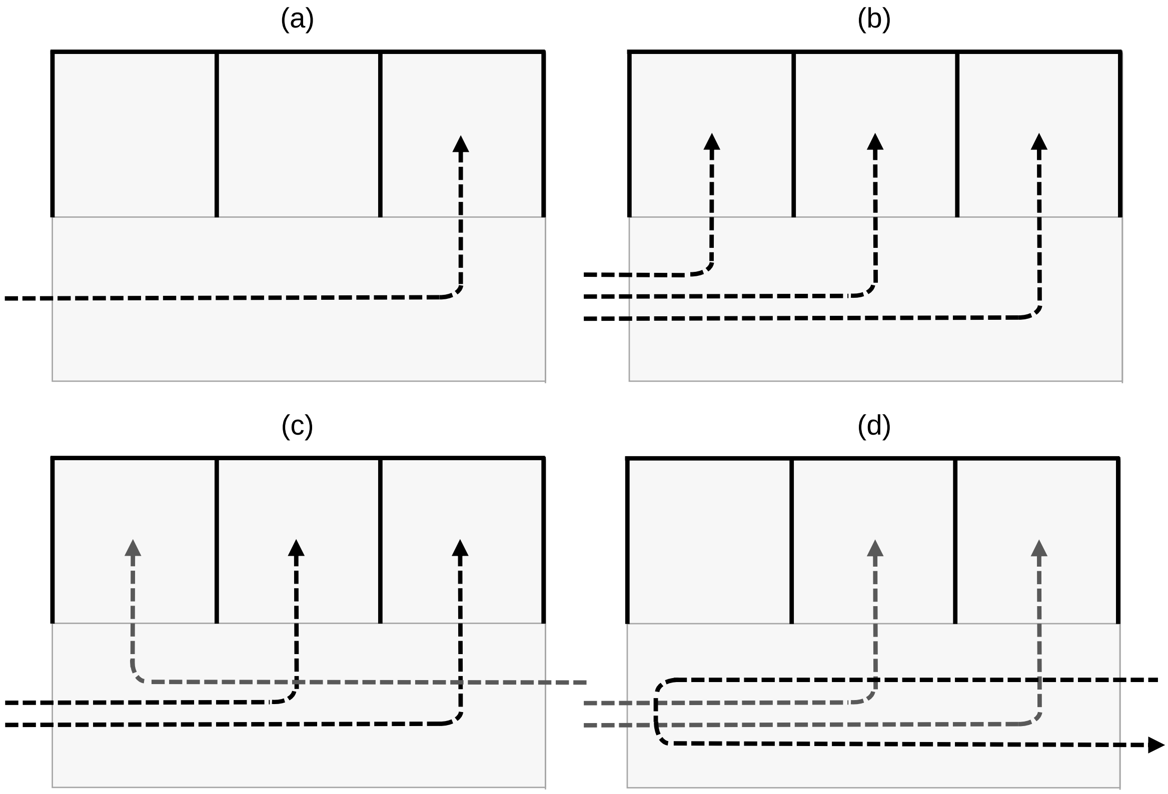

Figure 6.4: Exemplary test cases with typical movement paths in retail stores

6.3.2 Evaluation Results

In this section, we first evaluate the item detection component, referring to the ability of the smart fitting room to detect the items carried into the individual cabins (i) reliably and (ii) in a timely fashion. In a second step, we assess the ability of the system to distinguish items that are being tried on from those that hang on coat hooks.

Evaluation of the Item Detection Component

To evaluate the ability of the system to detect the items carried into the individual cabins, we first have to assess the localization model's ability to correctly predict the locations of the tagged items within the RF field with possible locations in our case being inside the (i) first, (ii) second, and (iii) third cabin, and (iv) outside the fitting room area. To ensure representative results, we performed 10-fold cross validation (with 90 % of the data used for training and the remaining 10 % used for evaluation in each round).

In general, the results of multiclass classification problems can be summarized in the form of a confusion matrix as depicted in Figure 6.5. We base our evaluation of the localization model on the performance measures *Precision* and *Recall*. The class precision ($Precision_i$) is defined as $c_{ii}/\sum_j^N c_{ij}$ and measures whether the items predicted as being in this class are correctly assigned to this class. The class recall ($Recall_i$), on the other hand, is defined as $c_{ii}/\sum_i^N c_{ij}$ and measures the fraction of items within this class that have been correctly predicted. The precision of the class 'Outside,' for example, measures what fraction of items predicted as being outside of the fitting room area are actually outside of this area; the recall of this class, on the other hand, measures the fraction of items outside of the fitting room area that have been correctly predicted as being outside of this area.

		True class		
		Class 1	Class j	Class N
Predicted class	Class 1	c_{11}	c_{1j}	c_{1N}
	Class i	c_{i1}	c_{ij}	c_{iN}
	Class N	c_{N1}	c_{Nj}	c_{NN}

Figure 6.5: Schematic multiclass confusion matrix for N classes

The performance indicators for the classifiers under consideration are summarized in Table 6.2 displaying the class precision values and Table 6.3 displaying the class recall values. XGBoost, SVM, ANN, und logistic regression turned out to be well suited classifiers for the localization model (in contrast to Naive Bayes which performed rather poorly). With an average precision value of 96.95 % and an average recall value of 96.32 %, the XGBoost model achieves the best classification results and was therefore chosen as the classifier for the artifact's item detection component.

Table 6.2: Precision values (%)

Classifier	Cabin 1	Cabin 2	Cabin 3	Outside
ANN	96.67	95.31	93.29	94.50
LogReg	94.29	95.79	93.37	94.63
SVM	97.69	91.66	94.57	95.06
XGBoost	98.40	98.04	95.64	95.73
Naive Bayes	88.01	84.65	68.74	93.67

Table 6.3: Recall values (%)

Classifier	Cabin 1	Cabin 2	Cabin 3	Outside
ANN	95.23	93.83	91.87	96.81
LogReg	96.40	92.77	92.19	95.75
SVM	94.17	95.11	91.31	96.63
XGBoost	97.16	95.52	94.06	98.56
Naive Bayes	87.70	84.85	90.83	77.64

As described above, we consider the first cabin into which a tagged item gets classified as the cabin the person carrying the item has entered. We regard a transition as correctly detected if this location is the cabin that the individual has actually entered. In the case that individuals do not enter any of the cabins, on the other hand, we consider transition detections of items that they are carrying as correct if the localization model never maps the associated data windows into one of the cabins. Following this approach, the system (based on the XGBoost localization model) was able to correctly determine for 95 % percent of the tagged items whether they were moved into a cabin and if so, which one.

Apart from reliability, the timeliness of the transition detections is important. As described above, we define transition time as the time a tagged item changes its location (i.e., the time it gets classified into one of the fitting room cabins by the localization model). The detection time is thus the difference between the time at which the transition detection component correctly classified a tagged item as being within the correct cabin and the time at which the light barrier was triggered by the person carrying the item. Figure 6.6 visualizes the temporal distribution of the detections of the individual garments carried into the cabins.[3] The histogram reveals that the system was able to detect almost all items within two seconds of the light barrier being triggered. The mean detection time is 0.85 s, the median 0.79 s, and the standard deviation is 0.88 s.

Interaction Detection

The evaluation of the interaction detection model's reliability is based on the criteria *Balanced Accuracy*, *Recall*, and *Precision*. Accuracy measures the proportion of instances that are correctly classified. We use balanced accuracy instead of accuracy because the

[3] For this evaluation, only test runs with items (i) that were actually carried into one of the cabins (in contrast to test runs with items that were carried in close proximity to the cabins without entering any of them) and (ii) whose location transition were correctly detected could be considered.

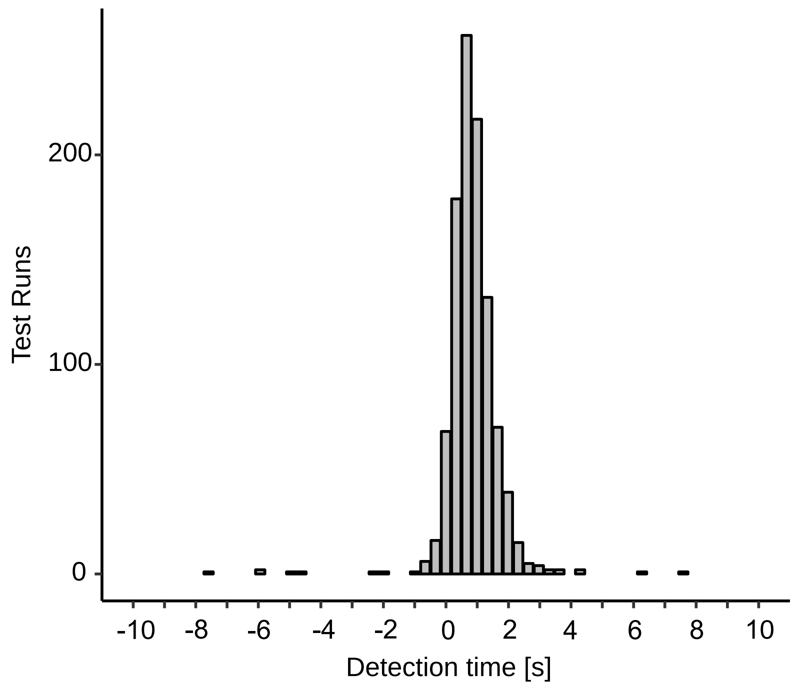

Figure 6.6: Detection time histogram (based on the XGBoost localization model)

data originating from the tests is highly imbalanced. Precision is the share of tagged items classified as 'moving' that actually were being tried on. In our application, if shirts that were hanging on coat hooks are erroneously classified as 'moving', precision is diminished. Recall, on the other hand, measures the proportion of correctly classified 'moving' instances. XGBoost arguably yielded the best results and was therefore chosen as the classifier for the artifact's interaction detection component. This model achieved the highest balanced accuracy at 94.9 % and the highest precision at 94.0 %. The recall of the XGBoost model, however, was 92.5 %, the same level as that of the SVM model.

6.4 Discussion

Our study shows that current limitations (e.g., need for shielding measures) of existing smart fitting room implementations can be tackled with software-based approaches. Our artifact automatically (i) detects the garments within the fitting room cabins and (ii) identifies those that are currently most relevant to the customers in the cabins (e.g., garments that they are currently trying on). The proposed artifact is a further refinement of the artifacts developed in the last two chapters. We particularly showed that item transitions can be detected with an RFID infrastructure whose antennas are not located at the transition areas. Such infrastructures are easier to integrate into existing (and, over time, constantly changing) retail environments and allow the use of one RFID reader for several fitting room cabins. In addition, we show that the low-level data gathered by the RFID infrastructure can also be leveraged to detect items customers are currently interacting with.

Naturally, there are limitations inherent to the presented research that offer opportunities for future research. First, we again see various potential model improvements (e.g., probabilistic models, ensemble methods, deep learning techniques). In addition, the integration of additional data can again improve the item detection approach. One possibility is the inclusion of product characteristics, which could be very useful for distinguishing items in neighboring fitting room cabins. Finally, the empirical data considered for the evaluation of the artifact was again not collected in the field but in a retail research laboratory. However, we recently installed several similar smart fitting rooms in stores of a leading German fashion retailer and are confident that the ongoing tests in these real-world environments will help us to further improve the proposed artifact.

7 Technology Maturity and Optimal System Configuration[1]

The previous three chapters focused on the development of classification models for the detection of object transitions in the context of three frequently discussed smart fashion store applications and different RFID infrastructures (see Figure 7.1). The assessment of the classification performance covered the accuracy of the models in isolation, that is, ignoring their interactions with the associated business processes. However, it is the performance of the technical artifacts embedded in the socio-technical service system that ultimately needs to be assessed. In particular, we want to assess the extent to which the classification models help to meet the business objectives (see description of CRISP-DM's evaluation phase in Section 3.2). This is particularly important because the purely technological assessment assumes constant, homogeneous cost factors for misclassified entities (e.g., Elkan 2001; Fan et al. 1999; Pazzani et al. 1994). As outlined in the introduction, we consider this is a reasonable approach for static settings in which the cost of rework or penalties is fixed (e.g., due to contractual arrangements or internal costing systems). In a retail store, however, the error costs can vary not only according to the type of event but also according to the state of the service system (e.g., time of day, number of customers in the store). The complexity of assessing the economic value of a smart fashion store application is therefore substantially higher than assessing the economic value of an isolated system.

The present chapter goes beyond purely technical performance criteria and shows how the impact of the proposed IT artifacts regarding service quality and costs can be analyzed by means of an analytical model of retail service operations. To exemplify the approach, we consider the simplest transition detection model developed in the course of this thesis (i.e., the classification model developed in Chapter 4) and two of the discussed application

[1]The chapter is adapted from Hauser, Flath, and Thiesse (2019) (see Section 1.5).

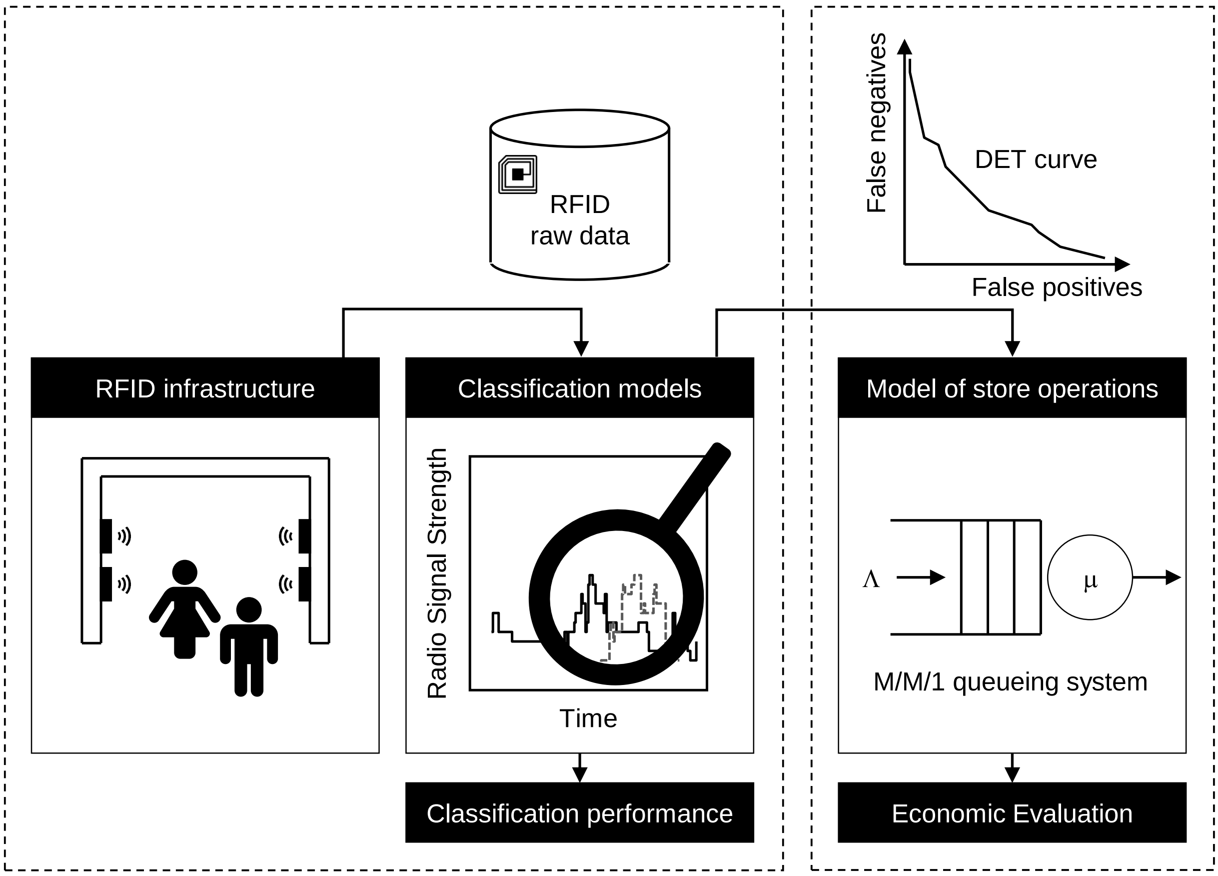

Figure 7.1: Approach followed in the previous three chapters (left) and this chapter (right)

scenarios, namely EAS and automated checkout. The EAS system triggers an alarm if a customer leaves the store with items that have not been paid for; the automated checkout system initiates payment processes when articles leave the store. While the automated checkout system presented in Chapter 5 aimed at (i) detecting RFID-tagged items customers want to purchase and (ii) assigning them to customers' shopping baskets, the present chapter is only concerned with the first of the two tasks.[2]

With regard to the two application cases, the analytical model allows us to (i) determine the costs associated with the different types of false classifications (customer dissatisfaction and unpaid products), (ii) identify the profit-maximizing configuration of the application's underlying transition detection models, and (iii) assess the application's technological

[2]The automated checkout system proposed in Chapter 5 is a direct extension of the EAS solution. The checkout system not only detects all the items customers leave the store with but also assigns these items to customers' shopping baskets. However, a solution that only detects the items could also be installed in fashion stores if customers are willing to leave the store one at a time in order to enable article allocation.

maturity. To appropriately model the interdependencies between the retail environment and the two detection systems, we take a two-step modeling approach: We first characterize the technical capabilities of the detection system using a Detection Error Trade-off (DET) curve (see Section 7.1). Subsequently, we embed this trade-off curve in a service operations model, which determines optimal detection system configurations by internalizing the inherent trade-off between system detection performance (registered true positive transitions) and customer service quality (avoidance of false alarms) (see Section 7.2).

To explore the interplay between detection system configuration and store service quality in a normative manner, we follow Oi (1992) and model a retailing system by means of a queuing model in which retail service capacity is consumed by shoppers to generate sales activity. Unlike standard queuing models, our analysis is not meant to describe a queue in the literal sense, but rather provides a concise means of modeling a congestion-sensitive service product (De Vany 1976). In particular, we follow the reasoning of Ho and Zheng (2004) that customers primarily care about the overall service experience and waiting time and we therefore subsume all retail processes in a single M/M/1 queuing system (i.e., a system with a single server, arrivals determined by a Poisson process, and exponentially distributed service times). System interruptions arising from false-positive events are incorporated in the form of server breakdowns (Krishnamoorthy, Pramod, and Chakravarthy 2014; Thiruvengadam 1963).

7.1 Performance Curve Fitting

Our results show that even a highly sophisticated classifier will still generate false-negative and false-positive misclassifications. In order to determine the optimal trade-off between the two misclassification events and evaluate the performance of the classification models from an economic perspective, we obviously need to better understand their influence on store processes. Knowledge about the actual economic impact of misclassifications can be used to find an optimal configuration for a classifier. In case of the considered binary classification models, a probability value for each detection event is assigned to one of two classes depending on the discrimination threshold. For example, a threshold of 0.5 implies that all events with a probability value smaller than 0.5 are assigned to one class (in our case ‘not through the gate’) and the rest to the other class (in our case ‘through the gate’). A change in the discrimination threshold results in the classifier becoming more sensitive to one type of misclassification, but less sensitive to the other. This way we can adjust the classification results towards points that are more favorable for a given application (Fawcett 2006). The trade-off between the two misclassification errors can be described by a DET curve, which Martin et al. (1997) aptly describe as "a means of representing performance on detection tasks that involve a trade-off of error types." While the previous three chapters were concerned with steps that sought to improve classification performance and thus shift the curve further towards the lower left corner, an optimization model reflecting the respective application context allows finding the best operating point on the DET curve.

To further investigate the trade-off between false-negative and false-positive classifications with the help of an analytical model, we need to encapsulate the empirical classification results in an analytic object. Prior research (Grey and Morgan 1972; Gönen 2006; Pearce and Ferrier 2000) has established ROC curve fitting as a powerful means of representing classification trade-off characteristics. ROC curves, however, plot the true-positive rate against the false-positive rate (see Figure 4.5 in Section 4.3). As we are mainly interested in the trade-off between false-negatives and false-positives, we prefer the information provided by the DET curve and put forward the idea of DET curve fitting. Leveraging the observation of Martin et al. (1997) that DET curves are close to linear in log-log-space, we apply a truncated power law relationship to obtain the false-negative rate $\beta(\alpha)$ as a function of the false-positive rate α, that is,

$$\beta(\alpha) = \max\left\{\mathbb{1}(\alpha = 1), \min\left\{1, a \cdot \alpha^{-k}\right\}\right\} \tag{7.1}$$

where the parameters a and k determine the shape of the curve while the maximum and minimum terms ensure that $\beta(\alpha)$ takes meaningful values on the interval $[0; 1]$.

As outlined above, we consider the transition detection models developed in Chapter 4 for demonstrating how the impact of the proposed IT artifacts on service quality and costs can be analyzed. The fitted curves for the SVM classifier (hereafter referred to as the strong classifier), the logistic regression classifier (hereafter referred to as the weak classifier), and the threshold logic are shown in Figure 7.2. In determining the curve's parameters a and k, we focused on the bottom left corner of the DET graph, with false-positive rate and false-negative rate smaller than 0.2, to accomplish the best possible fit for further analysis. The fitted curves are $\beta(\alpha) = 0.001\alpha^{-0.71}$ (strong classifier), $\beta(\alpha) = 0.005\alpha^{-0.7}$ (weak classifier), and $\beta(\alpha) = 0.014\alpha^{-0.9}$ (threshold logic).

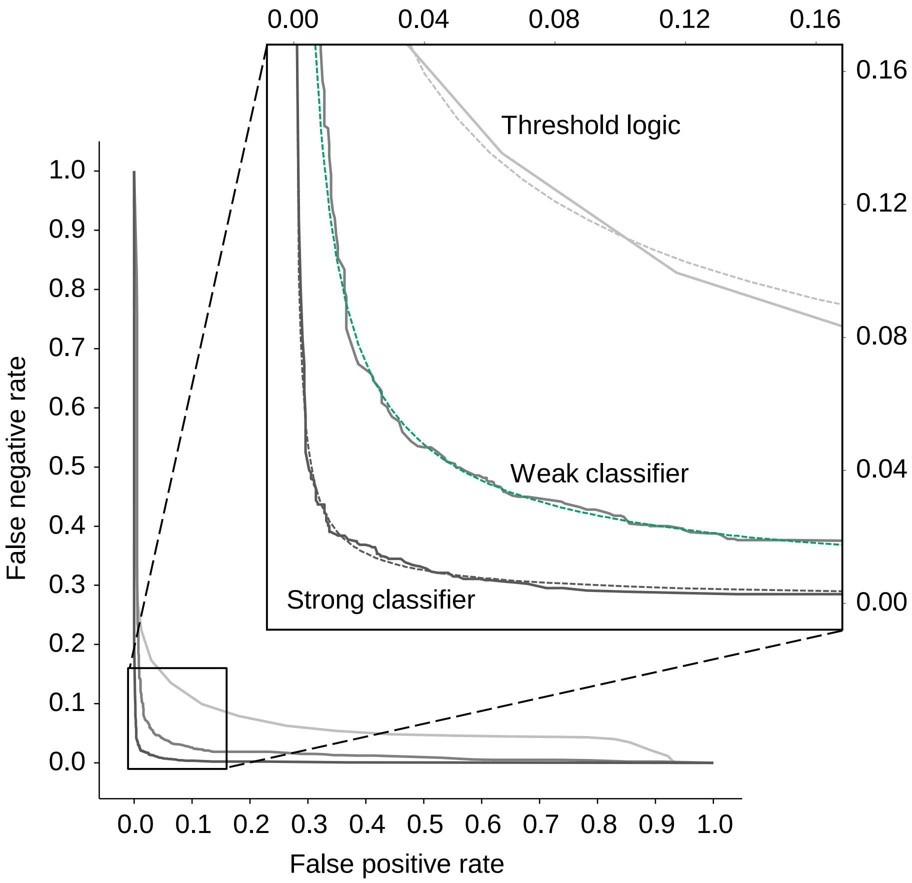

Figure 7.2: DET graph with fitted performance curves

7.2 Service Operations Model

The store traffic stream Λ comprises potential customers. However, this gross potential does not directly translate into sales. Whereas the store traffic stream Λ is exogenous (e.g., depending on the day of the week or the weather), the customer rate λ arises endogenously from the interaction between store traffic and service quality: Individual customers decide whether to buy from the given store after taking into account the store brand as well as service quality as exemplified by the shopping time (Grewal et al. 2003; Messinger and Narasimhan 1997). Customers are served through the store's serving capacity μ, which we normalize at unity. The effects of different classification results of the transition detection models on store processes are different for each of the two application cases:

- In the EAS case, the detection system helps mitigate possible losses from shoplifting. To model this effect, we consider in addition to potential customers a shoplifter population of size $t\Lambda$ where t is the theft rate. A lack of discriminatory power for the detection system may diminish service quality for regular customers and in turn revenues due to false alarms. This is because false-positive events reflect the situation of a customer triggering an alarm despite not being a thief (see Figure 7.3).

- For automated checkout, the detection system seeks to properly register purchases while avoiding the incorrect initialization of transactions when no purchase has taken place. An automated checkout world does not distinguish between thieves and honest buyers as everybody just walks out of the store. Similar to the EAS application, false-positive events again impair service quality (see again Figure 7.3).

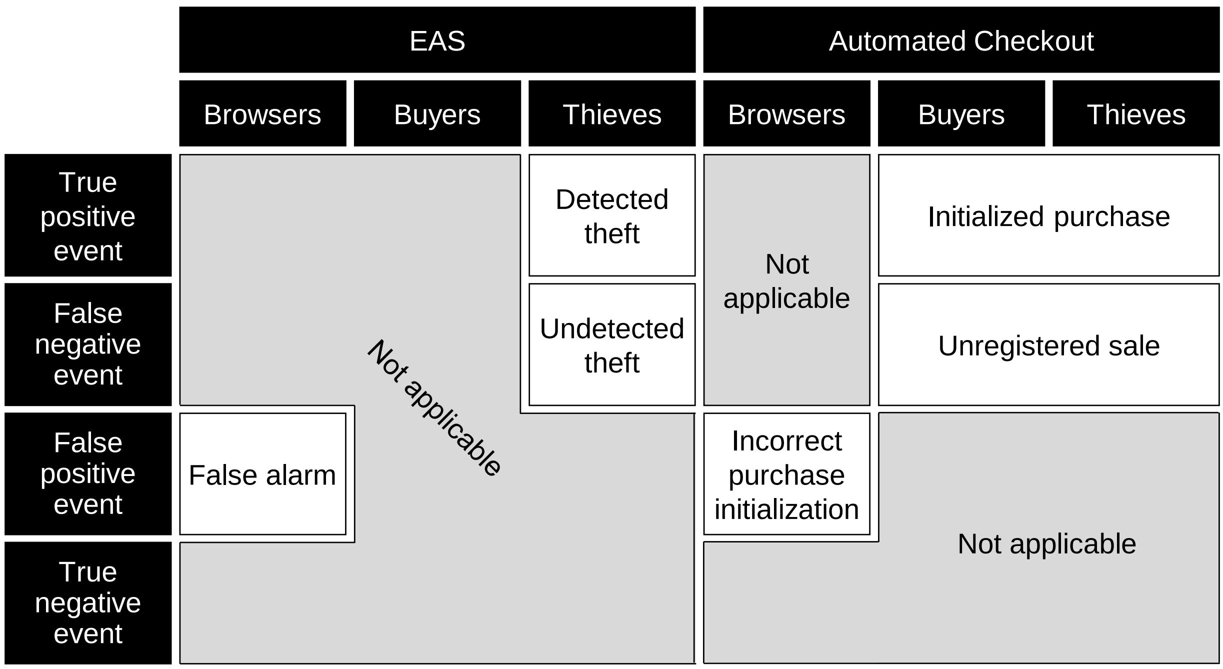

Figure 7.3: Event types for the EAS and automated checkout scenario

Figure 7.4 illustrates the basic elements and relationships of the service operations model. A purchase leads to an average sales margin of m. Similarly, an undetected theft (EAS case) or an unregistered sale (automated checkout case) result in an average loss of c. In both settings, a false-positive detection event occurs with probability α. Leveraging a classifier's DET curve, we can express false-negatives β as a function of false-positives α. In the EAS scenario, a shoplifter remains undetected with probability $\beta(\alpha)$ (false-negative event) and is detected with probability $(1 - \beta(\alpha))$ (true-positive event). In the automated checkout case, on the other hand, a purchase event is registered with probability $(1 - \beta(\alpha))$ (true-positive event) and overlooked with probability $\beta(\alpha)$ (false-negative event). The central trade-off arises from the interplay between false-negative and false-positive events as determined by the detection system sensitivity: Whereas undetected theft or unregistered sales result in direct losses, the economic impact of false alarms is indirect, manifesting itself in reduced service quality, which in turn reduces future sales. Going forward, we seek to characterize the optimal (profit-maximizing) detection system configuration. This setup incorporates results from Grewal et al. (2003), who empirically show that wait-time expectations have a negative effect on store patronage decisions. The notion of service quality deterioration follows Hayes and Blackwood (2006, p. 276), who note that false alarms from detection systems "create collateral damage such as store disruption and customer irritation."

7.2.1 Notations

The notations used throughout the paper are as follows:

Λ	Raw customer stream
b_0	Customer base utility
b_Q	Customer service quality sensitivity
U	Customer utility
λ	Realized store traffic
μ	Service capacity
η	Arrival rate
θ	Service restoration rate
W	Sojourn time
π	Profit
$\mathbb{P}$	Purchase probability
L	Service queue length
λ	Arrival rate
m	Average sales margin
c	Average loss from theft (EAS) or unregistered sale (automated checkout)
t	Theft rate
γ	Capture rate
α	Probability of false-positive classification
$\beta(\alpha)$	Probability of false-negative classification as function of α

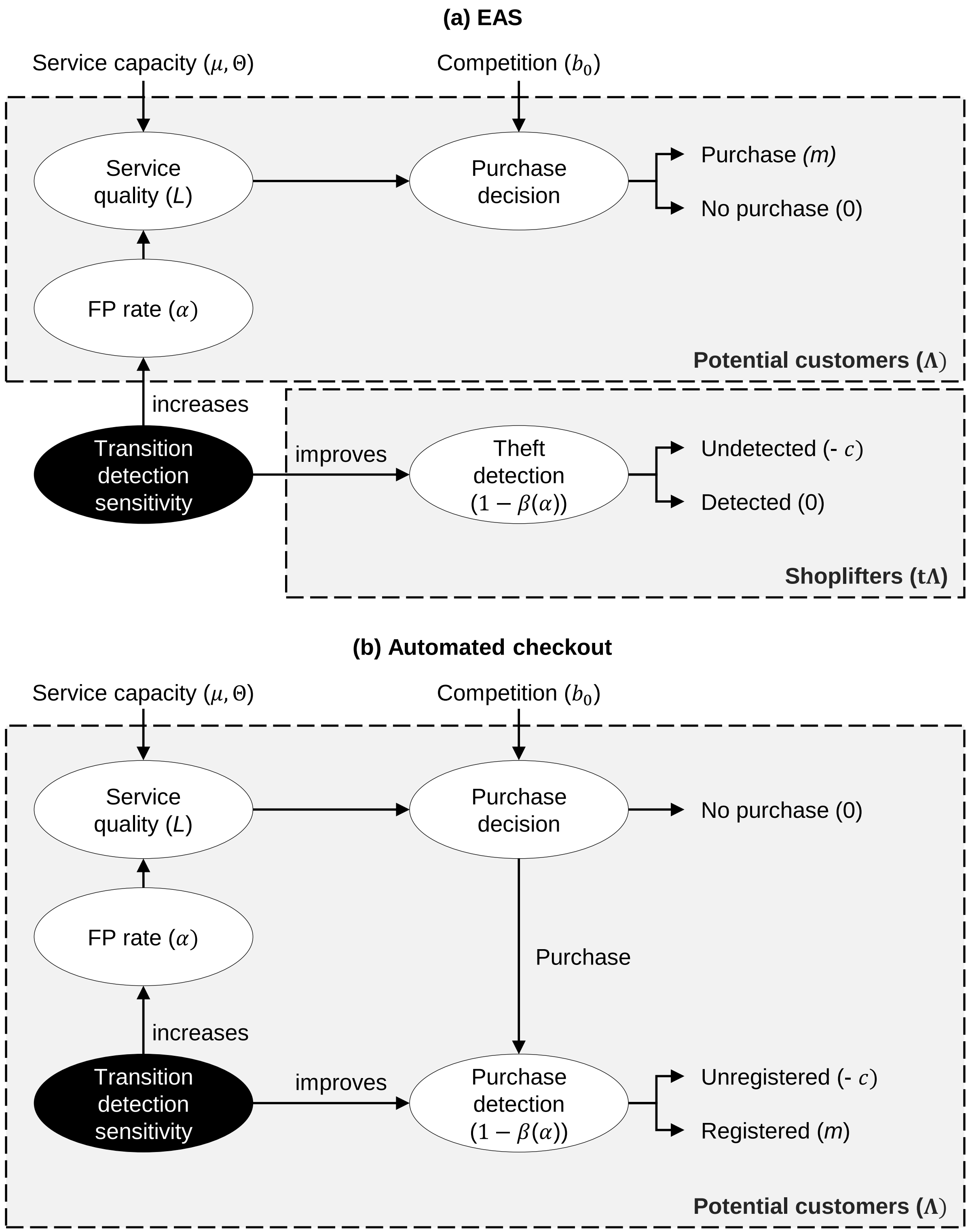

Figure 7.4: Service operations model for (a) EAS and (b) automated checkout

7.2.2 Average Queue Length Determination

In our model, the service queue length L is impacted by two special effects—system interruptions and customer choice. Assuming independent behavior of store customers, it is reasonable to assume that false-positive events constitute a Poisson arrival process with arrival rate $\eta(\alpha) = \Lambda\alpha$, where α is the probability of an false-positive event triggered by the detection system. The false-positive probability α obtains from the detection system configuration. Naturally, any choice of α fixes the corresponding value for the false-negative probability $\beta(\alpha)$ according to the DET function. To retain a lightweight model, the incident management time is assumed to be exponentially distributed with mean $1/\theta$.

The assumption of a Poisson structure for the occurrence and clearance of interruption events allows us to leverage the PASTA property (Wolff 1982) to determine queuing metrics without knowledge of the state probabilities. Specifically, we adapt the solution characterization provided by Adan and Resing (2015, pp. 102-103). We can observe that a customer who decides to shop at the store and thus joins the store service system will expect the average queue length L. Each customer in the queue will require a mean (non-interrupted) service time $1/\mu = 1$. As noted before, service interruptions due to false-positive events occur with a rate of $\eta(\alpha)$. Consequently, each customer will on average experience $\eta(\alpha)L(\alpha, \lambda)$ breakdowns with an average incident management time of $1/\theta$. False-positive events are driven by store traffic Λ and hence may occur even in moments at which there is no store service queue. This leads to an additional service time increase of $\frac{\theta}{(1+\eta(\alpha)/\theta)}$, which allows us to express the expected sojourn time W as

$$W(\alpha, \lambda) = \underbrace{1 + L(\alpha, \lambda)}_{\text{Time waiting for others to be served}} + \underbrace{\frac{\eta(\alpha) L(\alpha, \lambda)}{\theta} + \frac{\theta}{1 + \frac{\eta(\alpha)}{\theta}}}_{\text{Waiting time caused by service disruptions}} . \tag{7.2}$$

Applying Little's law, which states that the long-term average number L of customers in a stationary system is equal to the long-term average effective arrival rate λ multiplied by the average time W that a customer spends in the system, gives us a direct expression for the average queue length $L(\alpha, \lambda)$:

$$L(\alpha, \lambda) = \frac{\lambda\left(\eta(\alpha)(\eta(\alpha) + 1) + 2\eta(\alpha)\theta + \theta^2\right)}{(\eta(\alpha) + \theta)\left((1 - \lambda)\theta - \eta(\alpha)\lambda\right)} \tag{7.3}$$

Lemma 1. *Queue length L is strictly increasing in α.*

Proof. Replacing $\eta(\alpha)$ in Equation (7.3) by $\alpha \cdot \Lambda$ and taking the first derivative with respect to α yields

$$\frac{\partial L}{\partial \alpha} = \frac{\Lambda\lambda\left(\alpha^2\Lambda^2(\lambda+\theta) + 2\alpha\Lambda\theta^2 + \theta^2(-\lambda+\theta+1)\right)}{(\alpha\Lambda+\theta)^2\left(\alpha\Lambda\lambda + (\lambda-1)\theta\right)^2}. \tag{7.4}$$

The denominator is a sum of quadratic terms and consequently is positive. The same applies to the nominator because $\lambda < 1$. □

7.2.3 Demand Equilibrium

We use the logit choice model proposed by Ho and Zheng (2004) and assume the representative customer's utility function for buying at the store to be given by

$$U(L) = b_0 - b_Q L. \tag{7.5}$$

Parameter b_Q reflects customer sensitivity to store service quality (captured by the service queue length L).[3] The other parameter b_0 subsumes the base utility offered by the store's other characteristics (i.e., the utility in the case of perfect service quality). Applying a standard logit choice model (Train 2009), the purchase probability $\mathbb{P}(L)$ obtains as $\mathbb{P}(L) = \frac{1}{1+e^{-U(L)}}$ and the store's realized customer stream is given by

$$\lambda(\alpha) = \mathbb{P}(L)\,\Lambda. \tag{7.6}$$

Ho and Zheng (2004) established that a raw customer stream, combined with a service quality-aware choice function, requires the derivation of a demand rate equilibrium to link the customer rate λ with store traffic Λ. To this end, we use the queue length characterization $L(\alpha, \lambda)$ and reformulate the store choice formulation:

$$\lambda = \mathbb{P}(L(\alpha, \lambda))\,\Lambda \tag{7.7}$$

Any value of λ^* which equates the right-hand side of Equation (7.7) with itself constitutes a demand rate equilibrium with respect to the chosen false-positive rate α. It can be

[3] Following Messinger and Narasimhan (1997), the key metric is of course waiting time and not queue length. However, following Little's law the two can be used interchangeably and using queue length simplifies some of the formulations.

shown that such a value can always be identified, potentially using numerical methods. Furthermore, the relationship between α and λ^* can be characterized and it obtains that a more sensitive transition detection system will result in a strictly lower customer rate.

Proposition 2. *For any α choice, a unique $\lambda^*(\alpha)$ exists such that Equation (7.7) holds.*

Proof. The proof is analogous to Proposition 1 in Ho and Zheng (2004). The function $\mathbb{P}\left(\Lambda\left(\alpha,\lambda\right)\right)\Lambda$ is decreasing and continuous in λ with $\mathbb{P}\left(\Lambda\left(\alpha,\lambda\right)\right)\Lambda > 0$. Therefore, $g(\lambda) = \mathbb{P}\left(\Lambda\left(\alpha,\lambda\right)\right)\Lambda - \lambda$ is strictly decreasing in λ with $g(0) > 0$ and $g(\infty) \rightarrow -\infty$. □

Lemma 3. *The equilibrium demand rate λ^* is strictly decreasing in α.*

Proof. Following Lemma 1, queue length L is increasing in α, and in turn, $\mathbb{P}\left(L\left(\alpha,\lambda\right)\right)$ is decreasing in α. Revisiting the proof of Proposition 2, for a given Λ value the corresponding λ^* realization will be strictly smaller for a ceteris paribus higher choice of α. Figure 7.5 illustrates the mechanics of this proof. □

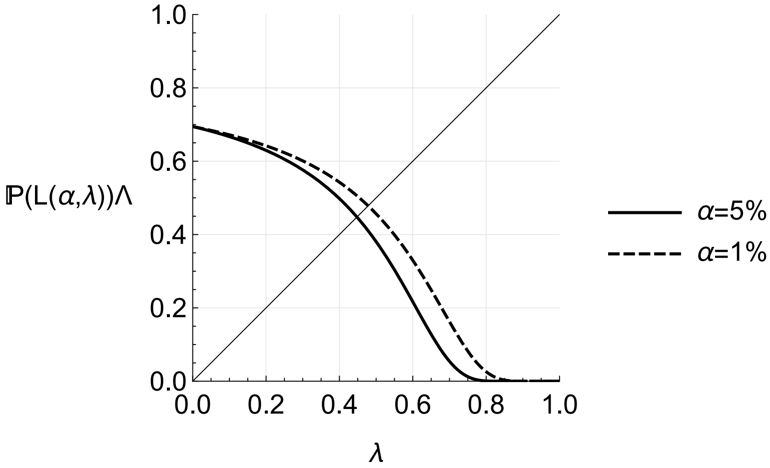

Figure 7.5: Demand rate equilibrium for different α values (Λ=0.95, b_Q=1, b_0=1, θ= 0.5)

In the following we drop λ from $L(\alpha,\lambda)$, implicitly assuming $\lambda^*(\alpha)$.

7.2.4 Profit-maximizing System Configuration

In both scenarios, the store profit function internalizes the trade-off between interruptions of store operations due to false alarms (driven by $L\left(\alpha\right)$ being increasing in α) and the direct costs of undetected theft or unregistered sales (driven by $\beta\left(\alpha\right)$ being decreasing in α).

Assuming the average customer contribution margin m the average loss per theft c, and a capture rate γ, the store's profit stream for the EAS case is

$$\pi(\alpha) = \underbrace{\Lambda m \mathbb{P}(L(\alpha))}_{\text{sales}} - \underbrace{\Lambda ct(\beta(\alpha) + (1-\beta(\alpha))(1-\gamma))}_{\text{theft}}. \tag{7.8}$$

For the sake of exposition, we assume a perfect capture rate $\gamma = 1$. For the automated checkout scenario, on the other hand, the store's profit stream is given by

$$\pi(\alpha) = \underbrace{\Lambda \mathbb{P}(L(\alpha))\, m(1-\beta(\alpha))}_{\text{registered sales}} - \underbrace{\Lambda \mathbb{P}(L(\alpha)\, c\beta(\alpha))}_{\text{unregistered sales}}. \tag{7.9}$$

Here, we assume that the payment process can be initiated and carried out correctly for each of the detected items. The optimal (profit-maximizing) transition detection configuration obtains as a maximization problem of π with respect to α.

Proposition 4. *There exists a unique profit-maximizing choice of α.*

Proof. We proceed by illustrating the EAS case. The automated checkout case obtains in an analogue fashion. The profit function is a linear combination of a sales term (hereafter referred to as π^C) and a theft term (hereafter referred to as π^T.). Due to Equation (**??**), the theft term is clearly concave and strictly increasing in α on the interval $(\underline{\alpha}, 1)$. From Lemma 3 and Equation (7.8), we can establish that the sales term is strictly decreasing in the α choice. Furthermore, because π^C is continuous and bounded, we know that $\frac{\partial \pi^C}{\partial \alpha}$ is finite at both $\alpha = 0$ and $\alpha = 1$ based on the mean value theorem. These observations lead to the following three generic cases:

- $\frac{\partial \pi}{\partial \alpha}|_{\alpha=1} > 0$: The most negative derivative of the sales term is too small to overcome the least positive value of the derivative of the theft term at $\alpha = 1$. In this setting, the most extreme EAS configuration ($\alpha = 1, \beta = 0$) is optimal.

- $\frac{\partial \pi}{\partial \alpha}|_{\alpha=\underline{\alpha}} < 0$: The most positive derivative of the theft term is too small to overcome the least negative derivative of the sales term at $\alpha = \underline{\alpha}$. Therefore, EAS is abandoned with $\alpha = 0$ and $\beta = 1$.

- $\frac{\partial \pi}{\partial \alpha}|_{\alpha=\underline{\alpha}} > 0$ and $\frac{\partial \pi}{\partial \alpha}|_{\alpha=1} < 0$: Applying the mean value theorem, we know that $\frac{\partial \pi}{\partial \alpha}$ has a unique root on the interval $(\underline{\alpha}, 1)$. Observing the derivative signs, we verify this to be the global store profit maximum. □

7.3 Technology Maturity Assessment

Any technological innovation being considered for deployment needs to demonstrate its usefulness by answering two basic questions: (i) Does it work and (ii) how much better is it than an alternative solution? To approach these questions, we combine the classification models and the service operations model (with base choices for the parameter values) to determine the store profitability for (i) the EAS application scenario (see Figure 7.6) and (ii) the automated application checkout scenario (see Figure 7.7). This approach allows assessing the performance of the applications embedded in retail environments and goes thus beyond the purely technical assessment presented in the previous three chapters.

Considering the EAS application scenario, we see that the accrued benefits are strictly increasing in detection system performance—the best classifier (i.e., the strong classifier) generates a 4 % profit increase over the simple threshold logic. Furthermore, the optimal system configuration shifts towards lower false-positive rates—the optimal α is around 5.5 % with the weak classifier and below 1 % with the strong data mining-based classifier.

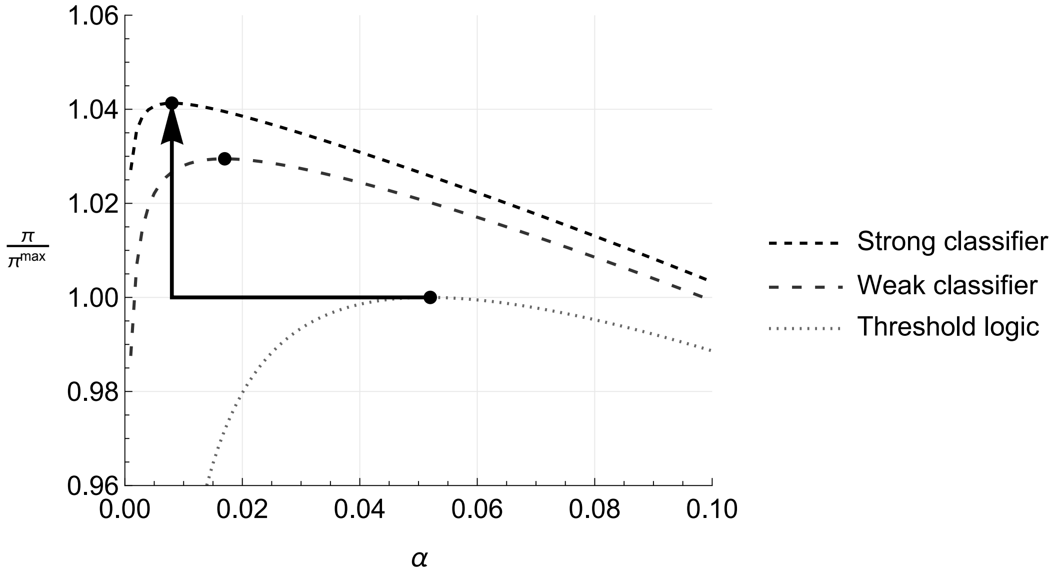

Figure 7.6: Store profitability for the EAS scenario
(t=0.01, m=0.1, c=1, b_Q=0.1, b_0=1, Λ=0.9, θ=0.5)

For the automated checkout scenario, the baseline threshold solution results in an optimal false-positive rate of almost 50 % which means that every second customer in reading range of the system must be checked manually, which essentially defeats the purpose of an automated checkout system. However, we see drastic improvements with regards to profitability (an increase of 33 % for the weak classifier and of 75 % for the strong classifier) and false-positive rate (around 23 % for the weak classifier and around 9 % for the strong classifier) for the better transition detection systems. Yet, even for the best classifier, the optimal false-positive rate still seems quite high (i.e., almost every tenth customer in reading range of the system must be checked manually).

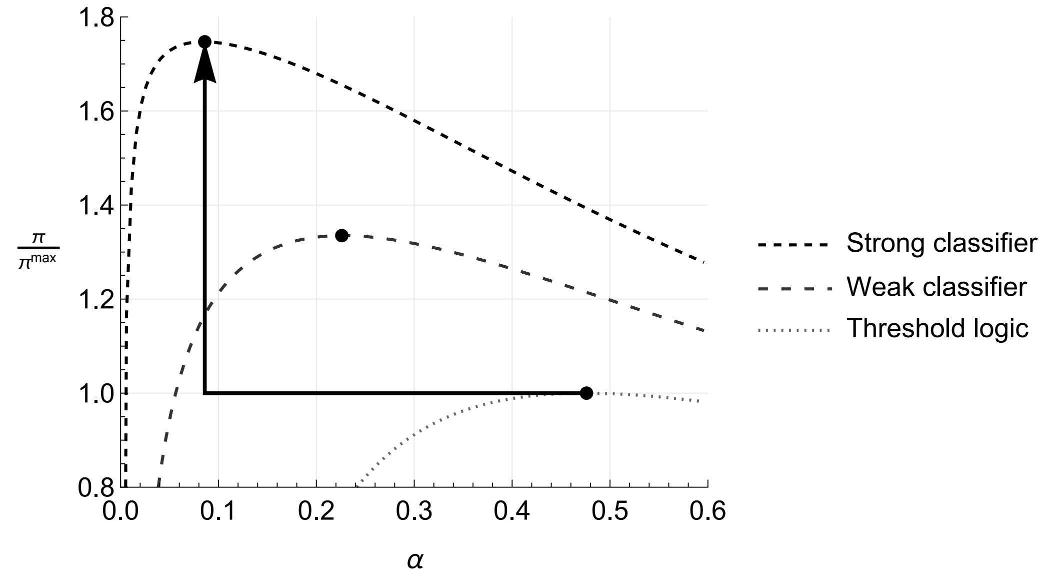

Figure 7.7: Store profitability for the automated checkout scenario (t=0.01, m=0.1, c=1, b_Q=0.1, b_0=1, Λ=0.9, θ=0.5)

The initial evaluation provides us with three key insights: (i) system configuration matters from an economic perspective, (ii) better classification performance directly translates into enhanced business value, and (iii) the benefit increase is more pronounced for more sophisticated application scenarios. The assessment of the technological maturity of two application scenarios, however, is very different. While the EAS results are promising and suggest a continued analysis of the application, the results of the automated checkout systems are—despite the great benefits from better transition detection—insufficient for any real-world deployment. Going ahead with the latter application will thus require improved hardware to increase the reliability of the transition detection system.

Given these initial assessments, we focus in the remainder of the chapter on the evaluation of the EAS application using the strong classifier.

7.4 Numerical Analysis and Optimal Configuration

Clearly, the model setup outlined in Figure 7.4 relies on qualitatively defined relationships of different model parameters. Because of this, we perform a sensitivity analysis that pursues two main objectives: We want to assess (i) the model's robustness and the effect of certain assumptions and parameter choices and (ii) whether adapting the system configuration to the current store situation provides benefits for retailers.

7.4.1 Effect of Incident Management Effectiveness

The first sensitivity analysis is concerned with service restoration upon false-positive events. For higher incident management effectiveness, service restoration will take less time leading to smaller reductions in service quality. Figure 7.8 illustrates the optimal choice of the EAS configuration α^* for different incident management effectiveness levels θ. For each effectiveness level, profitability exhibits a unique maximum value. Any upward or downward deviations from the optimal α choice result in markedly reduced store profits. The α^* values are decreasing in incident management times $1/\theta$. This is because more effective incident management renders store interruptions less harmful with respect to service quality.

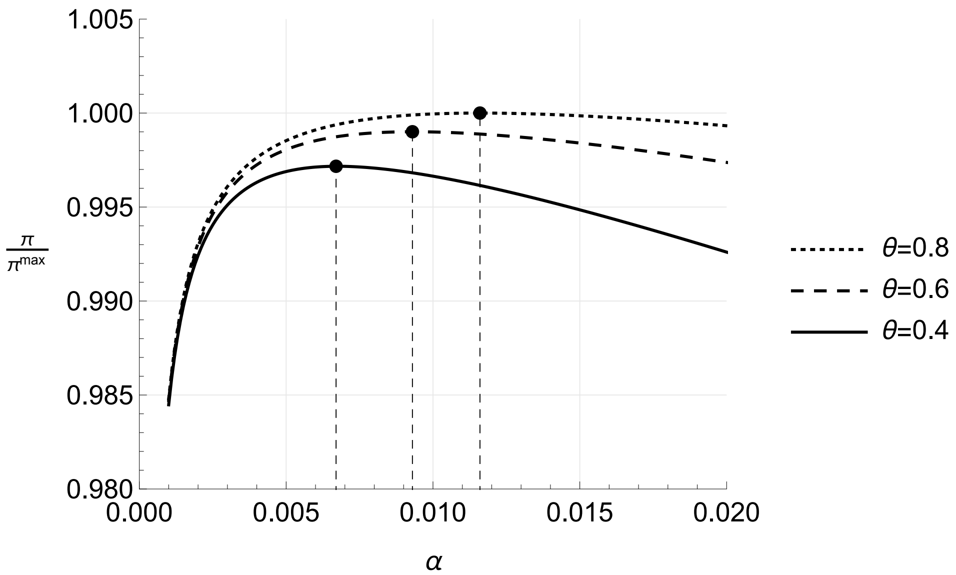

Figure 7.8: Normalized profits for varying α and different service restoration rates θ (t=0.01, m=0.1, c=1, b_Q=0.1, b_0=1, Λ=0.9)

7.4.2 Effect of Store Profitability and Theft Rate

Figure 7.9 illustrates the profit-maximizing EAS configuration α^* in response to varying levels of the store's average profit margin m for different store traffic intensity. It can be seen that α^* is ceteris paribus decreasing in product profitability. This is because as margins and store traffic increase, avoidance of congestion becomes more important (at the expense of more undetected theft). Consequently, the optimal configuration is determined by a characteristic of the store (i.e., profitability). However, this needs to be reflected upon as higher store margins will most likely also increase the cost of theft.

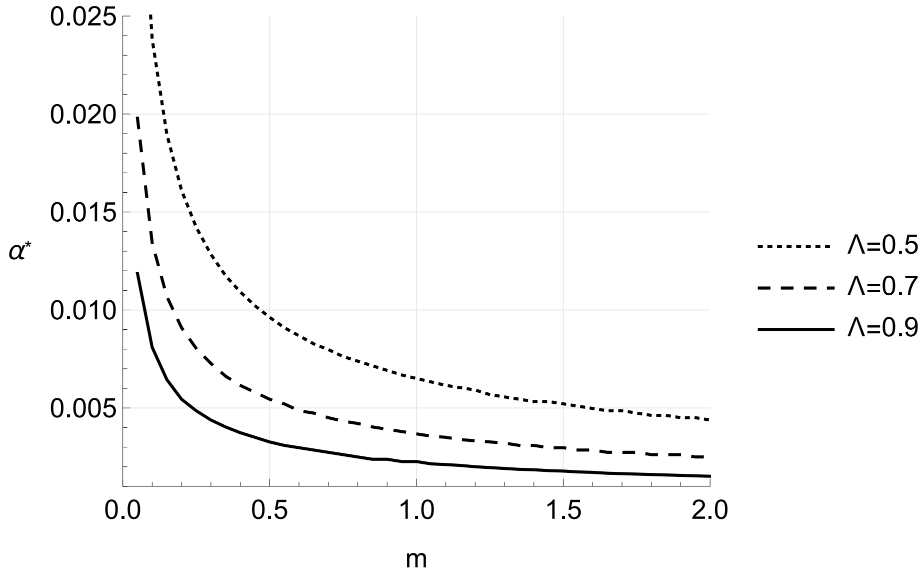

Figure 7.9: Optimal false-positive rate α^* for varying profit margins m (t=0.01, c=1, b_Q=0.1, b_0=1, θ=0.5)

Figure 7.10 illustrates the profit-maximizing EAS configuration α^* in response to varying theft rates t for different store traffic intensity. The false-positive rate α^* is unsurprisingly increasing in the theft rate as the store improves theft protection when facing more frequent store theft. This comes at the expense of more frequent server breakdowns. For low theft rates, the number of customers in the store has little effect on the optimal false-positive rate. With an increasing theft rate, however, the customer stream has a significant impact on the optimal system configuration. If the store's utilization level is low, the store queue is obviously relatively short. In contrast, with increasing utilization, the queue length increases, which impairs purchase probability due to the non-linear relationship between customer stream and store revenue. The higher the store utilization, the more important is it to avoid false alarms which will further stress service quality due to server breakdowns.

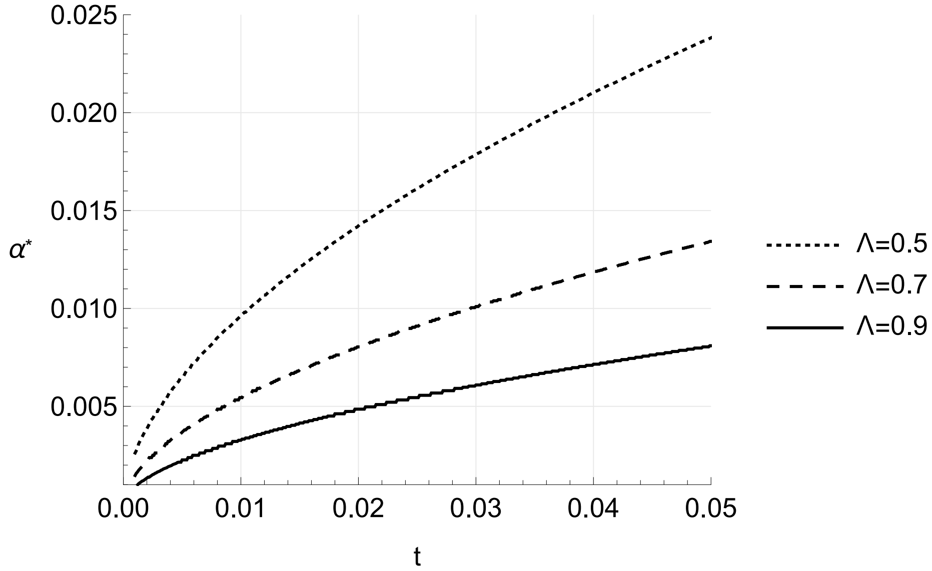

Figure 7.10: Optimal false-positive rate α^* for varying theft rates t (m=0.5, c=1, b_Q=0.1, b_0=1, θ=0.5)

7.4.3 Effect of Competitive Pressure

Besides store properties, we also consider the effect of the underlying customer choice model to understand the effects of false alarms on store profitability. Figure 7.11 illustrates the influence of the base conversion rate b_0 on the optimal false-positive rate for different levels of customer service quality sensitivity b_Q. It turns out that the optimal false-positive rate α^* is decreasing in b_Q. In contrast, we find a non-monotone relationship between store base utility and α^*. This requires a more careful explanation: In the case of a lower base conversion rate, only a small percentage of customers ultimately purchase products. In that case, customers' general purchase intention is so low that service queues will be extremely rare. A case in point is exclusive department stores. Although these stores are often very crowded, only a small fraction of the visitors are potential buyers; most of them just want to look around. In this case, the store will optimally increase security measures because the impact on queue length will be minimal. In contrast, situations with very high base conversion rates reflect circumstances whereby customers want to purchase products despite an extremely poor service offering (e.g., the sales launch of a new iPhone generation). In this situation, the adverse effect of tightened security will certainly increase queues, but customers do not care. For intermediate values of b_0, a store can actively influence purchase decisions through an improved service experience. Consequently, false-positive events matter, and store security is relaxed in favor of decreased service disruptions.

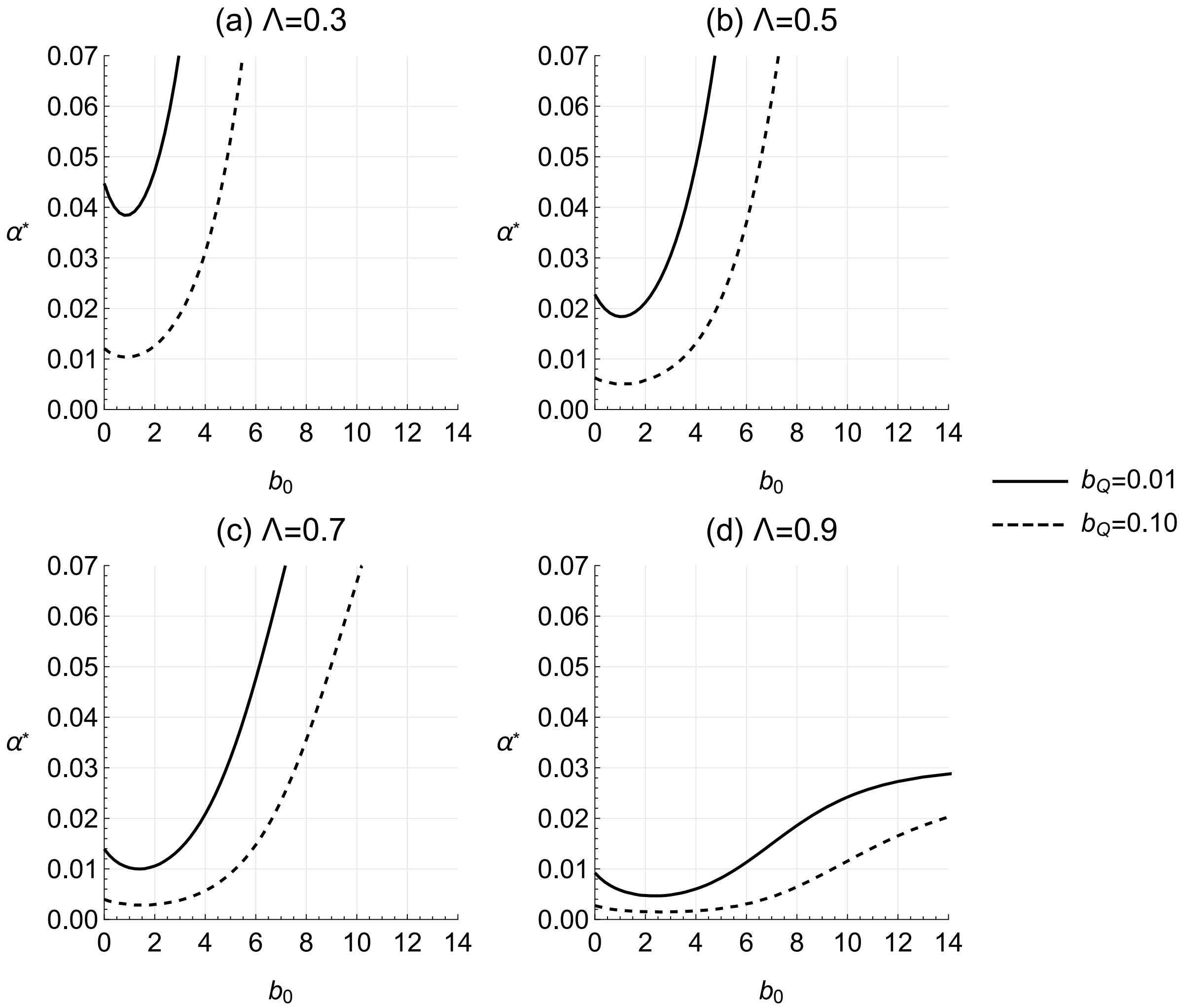

Figure 7.11: Optimal false-positive rate α^* as a function of competitive position b_0 for different values of b_Q and Λ (t=0.01, m=0.5, c=1, θ=0.25)

7.4.4 Adaptive System Configurations

The above relationships suggest a clear advantage from adapting a smart EAS configuration to the current store situation. Of particular importance in this context is the utilization of the service system. This observation corresponds to the results of Mani, Kesavan, and Swaminathan (2015), who argue that dynamic capacity management based on store traffic information is important to the optimization of profits. However, in contrast to short-term staffing decisions, adapting the optimal EAS gate configuration may be achieved rather easily. Figure 7.12 provides an illustration how an integrated service optimization model can inform such adaptive configuration decisions: Depending on current store traffic Λ (as measured by the number of entries), profitability m (presence of costly store promotions), and competitive pressure b_0 (presence of competitor promotions) the model determines different optimal false-positive rates α^*.

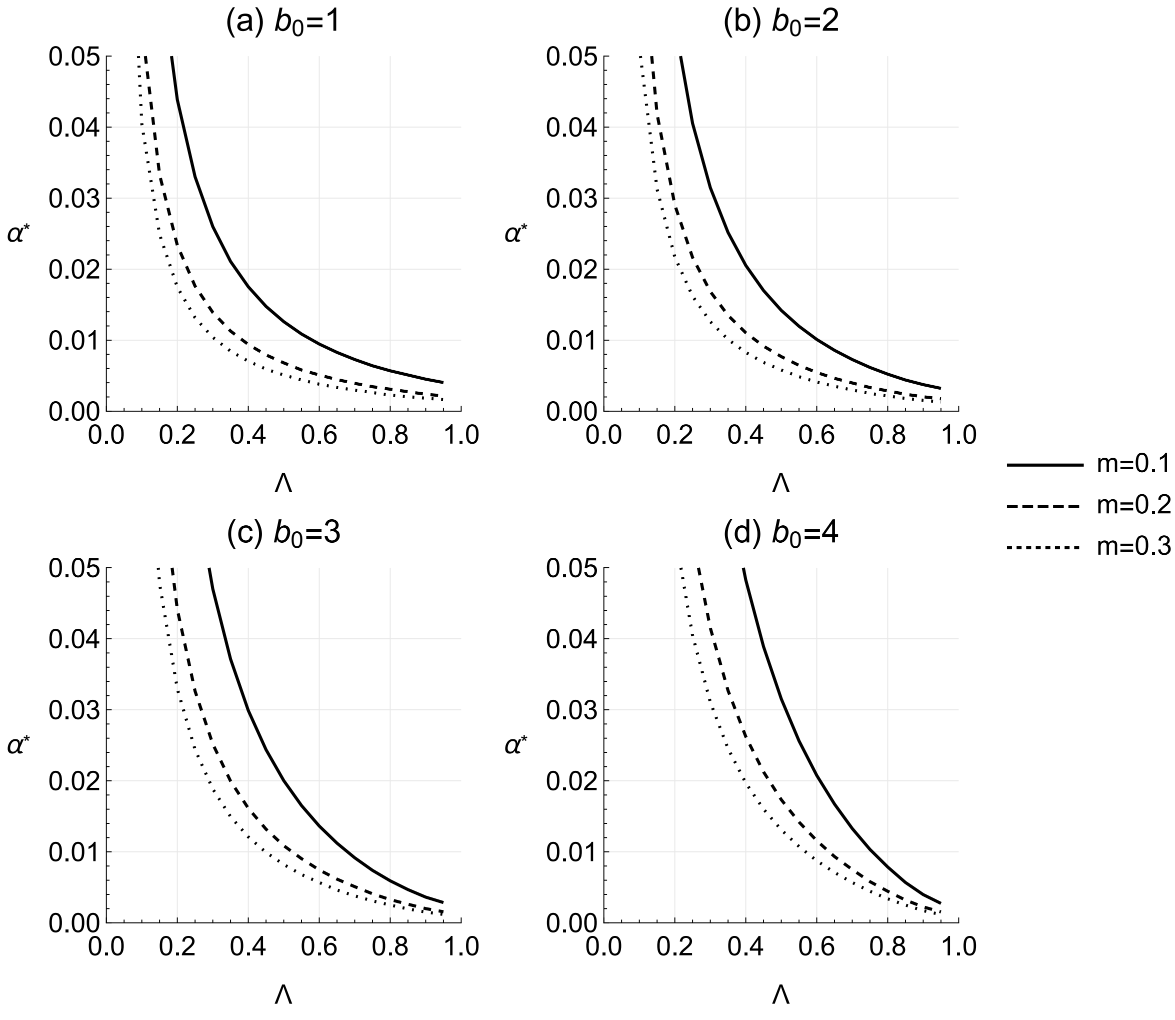

Figure 7.12: Optimal false-positive rate α^* as a function of current store traffic Λ for different values of m and b_0 (t=0.01, c=1, b_Q=0.1, θ=0.25)

7.5 Discussion

While the previous three chapters (as well as prior research on RFID data analytics) assessed the performance of developed machine learning models using standard performance metrics for predictive power in terms of accuracy, the present chapter introduced a service operations model that allows for the evaluation of the business performance of such models embedded into socio-technical systems. In this respect, we followed Lee and Özer (2007) and were able to map technical detection characteristics to business-relevant metrics such as service quality or profitability. Moreover, the service operations model allowed us to determine the optimal configuration of the machine learning models, leveraging the models' freedom of fine-tuning detection sensitivity. We focused on two frequently discussed application cases, namely (i) EAS and (ii) automated checkout. Our results highlight the importance of a comprehensive evaluation of such smart applications: While the EAS application achieved high operational performance with the best classification model, the automated checkout solution failed this initial litmus test (although both applications are based on the same underlying machine learning models for transition detection).

Naturally, there are limitations inherent to the presented research. First, we considered only transition detection, which—despite its crucial importance for many smart applications—is only one facet of a broader range of detection tasks. In particular, applications relying on indoor localization cannot directly benefit from the proposed service operations model. As a matter of fact, the chapter was only concerned with (i) the transition detection models presented in Chapter 4 (i.e., the chapter presenting the EAS artifact) and (ii) two of the three application scenarios discussed in this thesis. In particular, we took into account machine learning models that were only optimized with regard to detection accuracy but not for speed of detection. An extension of the service operations model would, however, easily be possible—in the EAS case, for example, by adjusting the capture rate γ (i.e., by considering a functional correlation between the speed of detection and the probability of successfully catching a thief). Similarly, an adaptation of the service operations model to optimally configure the underlying transition detection model of the smart fitting room application would be possible, as the underlying optimization problem is based on a similar trade-off between misclassification events. A product that is brought into a cabin but not detected by the system (i.e., a false-negative event) results in missed opportunities for cross- and up-selling; a product that is accidentally detected (i.e., a false-positive event) impairs service quality as it causes the system to display misleading information (e.g., additional information on products outside the fitting room).

With regard to the actual model, we focused solely on capacity utilization as the single measure of service quality. This is arguably a somewhat limited focus and fails to account for "broader and non-financial consequences" (Ostrom et al. 2015). Without access to comprehensive store data outside the experimental study, we parametrized the service operations model in a generic fashion. Similarly, the chosen model setup may be too restrictive for proper representation of real-world retail service systems. A richer model could increase the number of servers, relax the distributional properties of the stochastic processes and potentially consider a queuing network with sequential processes (e.g., browsing, trying, buying). For instance, the theft rate may not be exogenously given, but rather dependent on the chosen EAS configuration, that is, shoplifters respond to the number of undetected thefts (Becker 1968; Gill 2007). Furthermore, the integration of empirical store traffic data akin to Kesavan, Deshpande, and Lee (2014), as well as information on margins and competitive pressure, would facilitate the calibration of model parameters. Future research could also move away from using the model primarily for system evaluation and instead assume an active optimization role. This would, in turn, necessitate a more dynamic modeling approach with non-stationary arrival processes.

8 Conclusion

Digital innovation offers a wide range of opportunities in the fashion retail sector (Manyika et al. 2015; PricewaterhouseCoopers 2015). However, the transformation of abstract concepts into real applications is a non-trivial challenge for which turnkey solutions typically do not exist. Reliable object detection and tracking are key capabilities for the transformation of traditional brick-and-mortar stores into smart stores "that are able to accommodate [customer] needs and wants when desired" (Kourouthanassis and Roussos 2003). Against this backdrop, this thesis presented three novel smart fashion store applications that are tied to the locations of garments on the sales floor and offer clear benefits for retail companies and their customers: (i) an EAS system that reliably distinguishes between theft and non-theft events, (ii) an automated checkout system that detects customers' purchases when they are leaving the store and associates them with individual shopping baskets to automatically initiate payment processes, and (iii) a smart fitting room that detects the items carried into individual cabins and identifies the items customers are currently most interested in. The practical relevance of the EAS system is demonstrated not only by the economic extent of losses due to theft but also by the cost reduction potential of replacing existing proprietary EAS systems. The automated checkout system, on the other hand, promises greater sales due to an improved customer experience and cost savings because less store personnel is needed (Manyika et al. 2015). The smart fitting room application represents great potential to enhance the customer shopping experience through additional offerings such as product recommendations or omnichannel services. This thesis focused on two interrelated challenges in the practical implementation of such applications. The first challenge refers to the dependencies between digital services and events in the real world; the second challenge to the optimal configuration and evaluation of such applications.

The three proposed smart applications leverage (i) RFID technology and (ii) machine learning techniques to address the first of the two interrelated challenges. RFID is chosen primarily because the technology is already widely used in the fashion retail industry for the automatic detection of logistical units in upstream and backroom processes. Consequently, solutions that piggyback on RFID present an opportunity to (i) expand the value potential of the technology and (ii) avoid investments in additional sensors and infrastructure. Machine learning techniques, on the other hand, provide a means to (i) address the problem of limited process control on the sales floor (e.g., unpredictable customer behavior, suboptimal store layouts, lack of space) and (ii) facilitate new processes (e.g., detection of customer interactions with RFID-tagged garments). The application of machine learning models to improve detection accuracy was already introduced in prior research (Buffi et al. 2017; Keller, Thiesse, and Fleisch 2014a; Keller et al. 2012; Ma, Wang, and Wang 2018). However, this thesis proposes novel machine learning features that facilitate the reliable and timely identification of multiple objects moving along uncontrolled paths. In doing so, this thesis goes beyond prior research on RFID data analytics, which has almost exclusively focused on standardized processes in controlled environments (e.g., production or logistics facilities). Moreover, this thesis shows that machine learning techniques can be leveraged to detect items at the very moment (or shortly after) they pass through a transition area and that, to achieve this, the antennas do not need to be located at the transition area.

The three proposed IT artifacts comprise different hardware and software components (see Figure 8.1). The EAS system uses an RFID gate with a gate-mounted RFID system and four far-field antennas, the automated checkout system an RFID gate and a ceiling-mounted RFID system with 52 far-field antenna beams, the smart fitting room application only a ceiling-mounted RFID system. The architecture of all three artifacts includes an item detection software component that reliably distinguishes between items that pass through a transition area and others (e.g., static items near the RFID reader). With increasing requirements, the software component was continuously upgraded. While the design of the EAS system focused on the reliable identification of tagged items leaving the shopping floor area, the design of the automated checkout system also aimed at identifying items at the very moment (or shortly after) they leave the store. In contrast to the design of the first two applications, the smart fitting room is based on an RFID infrastructure with antennas that are not located at the transition areas, an arrangement which required a different approach to enable reliable and timely item detection and easy integration into existing retail environments. While the reliable detection of item transitions is of utmost importance for the developed smart applications, this thesis also showed that more information can be extracted from the low-level RFID data to improve the provided services. The architecture of the automated checkout system also comprises a 'product assignment' component and the smart fitting room application an 'interaction detection' software component. The product assignment component assigns items leaving the store (identified by the item detection component) to individual shopping baskets; the interaction detection component distinguishes between items customers are currently interacting with and others.

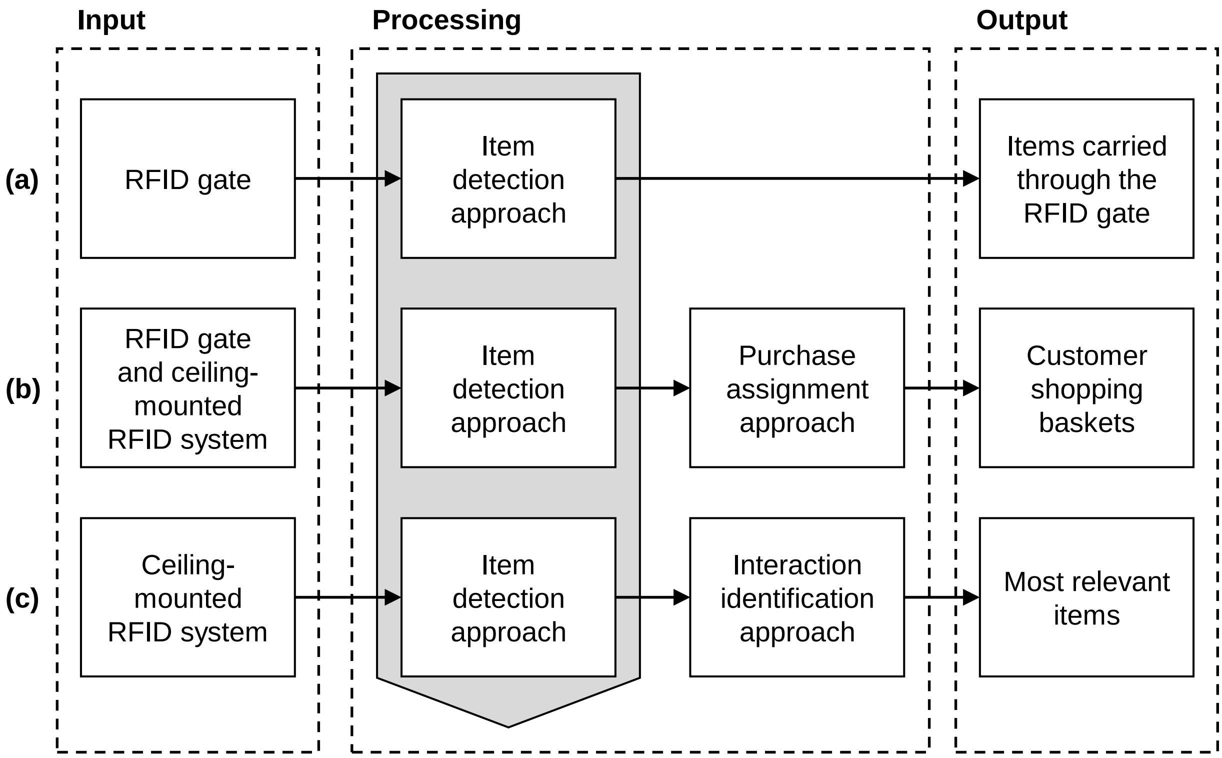

Figure 8.1: Architecture of the (a) EAS system, (b) automated checkout system, and (c) smart fitting room application

While prior research on RFID-based transition detection systems has assessed their performance using standard performance metrics for predictive power in terms of classification accuracy, this thesis demonstrates how the degree to which the systems' underlying classification models help to meet the business objectives can be assessed. The focus on accuracy in prior research neglects (i) the economic impact of misclassifications (which not only depends on the type of error but also on the current state of the service system) and (ii) the inherent trade-off between different false-negative and false-positive classification events (i.e., configuring a classifier for fewer false-positive events typically increases the occurrence of false-negative events and vice versa). Regarding classification models, this also implies that a major degree-of-freedom of these models, the detection sensitivity, is not used. To assess the suitability of the detection results in the context of a service system, this thesis proposed evaluating the business performance of the cyberphysical systems by means of a service operations model. To exemplify the approach, this thesis considered the simplest transition detection classifier developed in the course of the research and two of the discussed application scenarios, namely EAS and automated checkout. To this end, this thesis first introduced a mathematical approximation of the classifier's DET curve, which describes the ratio between false-positive and false-negative events as a function of the model configuration. In a second step, the function was integrated into an optimization model, which determines optimal detection system configurations by internalizing the inherent trade-off between system detection performance and customer service quality. The results highlight the importance of a comprehensive evaluation of smart applications. While the EAS system achieved high operational performance with the best classification system, the automated checkout solution failed the initial litmus test.

Beyond the specific use cases, this thesis provides contributions along different dimensions. First, it demonstrates how data analytics techniques can be leveraged to extract valuable information from RFID data streams. In this context, the data preparation phase has proven particularly important. This thesis shows (i) that one should apply a sliding window approach to enable classifications in a timely fashion and (ii) that features that augment pure signal strength readings with spatial and temporal information help to improve detection accuracy (e.g., temporal shifts between maximum RSSI values of different antennas). Secondly, this thesis shows that RFID data analytics can not only be used to improve existing RFID-based processes but also to facilitate new processes (e.g., fitting rooms that are able to detect garments customers are currently trying on). This is particularly interesting for retailers with RFID in productive use who can leverage the developed software components to provide novel retail services and enhance customers' shopping experience. The software components can thus serve as an incubator for service innovations in the retail industry and beyond. Finally, the thesis demonstrates (i) how technical artifacts embedded in socio-technical systems can be optimally configured and (ii) how they can be economically evaluated. In this context, the use of fitted performance curves establishes a novel link between data mining results and service operations management. This thesis followed Lee and Özer (2007) and was able to substantiate empirical results by means of a thorough analysis of the underlying economic trade-off. This approach allowed us to map technical detection characteristics to business relevant metrics.

There are limitations and opportunities for future research inherent to the presented research that go beyond the discussion sections in the individual chapters. First, this thesis focused on the development of software components that leverage data mining techniques to distinguish between different events based on RFID data. However, successful deployment of the proposed smart fashion store applications necessitates additional system components. Further tests of the automated checkout system should include consideration of the payment initialization process; further tests of the smart fitting room application should comprise customer services that are based on the information provided by the developed software components (e.g., product recommendations based on the garments that are currently of greatest interest to a customer).[1] Secondly, future research should be focused on creating entire service systems instead of designing individual system components. The software components of the discussed artifacts can be applied in instances beyond the presented use cases. Automated detection systems that can be implemented in environments with limited process control, for example, offer various opportunities for additional use cases. Item path information, on the other hand, can be used to trigger automatic stock replenishment or to improve product recommendations as it could help answer help answer various interesting questions (e.g., “Did the customer spend a lot of time in a particular section of the fashion store?”, or “Which items are often tried on together?”). Such generalizations of the developed systems are key for the successful introduction of novel cyberphysical systems. Finally, future research should investigate whether the proposed smart fashion store applications may be perceived as a potential privacy threat by customers.[2] This is important because introductions of new technology in retail environments in the past have shown that a failure to sufficiently consider privacy concerns can have severe consequences for retailers. When retailers in North America and Europe started to roll out RFID in the early 2000s, for example, a public debate arose concerning the potential misuse of the data that could be collected with the technology (Thiesse 2007).

[1]Preliminary results with regard to this issue have been published in Hanke et al. (2018). This paper is concerned with the question of whether and to what extent the sensing capabilities of smart fitting rooms and the integration of contextual information can improve the quality of product recommendations.

[2]Preliminary results with regard to this research question have been published in Weinhard, Hauser, and Thiesse (2017). In this paper, we investigate the antecedents of customers’ usage intention towards smart fitting rooms and the associated trade-off between the perceived benefits and the perceived privacy costs. To this end, we propose a research model based on the most recent version of the Unified Theory of Acceptance and Use of Technology (UTAUT2) and the Extended Privacy Calculus Theory.

Bibliography

Adan, I. and J. Resing (2015). *Queuing Systems*. URL: https://www.win.tue.nl/~iadan/queueing.pdf (visited on 06/21/2019).

Amazon (2018). *Amazon Books*. URL: https://www.amazon.com/amazon-books/b?ie=UTF8&node=13270229011 (visited on 02/05/2018).

Amazon (2018). *Amazon Go: Frequently Asked Questions*. URL: https://www.amazon.com/b?node=16008589011 (visited on 04/15/2016).

Armistead, C. and G. Clark (1994). "The Coping Capacity Management Strategy in Services and the Influence on Quality Performance". In: *International Journal of Service Industry Management* 5.2, pp. 5–22. DOI: 10.1108/09564239410057654.

Bai, Y., F. Wang, and P. Liu (2006). "Efficiently Filtering RFID Data Streams." In: *CleanDB*.

Bamfield, J (2011). *The Global Retail Theft Barometer 2011*. Nottingham, UK: Centre for Retail Research.

Barthel, R., A. Hudson-Smith, and M. de Jode (2014). *Future Retail Environments and the Internet of Things (IoT)*. London Global University, UCL.

Baskerville, R., A. Baiyere, S. Gregor, A. Hevner, and M. Rossi (2018). "Design Science Research Contributions: Finding a Balance between Artifact and Theory". In: *Journal of the Association for Information Systems* 19.5, pp. 358–376. DOI: 10.17705/1jais.00495.

Baxter, G. and I. Sommerville (2011). "Socio-Technical Systems: From Design Methods to Systems Engineering". In: *Interacting with Computers* 23.1, pp. 4–17. DOI: 10.1016/j.intcom.2010.07.003.

Becker, G. S. (1968). "Crime and Punishment: An Economic Approach". In: *The Economic Dimensions of Crime*. Springer, pp. 13–68.

Betzing, J. H., A.-Q. M. Hoang, and J. Becker (2018). "In-Store Technologies in the Retail Servicescape". In: Multikonferenz Wirtschaftsinformatik. Lüneburg.

Beverungen, D., O. Müller, M. Matzner, J. Mendling, and J. vom Brocke (2017). "Conceptualizing Smart Service Systems". In: *Electronic Markets* 29.1, pp. 7–18. DOI: 10.1007/s12525-017-0270-5.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning.* Information Science and Statistics. New York: Springer.

Bitner, M. J., V. A. Zeithaml, and D. D. Gremler (2010). "Technology's Impact on the Gaps Model of Service Quality". In: *Handbook of Service Science*. Springer, pp. 197–218.

Blázquez, M. (2014). "Fashion Shopping in Multichannel Retail: The Role of Technology in Enhancing the Customer Experience". In: *International Journal of Electronic Commerce* 18.4, pp. 97–116. DOI: 10.2753/JEC1086-4415180404.

Borgia, E. (2014). "The Internet of Things Vision: Key Features, Applications and Open Issues". In: *Computer Communications* 54, pp. 1–31. DOI: 10.1016/j.comcom.2014.09.008.

Bottani, E., G. Ferretti, R. Montanari, A. Rizzi, and A. Volpi (2012). "Performances of RFID, Acousto-Magnetic and Radio Frequency Technologies for Electronic Article Surveillance in the Apparel Industry in Europe: A Quantitative Study". In: *International Journal of RF Technologies* 3.2, pp. 137–158. DOI: 10.3233/RFT-2012-024.

Brandt, T., S. Feuerriegel, and D. Neumann (Apr. 2017). "Modeling Interferences in Information Systems Design for Cyberphysical Systems: Insights from a Smart Grid Application". In: *European Journal of Information Systems*, pp. 1–14. DOI: 10.1057/s41303-016-0030-1.

Brunato, M. and R. Battiti (2005). "Statistical Learning Theory for Location Fingerprinting in Wireless LANs". In: *Computer Networks* 47.6, pp. 825–845. DOI: https://doi.org/10.1016/j.comnet.2004.09.004.

Brusey, J., C. Floerkemeier, M. Harrison, and M. Fletcher (2003). "Reasoning about Uncertainty in Location Identification with RFID". In: *Workshop on Reasoning with Uncertainty in Robotics at IJCAI*, pp. 23–30.

Brynjolfsson, E., Y. J. Hu, and M. S. Rahman (2013). "Competing in the Age of Omnichannel Retailing". In: *MIT Sloan Management Review.*

Buffi, A., E. D'Andrea, B. Lazzerini, and P. Nepa (2017). "UHF-RFID Smart Gate: Tag Action Classifier by Artificial Neural Networks". In: *2017 IEEE International Conference on RFID Technology Application*, pp. 45–50. DOI: 10.1109/RFID-TA.2017.8098900.

Böhmann, T., J. M. Leimeister, and K. Möslein (2014). “Service Systems Engineering: A Field for Future Information Systems Research”. In: *Business & Information Systems Engineering* 6.2, pp. 73–79. DOI: 10.1007/s12599-014-0314-8.

— (2018). “The New Frontiers of Service Systems Engineering”. In: *Business & Information Systems Engineering* 60.5, pp. 373–375. DOI: 10.1007/s12599-018-0553-1.

Cao, L. (2014). “Business Model Transformation in Moving to a Cross-Channel Retail Strategy: A Case Study”. In: *International Journal of Electronic Commerce* 18.4, pp. 69–96. DOI: 10.2753/JEC1086-4415180403.

Cao, L. and L. Li (2015). “The Impact of Cross-Channel Integration on Retailers’ Sales Growth”. In: *Journal of Retailing* 91.2, pp. 198–216. DOI: 10.1016/j.jretai.2014.12.005.

Chang, C.-C. and C.-J. Lin (2011). “LIBSVM: A Library for Support Vector Machines”. In: *ACM Transactions on Intelligent Systems and Technology* 2.3, pp. 1–27. DOI: 10.1145/1961189.1961199.

Chapman, P., J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer, and R. Wirth (2000). *CRISP-DM 1.0 Step-by-Step Data Mining Guide*. SPSS Inc.

Chaves, L. W. F., E. Buchmann, and K. Böhm (2010). “Finding Misplaced Items in Retail by Clustering RFID Data”. In: *Proceedings of the 13th International Conference on Extending Database Technology*. New York, NY, USA: ACM, pp. 501–512. DOI: 10.1145/1739041.1739102.

Chen, T. and C. Guestrin (2016). “XGBoost: A Scalable Tree Boosting System”. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, pp. 785–794.

Davis, M. M. and T. E. Vollmann (1990). “A Framework for Relating Waiting Time and Customer Satisfaction in a Service Operation”. In: *Journal of Services Marketing* 4.1, pp. 61–69. DOI: 10.1108/EUM0000000002506.

Dawson, S. (1993). “Consumer Responses to Electronic Article Surveillance Alarms”. In: *Journal of Retailing* 69.3, pp. 353–362. DOI: 10.1016/0022-4359(93)90011-7.

De Vany, A. (1976). “Uncertainty, Waiting Time, and Capacity Utilization: A Stochastic Theory of Product Quality”. In: *Journal of Political Economy* 84.3, pp. 523–541. DOI: 10.1086/260457.

Deyle, E. (2015). *The Global Retail Theft Barometer 2014-2015*.

Domingos, P. (2012). "A Few Useful Things to Know about Machine Learning". In: *Communications of the ACM* 55.10, pp. 78–87. DOI: 10.1145/2347736.2347755.

Donaldson, T. (2015). *More Manufacturers, Retailers Adopt RFID for Better Inventory Visibility*. URL: https://sourcingjournal.com/topics/retail/survey-manufacturers-retailers-adopt-rfid-better-inventory-visibility-26379/ (visited on 09/16/2018).

Elkan, C. (2001). "The Foundations of Cost-Sensitive Learning". In: *Proceedings of the 17th International Joint Conference on Artificial Intelligence*. 1. Lawrence Erlbaum Associates Ltd, pp. 973–978.

Fan, T.-J., X.-Y. Chang, C.-H. Gu, J.-J. Yi, and S. Deng (2014). "Benefits of RFID Technology for Reducing Inventory Shrinkage". In: *International Journal of Production Economics* 147, pp. 659–665. DOI: 10.1016/j.ijpe.2013.05.007.

Fan, W., S. J. Stolfo, J. Zhang, and P. K. Chan (1999). "AdaCost: Misclassification Cost-Sensitive Boosting". In: *Proceedings of the Sixteenth International Conference on Machine Learning*, pp. 97–105.

Fawcett, T. (2006). "An Introduction to ROC Analysis". In: *Pattern Recognition Letters* 27.8, pp. 861–874. DOI: 10.1016/j.patrec.2005.10.010.

Finkenzeller, K. (2015). *RFID-Handbuch: Grundlagen Und Praktische Anwendungen von Transpondern, Kontaktlosen Chipkarten Und NFC*. 7th, revised and expanded. München: Carl Hanser Verlag.

Fishkin, K. P., B. Jiang, M. Philipose, and S. Roy (2004). "I Sense a Disturbance in the Force: Unobtrusive Detection of Interactions with RFID-Tagged Objects". In: *International Conference on Ubiquitous Computing*. Springer, pp. 268–282.

Flath, C. M. and N. Stein (2018). "Towards a data science toolbox for industrial analytics applications". In: *Computers in Industry* 94, pp. 16–25. DOI: 10.1016/j.compind.2017.09.003.

Frost, R. and K. Lyons (2017). "Service Systems Analysis Methods and Components: A Systematic Literature Review". In: *Service Science* 9.3, pp. 219–234. DOI: 10.1287/serv.2017.0180.

Gallino, S. and A. Moreno (2014). "Integration of Online and Offline Channels in Retail: The Impact of Sharing Reliable Inventory Availability Information". In: *Management Science* 60.6, pp. 1434–1451. DOI: 10.1287/mnsc.2014.1951.

Gill, M. (2007). *Shoplifters on Shop Theft: Implications for Retailers.* Leicester, UK: Perpetuity Research & Consultancy International.

Goller, M. and M. Brandner (2011a). "Experimental Evaluation of RFID Gate Concepts". In: *2011 IEEE International Conference on RFID*. IEEE, pp. 26–31.

— (2011b). "Improving Classification Performance of RFID Gates Using Hidden Markov Models". In: *2011 IEEE International Instrumentation and Measurement Technology Conference*. IEEE, pp. 1–5.

— (2012). "Evaluation of Feature Attributes for an RFID Conveyor Belt Application Using Probabilistic Models". In: *European Conference on Smart Objects, Systems and Technologies*. VDE, pp. 1–7.

Gregor, S. and A. R. Hevner (2013). "Positioning and Presenting Design Science Research for Maximum Impact." In: *MIS Quarterly* 37.2, pp. 337–355. DOI: 10.25300/MISQ/2013/37.2.01.

Gregory, J. (2015). *The Internet of Things, Revolutionizing the Retail Industry*. Accenture.

Grewal, D., A. L. Roggeveen, and J. Nordfält (2017). "The Future of Retailing". In: *Journal of Retailing* 93.1, pp. 1–6. DOI: 10.1016/j.jretai.2016.12.008.

Grewal, D., J. Baker, M. Levy, and G. B. Voss (2003). "The Effects of Wait Expectations and Store Atmosphere Evaluations on Patronage Intentions in Service-Intensive Retail Stores". In: *Journal of Retailing* 79.4, pp. 259–268. DOI: 10.1016/j.jretai.2003.09.006.

Grey, D. and B. J. Morgan (1972). "Some Aspects of ROC Curve-Fitting: Normal and Logistic Models". In: *Journal of Mathematical Psychology* 9.1, pp. 128–139. DOI: 10.1016/0022-2496(72)90009-0.

GS1 Germany (2007). *Supply Chain Management in the European Textile Industry*, pp. 1–39.

Gummesson, E. (1998). *Qualitative Methods in Management Research.* Lund: Studentlitterature Chartwell-Bratt.

Gönen, M. (2006). "Receiver Operating Characteristic (ROC) Curves". In: *Proceedings of the Thirty-First Annual SAS Users Group International Conference*, pp. 210–231.

Halevy, A., P. Norvig, and F. Pereira (2009). "The Unreasonable Effectiveness of Data". In: *IEEE Intelligent Systems* 24.2, pp. 8–12. DOI: 10.1109/MIS.2009.36.

Hanke, J., M. Hauser, A. Dürr, and F. Thiesse (2018). "Redefining the Offline Retail Experience: Designing Product Recommendation Systems for Fashion Stores". In: *Proceedings of the Twenty-Sixth European Conference on Information Systems*, pp. 1–14.

Hardgrave, B., J. Aloysius, and S. Goyal (2013). “RFID-Enabled Visibility and Retail Inventory Record Inaccuracy: Experiments in the Field”. In: *Production and Operations Management* 22.4, pp. 843–856. DOI: 10.1111/poms.12010.

Hauser, M., C. M. Flath, and F. Thiesse (2019). “Catch Me If You Scan: Enabling Smart Retailing Through RFID-Based Item Detection”.

Hauser, M., M. Griebel, and F. Thiesse (2017). “A Hidden Markov Model for Distinguishing between RFID-Tagged Objects in Adjacent Areas”. In: *IEEE International Conference on RFID*. IEEE. DOI: 10.1109/RFID.2017.7945604.

Hauser, M., D. Zügner, C. M. Flath, and F. Thiesse (2015). “Pushing the Limits of RFID: Empowering RFID-Based Electronic Article Surveillance with Data Analytics Techniques”. In: *Proceedings of the Thirty Sixth International Conference on Information Systems*, pp. 1–9.

Hauser, M., S. A. Günther, C. M. Flath, and F. Thiesse (2017a). “Designing Pervasive Information Systems: A Fashion Retail Case Study”. In: *Proceedings of the Thirty Eighth International Conference on Information Systems*, pp. 1–16.

Hauser, M., M. Griebel, J. Hanke, and F. Thiesse (2017b). “Empowering Smarter Fitting Rooms with RFID Data Analytics”. In: *Proceedings of the 13th International Conference on Wirtschaftsinformatik*, pp. 1299–1302.

Hauser, M., S. A. Günther, C. M. Flath, and F. Thiesse (2019). “Towards Digital Transformation in Fashion Retailing: A Design-Oriented IS Research Study of Automated Checkout Systems”. In: *Business & Information Systems Engineering* 61.1, pp. 51–66. DOI: 10.1007/s12599-018-0566-9.

Hayes, R. and R. Blackwood (2006). “Evaluating the Effects of EAS on Product Sales and Loss: Results of a Large-Scale Field Experiment”. In: *Security Journal* 19.4, pp. 262–276. DOI: 10.1057/palgrave.sj.8350025.

Herhausen, D., J. Binder, M. Schoegel, and A. Herrmann (2015). “Integrating Bricks with Clicks: Retailer-Level and Channel-Level Outcomes of Online–Offline Channel Integration”. In: *Journal of Retailing* 91.2, pp. 309–325. DOI: 10.1016/j.jretai.2014.12.009.

Herzer, G. (2003). “Magnetic Materials for Electronic Article Surveillance”. In: *Journal of Magnetism and Magnetic Materials* 254, pp. 598–602. DOI: 10.1016/S0304-8853(02)00930-7.

Hevner, A. R., S. T. March, J. Park, and S. Ram (2004). "Design Science in Information Systems Research". In: *MIS Quarterly* 28.1, pp. 75–105. DOI: 10.2307/25148625.

Ho, T. H. and Y.-S. Zheng (2004). "Setting Customer Expectation in Service Delivery: An Integrated Marketing-Operations Perspective". In: *Management Science* 50.4, pp. 479–488. DOI: 10.1287/mnsc.1040.0170.

Hoang, M. K., S. Schmitz, C. Drueke, D. H. T. Vu, J. Schmalenstroeer, and R. Haeb-Umbach (2013). "Server Based Indoor Navigation Using RSSI and Inertial Sensor Information". In: *10th Workshop on Positioning Navigation and Communication*. IEEE, pp. 1–6.

Impinj Inc. (2017a). *SpeedwayR Installation and Operations Guide Version 5.12.0*.

— (2017b). *xArray/xSpan Installation and Operations Guide Version 5.12.0*.

Ingilizian, Z., M. Ghosh, S. Shellaby, C. Donnelly, O. Wright, K. Fang Grant, A. Shah, H. Jolly, and S. Shroff (2017). *Shaping the Future of Retail for Consumer Industries*. World Economic Forum.

Jeffery, S. R., M. Garofalakis, and M. J. Franklin (2006). "Adaptive Cleaning for RFID Data Streams". In: *Proceedings of the 32nd International Conference on Very Large Data Bases*, pp. 163–174.

Ju Tu, Y. and S. Piramuthu (2008). "Reducing False Reads in RFID-Embedded Supply Chains". In: *Journal of Theoretical and Applied Electronic Commerce Research* 3.2, pp. 60–70. DOI: 10.4067/S0718-18762008000100006.

Kalish, I. and V. Eng (2018). *Global Powers of Retailing 2018 - Transformative Change, Reinvigorated Commerce*. Deloitte Touche Tohmatsu Limited.

Kang, Y. and S. B. Gershwin (2005). "Information Inaccuracy in Inventory Systems: Stock Loss and Stockout". In: *IIE Transactions* 37.9, pp. 843–859. DOI: 10.1080/07408170590969861.

Kasiri, N., R. Sharda, and B. Hardgrave (2012). "A Balanced Scorecard for Item-Level RFID in the Retail Sector: A Delphi Study". In: *European Journal of Information Systems* 21.3, pp. 255–267. DOI: 10.1057/ejis.2011.33.

Keller, T., F. Thiesse, and E. Fleisch (2014a). "Classification Models for RFID-Based Real-Time Detection of Process Events in the Supply Chain: An Empirical Study." In: *ACM Transactions on Management Information Systems* 5.4, pp. 1–30. DOI: 10.1145/2629449.

— (2014b). "Using Dynamic Time Warping to Identify RFID Tag Movement in a Logistics Scenario with and without Additional Process Knowledge". In: *Proceedings of the Thirty-Fifth International Conference on Information Systems*, pp. 1–16.

Keller, T., F. Thiesse, J. Kungl, and E. Fleisch (Nov. 2010). "Using Low-Level Reader Data to Detect False-Positive RFID Tag Reads". In: *IEEE International Conference on the Internet of Things*. IEEE, pp. 1–8. DOI: 10.1109/IOT.2010.5678439.

Keller, T., F. Thiesse, A. Ilic, and E. Fleisch (2012). "Decreasing False-Positive RFID Tag Reads by Improved Portal Antenna Setups". In: *IEEE International Conference on the Internet of Things*. IEEE, pp. 99–106. DOI: 10.1109/IOT.2012.6402310.

Kern, C. (2007). *Anwendung von RFID-Systemen*. Springer.

Kesavan, S., V. Deshpande, and H. S. Lee (2014). "Increasing Sales by Managing Congestion in Self-Service Environments: Evidence from a Field Experiment".

Kourouthanassis, P. and G. Roussos (Apr. 2003). "Developing Consumer-Friendly Pervasive Retail Systems". In: *IEEE Pervasive Computing* 2.2, pp. 32–39. DOI: 10.1109/MPRV.2003.1203751.

Krishnamoorthy, A, P. Pramod, and S. Chakravarthy (2014). "Queues with Interruptions: A Survey". In: *Top* 22.1, pp. 290–320. DOI: 10.1007/s11750-012-0256-6.

Kumar, D., E. M. Kornfield, A. C. Prater, S. Boyapati, X. Ren, and C. Yuan (Jan. 15, 2015). "Detecting Item Interaction and Movement". U.S. pat. 20150019391A1.

Kumar, P., H. Reinitz, J. Simunovic, K. Sandeep, and P. D. Franzon (2009). "Overview of RFID Technology and Its Applications in the Food Industry". In: *Journal of Food Science* 74.8, pp. 101–106. DOI: 10.1111/j.1750-3841.2009.01323.x.

Kwon, W.-S. and S. J. Lennon (2009). "Reciprocal Effects between Multichannel Retailers' Offline and Online Brand Images". In: *Journal of Retailing* 85.3, pp. 376–390. DOI: 10.1016/j.jretai.2009.05.011.

Lampe, M., C. Flörkemeier, and S. Haller (2005). "{Einführung in Die RFID-Technologie}". In: *Das Internet Der Dinge*. Springer, pp. 69–86.

Lasi, H., P. Fettke, H.-G. Kemper, T. Feld, and M. Hoffmann (2014). "Industry 4.0". In: *Business & Information Systems Engineering* 6.4, pp. 239–242. DOI: 10.1007/s12599-014-0334-4.

Lee, E. A. (2008). "Cyber Physical Systems: Design Challenges". In: *IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing*. IEEE, pp. 363–369.

Lee, H. and Özer (2007). "Unlocking the Value of RFID". In: *Production and Operations Management* 16.1, pp. 40–64. DOI: 10.1111/j.1937-5956.2007.tb00165.x.

Lee, I. and O. Sokolsky (2010). "Medical Cyber Physical Systems". In: *Design Automation Conference*. IEEE, pp. 743–748.

Lee, J., B. Bagheri, and H.-A. Kao (2015). "A Cyber-Physical Systems Architecture for Industry 4.0-Based Manufacturing Systems". In: *Manufacturing Letters* 3, pp. 18–23. DOI: 10.1016/j.mfglet.2014.12.001.

Levy, Y. and T. J. Ellis (2006). "A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research." In: *Informing Science* 9, pp. 181–212. DOI: 10.28945/479.

Li, H., C. Ye, and A. P. Sample (2015). "IDSense: A Human Object Interaction Detection System Based on Passive UHF RFID". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, pp. 2555–2564.

Litfin, T. and G. Wolfram (2006). "New Automated Checkout Systems". In: *Retailing in the 21st Century: Current and Future Trends*. Ed. by M. Krafft and M. K. Mantrala. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 143–157.

Liu, H., H. Darabi, P. Banerjee, and J. Liu (2007). "Survey of Wireless Indoor Positioning Techniques and Systems". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37.6, pp. 1067–1080. DOI: 10.1109/TSMCC.2007.905750.

Lu, Y., A. Musalem, M. Olivares, and A. Schilkrut (2013). "Measuring the Effect of Queues on Customer Purchases". In: *Management Science* 59.8, pp. 1743–1763. DOI: 10.1287/mnsc.1120.1686.

Ma, H., Y. Wang, and K. Wang (2018). "Automatic Detection of False Positive RFID Readings Using Machine Learning Algorithms". In: *Expert Systems with Applications* 91, pp. 442–451. DOI: 10.1016/j.eswa.2017.09.021.

Maglio, P. P., S. Vargo, N. Caswell, and J. Spohrer (2009). "The Service System Is the Basic Abstraction of Service Science". In: *Information Systems and e-business Management* 7.4, pp. 395–406. DOI: 10.1007/s10257-008-0105-1.

Mani, V., S. Kesavan, and J. M. Swaminathan (2015). "Estimating the Impact of Understaffing on Sales and Profitability in Retail Stores". In: *Production and Operations Management* 24.2, pp. 201–218. DOI: 10.1111/poms.12237.

Manyika, J., M. Chui, P. Bisson, J. Woetzel, R. Dobbs, J. Bughin, and D. Aharon (2015). *The Internet of Things: Mapping the Value Beyond the Hype*. McKinsey Global Institute.

Martin, A., G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki (1997). *The DET Curve in Assessment of Detection Task Performance*. DTIC Document.

Martin, D., R. Hirt, and N. Kühl (2019). "Service Systems, Smart Service Systems and Cyber-Physical Systems—What's the Difference? Towards a Unified Terminology". In: *Proceedings of the 14th International Conference on Wirtschaftsinformatik*, pp. 17–31.

Massawe, L. V., J. D. Kinyua, and H. Vermaak (2012). "Reducing False Negative Reads in RFID Data Streams Using an Adaptive Sliding-Window Approach". In: *Sensors* 12.4, pp. 4187–4212. DOI: 10.3390/s120404187.

Medina-Borja, A. (2015). "Smart Things as Service Providers: A Call for Convergence of Disciplines to Build a Research Agenda for the Service Systems of the Future". In: *Service Science* 7.1, pp. ii–v. DOI: 10.1287/serv.2014.0090.

Melià-Segui, J., R. Pous, A. Carreras, M. Morenza-Cinos, R. Parada, Z. Liaghat, and R. De Porrata-Doria (2013). "Enhancing the Shopping Experience through RFID in an Actual Retail Store". In: *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*. ACM, pp. 1029–1036.

Menard, S. (2018). *Applied Logistic Regression Analysis*. Vol. 106. SAGE publications.

Messinger, P. R. and C. Narasimhan (1997). "A Model of Retail Formats Based on Consumers' Economizing on Shopping Time". In: *Marketing Science* 16.1, pp. 1–23. DOI: 10.1287/mksc.16.1.1.

Meuter, M. L., A. L. Ostrom, R. I. Roundtree, and M. J. Bitner (2000). "Self-Service Technologies: Understanding Customer Satisfaction with Technology-Based Service Encounters". In: *Journal of Marketing* 64.3, pp. 50–64. DOI: 10.1509/jmkg.64.3.50.18024.

Mooney, J. G., V. Gurbaxani, and K. L. Kraemer (1996). "A Process Oriented Framework for Assessing the Business Value of Information Technology". In: *ACM SIGMIS Database* 27.2, pp. 68–81.

Morrell, L. (2017). *Getting to the Roots of What Digital Innovation Means. Internet Retailing*. URL: http://internetretailing.net/issue/digital-innovation-report-october-2015/getting-to-the-roots-of-what-digital-innovation-means/ (visited on 11/14/2018).

Motamedi, A., M. M. Soltani, and A. Hammad (2013). "Localization of RFID-Equipped Assets during the Operation Phase of Facilities". In: *Advanced Engineering Informatics* 27.4, pp. 566–579. DOI: 10.1016/j.aei.2013.07.001.

National Science Foundation (2014). *Partnerships for Innovation: Building Innovation Capacity (PFI:BIC)*. Program Solicitation NSF14-610. Arlingtion, Virginia: National Science Foundation.

Nedap Retail (2018). *iD TOP*. URL: `http://www.nedap-retail.com/solutions/idtop/` (visited on 03/27/2018).

Nordic ID (2018). *Nordic ID AR Series*. URL: `https://www.nordicid.com/en/downloads/products/nordic-id-ar-series/` (visited on 03/27/2018).

Oi, W. Y. (1992). "Productivity in the Distributive Trades: The Shopper and the Economies of Massed Reserves". In: *Output Measurement in the Service Sectors*. University of Chicago Press, pp. 161–193.

Oliva, R. and J. D. Sterman (2001). "Cutting Corners and Working Overtime: Quality Erosion in the Service Industry". In: *Management Science* 47.7, pp. 894–914. DOI: `10.1287/mnsc.47.7.894.9807`.

Orel, F. D. and A. Kara (2014). "Supermarket Self-Checkout Service Quality, Customer Satisfaction, and Loyalty: Empirical Evidence from an Emerging Market". In: *Journal of Retailing and Consumer Services* 21.2, pp. 118 –129. DOI: `10.1016/j.jretconser.2013.07.002`.

Österle, H., J. Becker, U. Frank, T. Hess, D. Karagiannis, H. Krcmar, P. Loos, P. Mertens, A. Oberweis, and E. J. Sinz (2011). "Memorandum on Design-Oriented Information Systems Research". In: *European Journal of Information Systems* 20.1, pp. 7–10. DOI: `10.1057/ejis.2010.55`.

Ostrom, A. L., M. J. Bitner, S. W. Brown, K. A. Burkhard, M. Goul, V. Smith-Daniels, H. Demirkan, and E. Rabinovich (2010). "Moving Forward and Making a Difference: Research Priorities for the Science of Service". In: *Journal of Service Research* 13.1, pp. 4–36. DOI: `10.1177/1094670509357611`.

Ostrom, A. L., A. Parasuraman, D. E. Bowen, L. Patricio, and C. A. Voss (2015). "Service Research Priorities in a Rapidly Changing Context". In: *Journal of Service Research* 18.2, pp. 127–159. DOI: `10.1177/1094670515576315`.

Papapostolou, A. and H. Chaouchi (2011). "RFID-Assisted Indoor Localization and the Impact of Interference on Its Performance". In: *Journal of Network and Computer Applications* 34.3, pp. 902–913. DOI: `10.1016/j.jnca.2010.04.009`.

Parada, R., J. Melià-Seguí, M. Morenza-Cinos, A. Carreras, and R. Pous (2015). "Using RFID to Detect Interactions in Ambient Assisted Living Environments". In: *IEEE Intelligent Systems* 30.4, pp. 16–22. DOI: 10.1109/MIS.2015.43.

Parasuraman, A. (2010). "Service Productivity, Quality and Innovation: Implications for Service-Design Practice and Research". In: *International Journal of Quality and Service Sciences* 2.3, pp. 277–286. DOI: 10.1108/17566691011090026.

Parlak, S. and I. Marsic (2013). "Detecting Object Motion Using Passive RFID: A Trauma Resuscitation Case Study". In: *IEEE Transactions on Instrumentation and Measurement* 62.9, pp. 2430–2437. DOI: 10.1109/TIM.2013.2258772.

Pazzani, M., C. Merz, P. Murphy, K. Ali, T. Hume, and C. Brunk (1994). "Reducing Misclassification Costs". In: *Machine Learning Proceedings*, pp. 217–225. DOI: 10.1016/B978-1-55860-335-6.50034-9.

Pearce, J. and S. Ferrier (2000). "Evaluating the Predictive Performance of Habitat Models Developed Using Logistic Regression". In: *Ecological Modelling* 133.3, pp. 225–245. DOI: 10.1016/S0304-3800(00)00322-7.

Peffers, K., T. Tuunanen, and B. Niehaves (2018). "Design Science Research Genres: Introduction to the Special Issue on Exemplars and Criteria for Applicable Design Science Research". In: *European Journal of Information Systems* 27.2, pp. 129–139. DOI: 10.1080/0960085X.2018.1458066.

Peffers, K., T. Tuunanen, M. A. Rothenberger, and S. Chatterjee (2007). "A Design Science Research Methodology for Information Systems Research". In: *Journal of Management Information Systems* 24.3, pp. 45–77. DOI: 10.2753/MIS0742-1222240302.

Perera, C., A. Zaslavsky, P. Christen, and D. Georgakopoulos (2014). "Context Aware Computing for the Internet of Things: A Survey". In: *IEEE Communications Surveys & Tutorials* 16.1, pp. 414–454. DOI: 10.1109/SURV.2013.042313.00197.

Peters, C., P. P. Maglio, R. Badinelli, R. R. Harmon, R. Maull, J. C. Spohrer, T. Tuunanen, S. L. Vargo, J. J. Welser, H. Demirkan, et al. (2016). "Emerging Digital Frontiers for Service Innovation". In: *Communications of the Association for Information Systems* 39, pp. 136–149. DOI: 10.17705/1CAIS.03908.

Piotrowicz, W. and R. Cuthbertson (2014). "Introduction to the Special Issue Information Technology in Retail: Toward Omnichannel Retailing". In: *International Journal of Electronic Commerce* 18.4, pp. 5–16. DOI: 10.2753/JEC1086-4415180400.

Porter, M. E. and J. E. Heppelmann (2014). "How Smart, Connected Products Are Transforming Competition". In: *Harvard Business Review* 92.11, pp. 64–88.

PricewaterhouseCoopers (2015). *Total Retail 2015: Retailers and the Age of Disruption*. URL: https://www.pwc.com/sg/en/publications/assets/total-retail-2015.pdf (visited on 08/02/2019).

Puerini, G. L., D. Kumar, and S. Kessel (Aug. 1, 2015). "Transitioning Items from a Materials Handling Facility". U.S. pat. 2015/0012396 A1. US Patent App. 14/495,818.

Reynolds, A. P., G. Richards, B. de la Iglesia, and V. J. Rayward-Smith (Mar. 2006). "Clustering Rules: A Comparison of Partitioning and Hierarchical Clustering Algorithms". In: *Journal of Mathematical Modelling and Algorithms* 5.4, pp. 475–504. DOI: 10.1007/s10852-005-9022-1.

Rigby, D. (2011). "The Future of Shopping". In: *Harvard Business Review* 89.12, pp. 65–76.

Roberti, M. (2016). *Signs RFID Is Nearing A Tipping Point in Apparel Retail*. URL: https://www.rfidjournal.com/articles/view?15152 (visited on 06/22/2019).

Rousseeuw, P. J. (1987). "Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis". In: *Journal of Computational and Applied Mathematics* 20, pp. 53–65. DOI: 10.1016/0377-0427(87)90125-7.

Roussos, G., P. Kourouthanasis, D. Spinellis, E. Gryazin, M. Pryzbliski, G. Kalpogiannis, and G. Giaglis (2003). "Systems Architecture for Pervasive Retail". In: *Proceedings of the 2003 ACM Symposium on Applied Computing*. ACM, pp. 631–636. DOI: 10.1145/952532.952656.

Ruiz-Garcia, L. and L. Lunadei (2011). "The Role of RFID in Agriculture: Applications, Limitations and Challenges". In: *Computers and Electronics in Agriculture* 79.1, pp. 42–50. DOI: 10.1016/j.compag.2011.08.010.

Rust, R. T. and M.-H. Huang (2014). "The Service Revolution and the Transformation of Marketing Science". In: *Marketing Science* 33.2, pp. 206–221. DOI: 10.1287/mksc.2013.0836.

Sein, M. K., O. Henfridsson, S. Purao, M. Rossi, and R. Lindgren (2011). "Action Design Research". In: *MIS Quarterly* 35.1, pp. 37–56. DOI: 10.2307/23043488.

Senecal, S. and J. Nantel (2004). "The Influence of Online Product Recommendations on Consumers' Online Choices". In: *Journal of Retailing* 80.2, pp. 159–169. DOI: 10.1016/j.jretai.2004.04.001.

Shankar, V., J. J. Inman, M. Mantrala, E. Kelley, and R. Rizley (2011). "Innovations in Shopper Marketing: Current Insights and Future Research Issues". In: *Journal of Retailing* 87.1, pp. 29–42. DOI: 10.1016/j.jretai.2011.04.007.

Sjøbakk, B., A. D. Landmark, and H. P. Hübert (2016). "The Adaptive Fitting Room". In: *Workshop on Business Models and ICT Technologies for the Fashion Supply Chain*. Springer, pp. 79–88.

Stankovic, J. A. (2014). "Research Directions for the Internet of Things". In: *IEEE Internet of Things Journal* 1.1, pp. 3–9. DOI: 10.1109/JIOT.2014.2312291.

Stein, N. and C. M. Flath (2017). "Applying data science for shop-floor performance prediction". In: *Proceedings of the Twenty-Sixth European Conference on Information Systems*.

Stein, N., J. Meller, and C. M. Flath (2018). "Big data on the shop-floor: sensor-based decision-support for manual processes". In: *Journal of Business Economics* 88.5, pp. 593–616. DOI: 10.1007/s11573-017-0890-4.

The Economist (2016). *The Future of Personal Transport: The Driverless, Car-Sharing Road Ahead*. URL: https://www.economist.com/business/2016/01/09/the-driverless-car-sharing-road-ahead (visited on 05/06/2017).

Thiesse, F. (2007). "RFID, Privacy and the Perception of Risk: A Strategic Framework". In: *The Journal of Strategic Information Systems* 16.2, pp. 214–232. DOI: 10.1016/j.jsis.2007.05.006.

Thiesse, F., J. Al-Kassab, and E. Fleisch (2009). "Understanding the Value of Integrated RFID Systems: A Case Study from Apparel Retail". In: *European Journal of Information Systems* 18.6, pp. 592–614. DOI: 10.1057/ejis.2009.33.

Thiesse, F. and T. Buckel (2015). "A Comparison of RFID-Based Shelf Replenishment Policies in Retail Stores under Suboptimal Read Rates". In: *International Journal of Production Economics* 159, pp. 126–136. DOI: 10.1016/j.ijpe.2014.09.002.

Thiruvengadam, K (1963). "Queuing with Breakdowns". In: *Operations Research* 11.1, pp. 62–71. DOI: 10.1287/opre.11.1.62.

Thottam, I. (2016). *A Trip to Amazon's Brick-and-Mortar Book Store*. URL: https://www.pastemagazine.com/articles/2016/03/a-trip-to-amazons-brick-and-mortar-book-store.html (visited on 03/26/2018).

Train, K. E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Uckelmann, D. and G. Romagnoli (2016). “RF-Based Locating of Mobile Objects”. In: *Proceedings of the 6th International Conference on the Internet of Things*. ACM, pp. 147–154. DOI: 10.1145/2991561.2991565.

Venable, J., J. Pries-Heje, and R. Baskerville (2012). “A Comprehensive Framework for Evaluation in Design Science Research”. In: *International Conference on Design Science Research in Information Systems*. Springer, pp. 423–438.

— (2016). “FEDS: A Framework for Evaluation in Design Science Research”. In: *European Journal of Information Systems* 25.1, pp. 77–89. DOI: 10.1057/ejis.2014.36.

Venkatesh, V., J. A. Aloysius, H. Hoehle, and S. Burton (2017). “Design and Evaluation of Auto-ID Enabled Shopping Assistance Artifacts in Customers’ Mobile Phones: Two Retail Store Laboratory Experiments”. In: *MIS Quarterly* 41.1, pp. 83–113. DOI: 10.25300/MISQ/2017/41.1.05.

Verhoef, P. C., P. Kannan, and J. J. Inman (2015). “From Multi-Channel Retailing to Omni-Channel Retailing: Introduction to the Special Issue on Multi-Channel Retailing”. In: *Journal of Retailing* 91.2, pp. 174–181. DOI: 10.1016/j.jretai.2015.02.005.

Verhoef, P. C., S. A. Neslin, and B. Vroomen (2007). “Multichannel Customer Management: Understanding the Research-Shopper Phenomenon”. In: *International Journal of Research in Marketing* 24.2, pp. 129–148. DOI: 10.1016/j.ijresmar.2006.11.002.

Wahba, P. (2014). *Nordstrom Taps eBay’s Tech to Build Fitting Rooms of the Future*. URL: http://fortune.com/2014/11/25/nordstrom-ebay-fitting-rooms/ (visited on 06/22/2019).

Walter, F. E., S. Battiston, M. Yildirim, and F. Schweitzer (2012). “Moving Recommender Systems from On-Line Commerce to Retail Stores”. In: *Information Systems and e-Business Management* 10.3, pp. 367–393. DOI: 10.1007/s10257-011-0170-8.

Want, R. (2006). “An Introduction to RFID Technology”. In: *IEEE Pervasive Computing* 5.1, pp. 25–33. DOI: 10.1109/MPRV.2006.2.

Weinhard, A., M. Hauser, and F. Thiesse (2017). “Explaining Adoption of Pervasive Retail Systems with a Model Based on UTAUT2 and the Extended Privacy Calculus”. In: *Proceedings of the 21st Pacific Asia Conference on Information Systems*.

Weiser, M. (1999). “The Computer for the 21st Century”. In: *Mobile Computing and Communications Review* 3.3, pp. 1–11. DOI: 10.1145/329124.329126.

Willems, K., A. Smolders, M. Brengman, K. Luyten, and J. Schöning (2017). “The Path-to-Purchase Is Paved with Digital Opportunities: An Inventory of Shopper-Oriented Retail

Technologies”. In: *Technological Forecasting and Social Change* 124, pp. 228–242. DOI: 10.1016/j.techfore.2016.10.066.

Winter, R. (2008). “Design Science Research in Europe”. In: *European Journal of Information Systems* 17.5, pp. 470–475. DOI: 10.1057/ejis.2008.44.

Witten, I. H., E. Frank, M. A. Hall, and C. J. Pal (2016). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann.

Wolff, R. W. (1982). “Poisson Arrivals See Time Averages”. In: *Operations Research* 30.2, pp. 223–231. DOI: 10.1287/opre.30.2.223.

Wong, W.-K., S. Leung, Z. Guo, X. Zeng, and P. Mok (2012). “Intelligent Product Cross-Selling System with Radio Frequency Identification Technology for Retailing”. In: *International Journal of Production Economics* 135.1, pp. 308–319. DOI: 10.1016/j.ijpe.2011.08.005.

Wortmann, F. and K. Flüchter (2015). “Internet of Things - Technology and Value Added”. In: *Business & Information Systems Engineering* 57.3, pp. 221–224. DOI: 10.1007/s12599-015-0383-3.

Yao, L., Q. Z. Sheng, W. Ruan, X. Li, S. Wang, and Z. Yang (2015). “Unobtrusive Posture Recognition via Online Learning of Multi-Dimensional RFID Received Signal Strength”. In: *International Conference on Parallel and Distributed Systems*. IEEE, pp. 116–123.

Yuanfeng, D., Y. Dongkai, Y. Huilin, and X. Chundi (2016). “Flexible Indoor Localization and Tracking System Based on Mobile Phone”. In: *Journal of Network and Computer Applications* 69, pp. 107–116. DOI: 10.1016/j.jnca.2016.02.023.

Zhong, R. Y., X. Xu, E. Klotz, and S. T. Newman (2017). “Intelligent Manufacturing in the Context of Industry 4.0: A Review”. In: *Engineering* 3.5, pp. 616–630. DOI: 10.1016/J.ENG.2017.05.015.

Zhu, X., S. K. Mukhopadhyay, and H. Kurata (2012). “A Review of RFID Technology and Its Managerial Applications in Different Industries”. In: *Journal of Engineering and Technology Management* 29.1, pp. 152–167. DOI: 10.1016/j.jengtecman.2011.09.011.