

RFID-Enabled Business Process Intelligence in Retail Stores: A Case Report

Jasser Al-Kassab^{1,2}, Frederic Thiesse³ and Thomas Buckel⁴

¹ University of St. Gallen, Institute of Technology Management (ITEM-HSG), Dufourstrasse, Switzerland, Jasser.Al-Kassab@unisg.ch

² University of Cambridge, Institute for Manufacturing, Cambridge, United Kingdom, ja454@cam.ac.uk
University of Wuerzburg, Chair of IS Engineering, Wuerzburg, Germany, ³frederic.thiesse@uni-wuerzburg.de,
⁴thomas.buckel@uni-wuerzburg.de

Received 1 December 2012; received in revised form 25 March 2013; accepted 22 April 2013

Abstract

The growing interest in Radio Frequency Identification (RFID) technology in recent years has sparked an intensive debate on the benefits to be expected. With the growth of RFID implementations in size and scope comes a shift away from infrastructural aspects to the question of how to draw value from the large amounts of collected data. However, the necessary procedures for the handling of massive RFID data sets are still an under-researched issue. Against this background, the study presents results from a real-world trial conducted by a large apparel retailer. The objective of the trial was to explore the opportunities for generating novel performance indicators and reports on the reality of store processes and customer behavior on the sales floor. We give an overview of the algorithms used for RFID data processing and the interpretation of the resulting insights from a practitioner's point of view. The case example thus provides an overview of the potential of RFID as a powerful tool for assortment optimization, customer research, store layout design, and other management tasks in retail.

Keywords: RFID, Apparel retail, Store logistics, Data management, Business analytics

1 Introduction

Automatic identification technologies have become omnipresent in retail operations [29], [47]. It was particularly the introduction of barcodes almost 40 years ago that fundamentally changed the industry and laid the foundation for many of today's supply chain management concepts [32]. However, the barcode's success was not assured from the outset. Among the obstacles to implementation in the 1970s were consumer resistance to the elimination of item price marking, high scanner equipment costs, and the manufacturers' initial reluctance to redesign product labels [10], [24]. On the pro side, the economic rationale of barcode implementation relied on only three benefit types: (i) time and cost savings by speeding up the check-out process; (ii) the reduction of checker errors; and (iii) the elimination of paper price labels.

Unfortunately though, it turned out that these hard benefits alone were not sufficient to drive rapid diffusion. However, soft benefits (e.g. improvements from more accurate inventory data, real-time tracking of sales) were to a large extent regarded as speculative because of the required changes in planning processes and the lack of data processing capacities [9], [22]. It was not until the end of the 1980s that companies worldwide recognized the economic potential of the data sets collected by means of the barcode. Barcode data eventually became the enabler of several novel supply chain innovations, such as direct store delivery, continuous replenishment, and vendor-managed inventories. As we know today, the value that can be drawn from the billions of barcode scans every day exceeds the initial expectations by orders of magnitude.

1.1 Motivation

In our times, history seems to repeat itself anew in the emergence of wireless sensors, particularly in the form of RFID technology. Driven by the vision of the *internet of things*, the growing interest in the use of RFID on the part of the retail industry in recent years has sparked an intensive debate in academia and practice on the benefits to be expected [34], [37]. Moreover, in Gartner's Hype-cycle 2009, RFID was among the 5 technologies, which was attributed to gain most value for organizations in the following years [17]. A positive influence on the RFID adoption process is exerted by top management support and forces within supply chain partners. Therefore, Retail giants in Europe and the U.S. have taken many efforts to convince their suppliers of the positive impact of RFID on supply chain performance. Notwithstanding these efforts, the overall adoption rate in retail in the early years was still rather slow [41]. A major reason for this phenomenon may be found in the narrow focus of prior RFID trials on applications, where RFID merely serves to save cost by automating object identification tasks. In contrast to that, the value of the collected data as a source of unprecedented managerial insights into the reality of operations and customer behaviour is still a widely untapped area of research [28], [40].

1.2 Objective and Structure

Against this background, the objective of this paper is to answer the following research question:

How can RFID data be used to provide value-added information in support of decision makers in retail organizations and, thus, enhance existing store management processes?

To answer this question we present algorithms, performance indicators, and visualizations developed in the context of a real-world trial conducted by a large apparel retailer. The apparel retail industry in general is characterized by high seasonality, high volatility, high-impulse purchasing, and complicated distribution and logistics operations [12]. Further management challenges encompass (i) short product lives of fashion products (typically only about three months), (ii) complexity of product management due to a large variety of apparel product characteristics, (iii) product counterfeiting in the supply chain, and (iv) lack of knowledge about the behavior of fashion shoppers, yet considered to be *mysterious and unpredictable* [31]. Multi-optional fashion shoppers are more demanding and their behavior has become increasingly difficult to assess [19]. Furthermore, the ageing of populations in Western countries forces retailers to consider the special needs of elderly people. At the same time, ethnic diversity in Western countries has increased during the last decades. Those demographic changes have contributed to shifts in decision patterns of consumers [36]. In sum, these challenges are a powerful motivator for the use of RFID as a novel source of unprecedented insights into store processes and customer behavior in apparel retail [40].

Against this background, we present and discuss filtering and aggregation mechanisms, which allow for turning raw RFID data sets into interpretable information that addresses the needs of retail managers. While the majority of prior research in the field consider the use of RFID for the identification of logistical units such as containers and pallets along the supply chain, the present case allows to extend the research focus on item-level tracking from the distribution center to the point of sales (POS), using a fine-meshed and integrated RFID data collection infrastructure, thus closing the *data void* which exists between the goods receipt and the POS. Hence, our study setting promises new insights into the value of RFID that go beyond the findings from prior work. Besides cost savings, the retailer expected new services for customers through RFID-based applications and new insights into customer behavior, by

the analyses of RFID data gathered on the sales floor. The results of these analyses are supposed to improve the retailers' store layout management, inventory management, category management, and department store processes. The remainder of this paper is organized as follows. In section 2 we provide an overview of the related literature. Section 3 presents the company background, including the presentation of the technical infrastructure of the trial, its goals and the analyses to be conducted. The data sources are presented in Section 4 while Section 5 describes the data filtering and aggregation mechanisms. Section 6 gives an overview of the conducted data analyses and performance metrics. Section 7 finally concludes with a discussion of our main findings, managerial implications, limitations, and suggestions for further research.

2 Literature Review

There is a large number of related studies dealing with the processing of RFID data. In chapter 2 we provide a short overview of related work as well as the research gap, where our study aims to fill.

2.1 Related Work

The various issues surrounding the processing of RFID data can be roughly classified into five categories. Many authors deal with purely technological issues such as the implementation of specialized RFID middleware components to handle large amounts of raw data collected from distributed RFID readers (e.g., [18], [27]). Another technology-oriented research streams sets the focus on error-correction mechanisms and other forms of RFID data filtering in order to ensure the quality of the collected data (e.g., [3], [43]). Another group of researchers have proposed approaches for the efficient storage of RFID data, query languages, and data structures as well as other concepts related to data storage and management (e.g., [2], [46]). Fourth, some of studies investigate the business value of the collected RFID data in different application settings. Fifth, some examples of prior work exist that consider the opportunities for generating value-added information from RFID data from the perspective of business intelligence and data analytics research. The present study contributes to the latter two research streams. Table 1 provides an overview of prior research.

Table 1: Overview of related work in business value, business intelligence and data analytics of RFID

	Reference	Described Value / Purpose / Central findings
Business Value of RFID	[4]	Accuracy, Cost savings, Efficiency, Security
	[13]	Process efficiency, Cost savings, Increased performance, Service level and quality
	[25]	Cost savings
	[31]	Responsiveness, Relatedness, Refinement
	[33]	Effectiveness, Efficiency, Innovation
	[38]	Timeliness, Currency, Accuracy, Completeness
	[39]	Supply chain efficiency, Innovation capability, Learning to adjust, Learning to transform
	[40]	Process Performance (Management and Operational), Capabilities (Information, Process and Information Technology (IT)) and firm Performance
	[44]	Cost savings, Increasing revenues per customer, Service quality
	[45]	Strategy, Processes, Information flow, IT infrastructure, Human and physical resources
[48]	Efficiency and Effectiveness, Customer satisfaction, Organizational performance	
Business intelligence & data analytics	[2]	An exploration of how enriched RFID data can be prepared in an integrated, multidimensional fashion for providing managerial support. Specifically, multidimensional data sets are designed for the analysis of a large volume of RFID data and for flexible aggregations along multiple dimensions (e.g., product, time, location, etc.).
	[1]	Different options for RFID data analysis are discussed, by examining a supply chain composed of a manufacturer of fast-moving consumer goods (FMCG), a goods consolidator, a distribution centre and some retailers. The authors develop BI systems, based on RFID data, that allow managing the following typical issues of the FMCG retailing: maximization shelf space utilization; out-of-stock; computation of order quantity and times; prioritization of deliveries and goods distribution (for consolidator).

Table 1: continuation

Business intelligence & data analytics	[5]	The issue of generating value from RFID data in the case of a FMCG warehouse is investigated. The author developed BI systems which allow for monitoring the flow time of items, deriving relevant KPIs of the warehouse assets utilization, and developing tracking & tracing functionalities for items handled in the warehouse.
	[6]	A Development of a set of relevant value-added information that can be derived from RFID reads in the context of FMCG; the study was performed exploiting real RFID data, gathered from a pilot study (the RFID Logistics Pilot project). The authors highlight that, by developing appropriate BI tools, value-added information can be derived in terms of product flow, flow time management, shelf life management, inventories and product traceability.
	[15]	Collecting data in an in-field analysis by a major US retailer is used, to derive interesting value-added information, concerning the actual movements of cases and pallets of products, from their arrival at the distribution centre to their end of life at the retail store. The authors also discuss the potential value of this visibility to supply chain players.
	[21]	Discussing the consequences of storing RFID data on transactional granularity in data warehouses and propose useful approaches for handling the resulting data volume.
	[22]	An in-field experiment, which compares 12 retail stores equipped with RFID to the resulting performance with additional 12 stores without RFID. The authors experienced a reduction of 16% in out-of-stock levels thanks to RFID deployment is experienced.

2.2 Research Gap

The contribution that we make to the literature is twofold. On the one hand, we provide significantly more detailed insights into the algorithms necessary for the processing of raw RFID data than prior studies. This approach is complementary to other studies that usually do not discuss the details of their algorithms at all and often rely on data gathered under unrealistic conditions. On the other hand, we investigate the interrelated topics of RFID data collection, data analysis, and generation of business value. In addition to the rather technological aspects of RFID data analytics, we also discuss the interpretation and use of the generated performance indicators and reports from a practitioner’s perspective. Not least, our research differs from the related work by the sheer size of the underlying data set, which was collected in a real-world environment in the course of more than one year.

3 Case Background

We consider a European retailer, which offers a wide assortment of products with apparel and footwear in more than 100 department stores. Besides the challenges pertinent to this industry, the company has particularly been facing strong competition in recent years by fast fashion retailers such as H&M and Zara (Inditex Group), which continually increase their market share and yield significantly higher margins due to tightly integrated supply chains [20], [30]. At the same time, its customers have not only become more demanding and price conscious, their behavior has also become increasingly difficult to assess, e.g., due to changing life patterns and the ageing of populations in Western countries [19], [36]. The retailer under study has been known for many years for its role as an early adopter of RFID in the supply chain due to a series of conducted RFID feasibility tests. Attracted by the promising benefits of RFID technology on the sales floor, the company started a large-scale trial to implement many of the RFID applications that have so far only been speculated about in one of its stores.

3.1 Basic Information of the RFID Trial

For our case study methodology, we collected RFID data during September 2007 and May 2009. We also prepared certain workshops and meetings with representatives of the retailer project team (see 3.4). Furthermore, members from different departments have been consulted with the aim to create a well-founded discussion. Additionally to these workshops, we conducted several telephone interviews to reflect the current project status at frequent intervals. Moreover, we were equipped with a detailed project documentation. With that supplement and during 4 extensive site visits of the retailer’s infrastructure and stores, we got a deeper insight in all RFID-relevant processes as well as an overview of various technological issues. For the 15 month trial starting in 2007, a total of \$150.000 was spent for the RFID infrastructure, which is described in the next section.

3.2 Description of the RFID Infrastructure

In order to evaluate the technology’s potential to address the previously mentioned issues, the retailer implemented an RFID infrastructure in the menswear department in one of its department stores with a total of 22,000 square feet in size. Figure 1 provides a schematic overview of the installation. All products to be sold in the store are shipped from a nearby distribution center (DC). Items are manually tagged at the DC and transported in trucks to the store, where they are read at the incoming gates. Goods intended for immediate sale are sent via the freight elevators to the menswear department, where they leave the back store and enter the front store via a transition gate. On the

average, 30,000 individual items equipped with transponder labels are constantly available to customers. Provided an item is not returned, the checkout reader usually poses the last part of the item's lifecycle in the store. Besides customer returns, items can also flow back from the front store to the back store if they are stored in the local storage room, returned to the supplier, or sent to other subsidiaries. Besides the clothing items, also merchandise fixtures were RFID-tagged. Clothing items were assigned to these fixtures, and the fixtures were assigned to a sales floor space, the so called *shop-in-shops* (see grey areas on the front store as indicated in Figure 1). Given this two-step aggregation relationship, every item was assigned to a fixture and thus to a shop-in-shop. Furthermore a total of 20 fitting rooms are aggregated into 5 fitting room clusters.

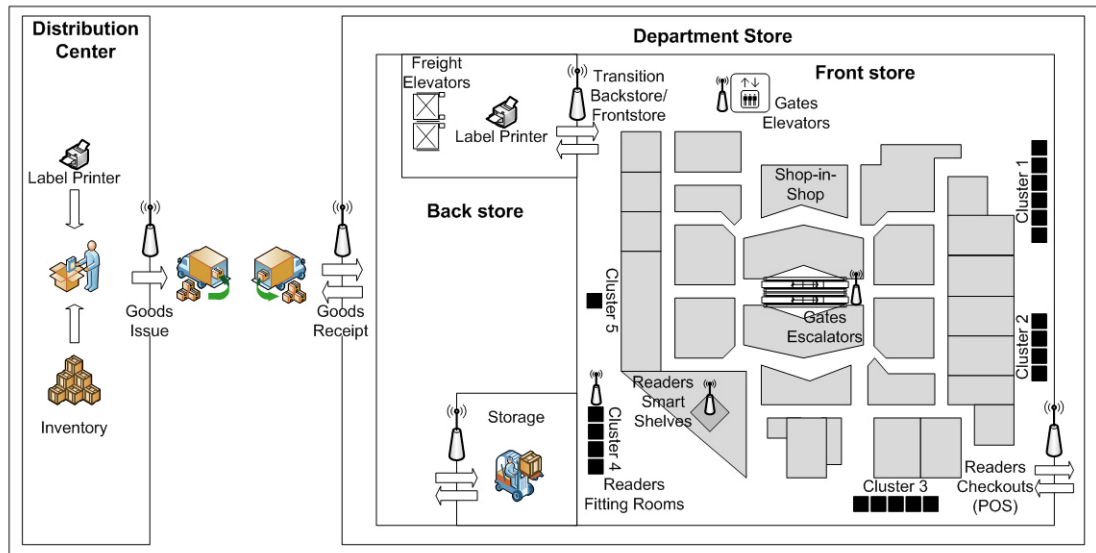


Figure 1: Schematic overview of the RFID installation

RFID readers are installed at the goods issue of the distribution center (2 readers), at the goods receipt of the department store (2 readers), at the labeling points (3 readers in the distribution center, and 3 at the department store), at the gateways between the sales floor and the backroom (2 readers), at escalators (3 readers) and elevators (2 readers), on *smart* shelves (10 readers), in all 20 fitting rooms (20 readers), at the storage (1 reader), and at the points of sales (7 readers). In total, the infrastructure includes 55 RFID readers and more than 200 antennas. Some RFID readers are equipped with photoelectric barriers and motion detectors in order to determine the direction of the flow of goods, e.g., upon transit from the back store. The installed RFID readers recorded the movements of thousands of RFID item-level tagged clothing. Table 2 provides an overview of the distribution of read events on the RFID readers during the trial. The absolute numbers represent a sufficiently large sample of read events, which improves the validity of our following analyses.

Table 2: Distribution of read events

Readers	Read Distribution	Absolute Numbers
Smart Shelves	41.22%	5053010
Front Store / Back Store Transition Reader	16.08%	1971094
Changing Room Readers	11.23%	1376753
Goods Receipt Reader	9.96%	1220624
Goods Issue Reader	9.69%	1188212
Escalator Exits	5.85%	717202
Labeling Point	2.90%	355528
Elevator Exits	1.73%	212009
Point of Sales	1.19%	146055
Storage	0.14%	16797

3.3 Managerial Perspective on RFID Data and Data Management

In order to address the strategic and managerial challenges of the apparel retail and the department store industry, a team of experts from the retailer identified and collected in various brainstorming sessions and internal workshops numerous data analyses in the area of layout, inventory, category, as well as department store process management. In several additional sessions and workshops, the identified analyses were clustered and categorized on a more general level. Finally, on the foundation of the collected data and the existing infrastructure, and going beyond what is already being done with the help of barcode-based sales and inventory data, the experts agreed on 15 data analyses that were most likely to create a financial impact and which had the biggest overall impact for the retail industry (see Table 3). Besides the data analyses, three infrastructure performance management indicators were defined in order to assess the data quality provided by the underlying infrastructure.

Table 3: Overview of identified and categorized RFID data analyses

Point of interest	Analysis	Purpose of the collected RFID Data
Infra-structure	Error Ratio of Bulk Reading at POS	Analysis of the performance indicators for the RFID installation and the gathered RFID data.
	Proportion of Unlabeled Merchandise Delivered	
	Assignment of Merchandise to Fixture	
Layout Mgmt.	Fitting Room Usage Distribution	Analysis of the number of try-ons on fitting room and cluster level.
	Catchment Area of Fitting Room Clusters	Analysis of the catchment area of fitting room clusters.
	Time-Dependent Visits of Fitting Room Clusters	Analysis of the number of try-ons in clusters according to time of the day.
Category Management	Correlation of Try-Ons and Sales	Analysis of the merchandise performance on different levels of detail, e.g., on the category-level or on the item-level. All analyses are based on the number of try-ons and the consequent sales. Moreover, analysis of complements and substitutes, as perceived by customers during their fitting room visits. Analyses of lead times of items on the front store or in the back store.
	Correlation of Try-Ons and Sales on the Category Level	
	Correlation of Try-Ons and Sales on the Item-Level	
Inv. Mgmt.	Out of Shelf but in Store	Analysis of the inventory management, including the number and duration of out of shelf or out of stock situations.
	Out of Stock – Store Replenishment	
Dept. Store Processes	Shelf Maintenance – Misplaced Merchandise	Analysis of the number and duration of misplaced situations on the shelves.
	Lead Time Analysis – Identification of Process Inefficiencies	Analysis of the process inefficiencies on the sales floor, for example.
	Front Store / Back Store Movements	Analysis of the occurrence of loops between front and back store.

For various reasons such as stray readings, missed reads, unreliable reads, and data redundancy, however, the collected RFID raw data had to be filtered and aggregated before it could be analyzed. Thus, in a second step, the necessary data filtering and aggregation techniques and mechanisms for the treatment of RFID raw data were defined in order to provide the necessary data basis for the business analyses. The following three sections 4 to 6 and its subsections of our paper follow the structure presented in Figure 2. Accordingly, all connections between the subsections are visualized and may be used to better understand the structure of our research.

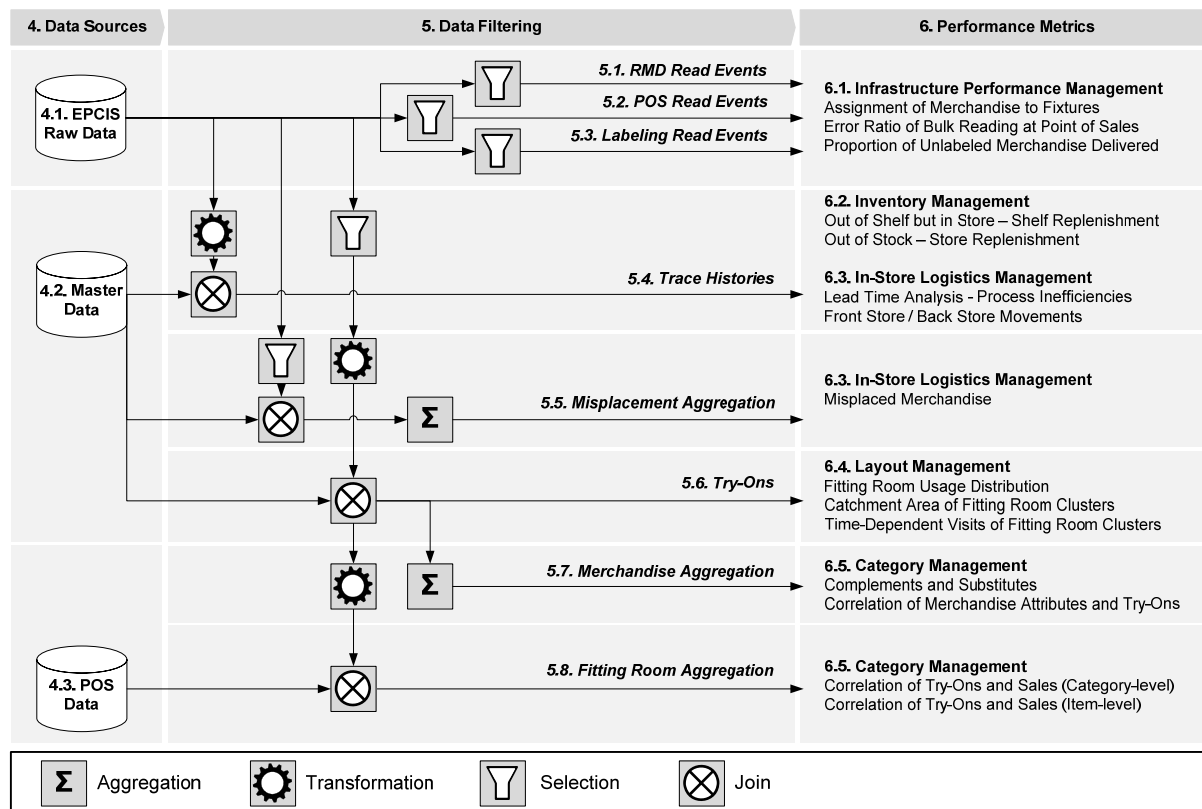


Figure 2: Overview of data sources, data filtering and performance metrics

4 Data Sources

Chapter 4 specifies the three different data sources we used and which are also shown in Figure 2. Altogether we give information about the use of Electronic Product Code Information System (EPCIS) Raw Data, Master Data as well as Point of Sales Data.

4.1 EPCIS Raw Data

EPCIS events are the data constructs stored in the EPCIS repository. Four generic event types are specified in the EPCIS data model. An *ObjectEvent* corresponds to the detection of Electronic Product Code (EPC) equipped items, e.g., on goods receipt upon arrival of a shipment. Since the granularity of individual EPC identification is not always needed, a *QuantityEvent* can alternatively be used that includes only information on the product type and the number of objects, which equals the level of detail obtained from barcode scanning. In contrast to that, an *AggregationEvent* is not associated to a particular tag read but rather denotes a (dis-)aggregation of a group of items, e.g., cases that are put on a pallet for shipping. In a similar way, a *TransactionEvent* links EPCs to a specific business transaction, e.g., a purchase order. Business documents can be linked to any recorded event, which are represented by instances of the *BizTransaction* type. All event types contain mandatory fields on date and time of an event, which objects or entities were subject to the event, where the event occurred, and in which business context the event happened. A complete description of EPCIS event types is provided by EPCglobal (2007).

4.2 Master Data

Master data are usually provided by the supplier or the retailer's procurement department. They describe a product on an article level, i.e., each data set refers to one specific European Article Number (EAN) code. This is in contrast to EPCIS event data, which refers to individual items. While the structure of master data records is the same for all product types (see Table 4), the range of values for single attributes (e.g., size and color codes) is supplier-specific, which complicates comparisons between products and categories.

Table 4: Master data description

Field	Description
EAN	European Article Number: a 13-digit article number identifies products on the article level, e.g., 5-pocket pants of a brand; the same pant in a different size or color can have a different EAN code.
Size	Article size information ranging in most cases from extra small (e.g. XS or 1 for shirts) to large (e.g. <i>waist=44</i> and <i>length=38</i> for jeans).
Department Number	Each clothing item is assigned to a department, such as <i>woolen</i> , <i>pants</i> , <i>accessories</i> , <i>suits</i> , or <i>shirts</i> . Each department has a three digit department number.
Article Group	Each clothing article is assigned to a article group number within the department number. This way, the shirts-department (123) can be subdivided into <i>short sleeve</i> or <i>long sleeve</i> shirts.
Segment	The segment combines the department number and the article group number in one string.
Color	Each clothing item's color, as preset by the supplier or manufacturer, e.g., <i>black 97</i> or <i>red 72</i>
Brand Name	Each clothing item's brand name.
Supplier	Each clothing item's supplier.
Season	Season to which clothing items are assigned to, e.g., <i>summer/spring 2008</i> .
Article text	Short description of the item (e.g., <i>5 pocket jeans</i> , <i>brand xyz</i>), comparable to <i>Segment</i> .
Retail Price	Retail price of the clothing item.
Wholesale price	Wholesale price of the clothing item.

4.3 Point of Sales Data

POS data are captured using barcodes, which encode traditional EAN numbers. Sales staff enters other data attributes, such as size, for example, manually. Despite the existence of master data, size information is collected redundantly, since some products carry the same EAN code for all sizes (see Table 5).

Table 5: Point of sales data

Field	Description
EAN	13 digit European Article Number.
Size	Size information, e.g., <i>M</i> , <i>L</i> , <i>XL</i> , or <i>44</i> , <i>52</i> , or <i>5</i> , <i>6</i> , <i>7</i> (see Table 4).
Time	Purchasing time.
Price	Retail price.
Amount	Amount of products purchased; a negative number indicates returns.
Department	Number that indicates which department the item is assigned to.
Checkout Number	Number of the checkout where item has been purchased.
Receipt number	Running number of receipt numbers, daily reset.

5 Data Filtering

Chapter 5 describes the different mechanisms which we used to filter the data generated by the sources outlined in the prior sections. The scripts we created for the data filtering can be found in the appendix.

5.1 RFID Mobile Device Read Events

Just like merchandise sales items on the sales floor, each merchandise fixture and each shop-in-shop area (cf. sub-section 3.2 and Figure 1) has a unique identifier. In a manual process on a regular basis, the RFID Mobile Devices (RMD) are used to assign each item to a merchandise fixture, and to assign each fixture to its shop. Therefore, the fixture's tag is scanned before all items lying on the fixture are scanned. Thereby, the items are assigned to that very fixture. To verify whether this process has been conducted on a regular basis, the RMD read events were selected from the EPCIS raw data and analyzed. This way the first performance indicator for the RFID installation was created.

5.2 POS Read Events

For the acceleration of the sale process by bulk reading, and thus the reduction of waiting lines at the point of sales, a total of eleven checkouts on different floors were equipped with an RFID reader that was integrated into the checkout counter. The readers were positioned in a way that only the items on the marked reading area were read. The sales personnel were required to swipe each article over the marked reading area, while at the same time activating the reader by pressing the activation foot pedal underneath the table. The according RFID POS events

were selected for this analysis and compared to the barcode POS events. This comparison provided a second performance indicator for the RFID infrastructure and data.

5.3 Labeling Read Events

All products to be sold in the menswear department of the retailer's trial department store were shipped via the distribution center (DC) where they were manually item-level tagged, using three labeling printers to print RFID tags. Three more printers were available in the department store for tagging items, which (i) miss their tag, (ii) have a broken tag, (iii) are returned and need to be retagged, or (iv) which arrived from the distribution center without a tag.

5.4 Trace Histories

Trace histories, which transform RFID raw data into product life cycles are necessary for analyses in the area of process execution and lead times (e.g., process inefficiencies or process cycles) as well as for inventory management analyses (e.g., *out of stock but in store* or the *out of stock – store replenishment* analysis). A typical product life cycle starts at the labeling point in the distribution center and ends at the check-out in the department store. In between, for example, the product is placed on a smart shelf, tried on, brought back into the back store, returned to the smart shelf, tried on until it is eventually sold. In order to generate trace histories, each EPC's read events are grouped and sorted by their time stamps (*eventTime*).

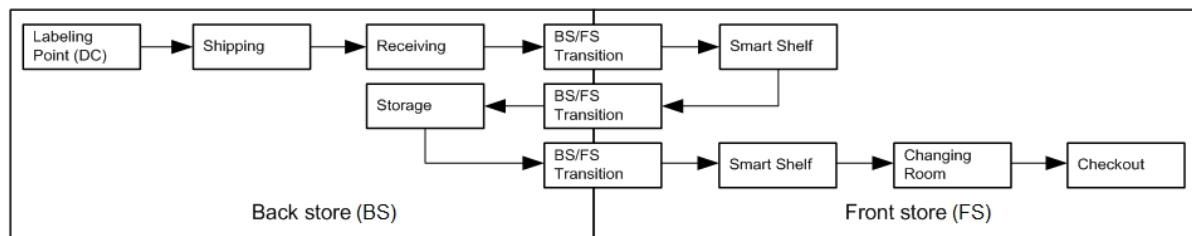


Figure 3: Physical item movements between back store and front store

In a second step, the item's movement data is transformed into a linked list by interconnecting the individual read events using the attributes *ReadPoint* and *eventTime*, and the knowledge about the store layout and the RFID installation. Figure 3 provides an overview of an exemplary product life cycle. In the present case, we examine the tagged product from receiving to checkout.

Table 6: Transforming EPC events to trace histories

EPC	eventTime_first	eventTime_prev	eventTime_next	Diff.	Readpoint	Readpoint_prev	Readpoint_next
EPC1	22-Feb-08 12:23	-	23-Feb-08 07:45	0.81	Labeling Point	-	Shipping
EPC1	23-Feb-08 07:45	22-Feb-08 12:23	23-Feb-08 09:36	0.08	Shipping	Labeling Point	Receiving
EPC1	23-Feb-08 09:36	23-Feb-08 07:45	23-Feb-08 10:23	0.03	Receiving	Shipping	BS/FS Transition
EPC1	23-Feb-08 10:23	23-Feb-08 09:36	23-Feb-08 10:45	0.02	BS/FS Transition	Receiving	Smart Shelf 1
EPC1	23-Feb-08 10:45	23-Feb-08 10:23	19-Mar-08 09:10	24.93	Smart Shelf 1	BS/FS Transition	BS/FS Transition
EPC1	19.03.2008 09:10	23-Feb-08 10:45	19-Mar-08 09:45	0.02	BS/FS Transition	Smart Shelf 1	Storage
EPC1	19.03.2008 09:45	19-Feb-08 09:10	15-Sep-08 08:20	179.94	Storage	BS/FS Transition	Storage
EPC1	15.09.2008 08:20	19-Feb-08 09:45	15-Sep-08 08:40	0.01	Storage	Storage	BS/FS Transition
EPC1	15.09.2008 08:40	15-Sep-08 08:20	15-Sep-08 08:58	0.01	BS/FS Transition	Storage	Smart Shelf 1
EPC1	15.09.2008 08:58	15-Sep-08 08:40	22-Sep-08 11:23	7.10	Smart Shelf 1	BS/FS Transition	Fitting Room 6
EPC1	22.09.2008 11:23	15-Sep-08 08:58	22-Sep-08 12:40	0.05	Fitting Room 6	Smart Shelf 1	Checkout
EPC1	22.09.2008 12:40	22-Sep-08 11:23	22-Sep-08 12:40	0.001	Checkout	Fitting Room 6	Checkout

In a third step, these traces are joined with the master data. This step includes the calculation of the temporal delays in days (see column 5 *Diff* in Table 6) between individual read events by subtracting the current event time from the previous one. This information can be seen in Table 6. Since the front and back store location information of each read point is known, lead times, inventory, and process execution analyses can be conducted based on the resulting table. Hence, Table 6 provides a starting point for trace history based analyses.

5.5 Misplacement Aggregation

RFID-enabled shelves (*smart shelves*) on the sales floor constantly interrogate tags and update their inventory level, including items' master data such as different sizes and colors, in the RFID database. In order to limit the amount of

data to be stored in the database, the periodicity of EPCIS data updates was set to 15 minutes. From these smart-shelf read events, inventory data was generated that allows for tracking inventory fluctuations over time. Besides analyses of *in stock but not on shelf* situations, this data allows for the identification of misplaced merchandise, since all items placed on the shelf are read, irrespectively of whether they are supposed to be there or not. All items on merchandise fixtures or shelves, which are not supposed to be on the shelf, can be considered as misplaced merchandise, as defined by the visual merchandise department. The read events of every single smart shelf were copied into a distinct database table and joined with the master data. The items, which are supposed to be on the smart shelf, were filtered in order to receive the misplacements on each particular shelf. The tables were sorted by item (EPC) and time using a script (see Appendix A). The number of consecutive read events was counted and stored in the database for every individual misplaced item. The used script counts the number of consecutive read events of a particular EPC on the shelf. This number multiplied by 15 equals the number of minutes an item has been misplaced on that shelf.

5.6 Try-Ons

For the analysis of customer try-ons in fitting rooms, we considered all fitting-room read events from the 20 fitting rooms on the sales floor. Fitting rooms were individually identifiable by the read event's attribute *ReadPoint*. With a basic database query, we selected all fitting room read events and copied them into a distinct table. Each fitting room was assigned to its cluster: According to Figure 1, Fitting Rooms (FR) 1-7 were assigned to Cluster 1, FR8-11 to Cluster 2, FR12-15 to Cluster 3, FR16-19 to Cluster 4, and FR20 to Cluster 5. As a result, we received a table with all visits (try-ons) in fitting rooms and in clusters, respectively. For example, by analyzing the life cycle of an item-level tagged 5pocket pants, we discovered that these pants had been read within several seconds in different fitting rooms of one cluster. The reason proved to be the broad read range of the fitting room readers, resulting in readings of neighboring fitting rooms and of people passing by the fitting rooms. Figure 4 provides an overview of these fitting room reading errors resulting: (I) reads by customers or employees passing by, (II) reads by adjacent cabins, (III) reads by left back items, and (IV) reads from nearby merchandise fixtures.

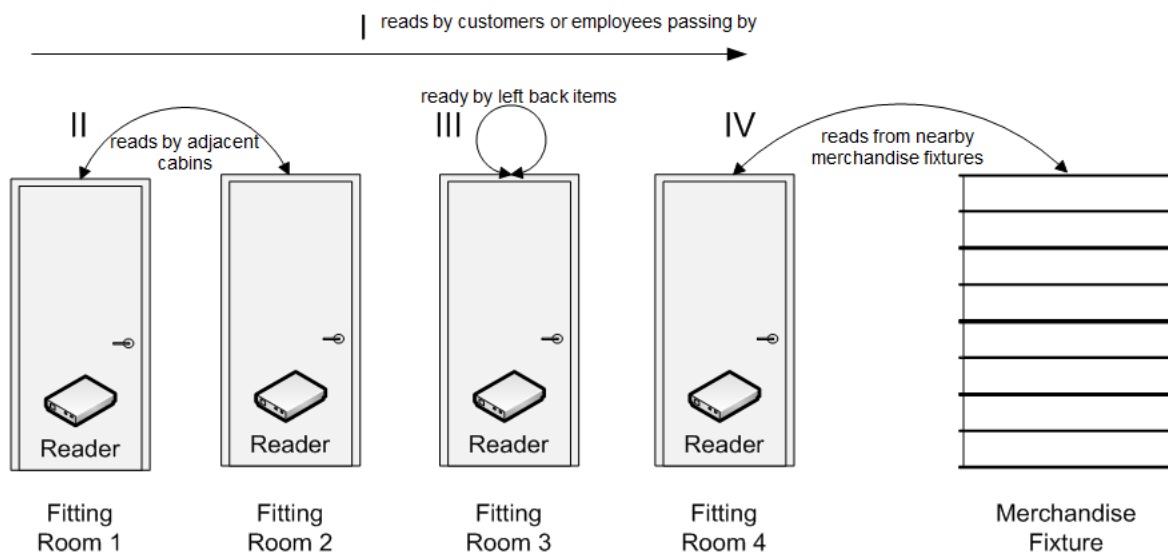


Figure 4: Fitting room reading errors

The noise from case I to IV had to be filtered for the try-on-based data analyses to be conducted. We did so by designing filtering algorithms that iterated through the table and filtered the described reading errors. From a filtering perspective, Cases I and II, as well as Case III and IV are comparable. For Cases I and II, a filtering algorithm was designed (see Appendix B) that filtered these reading errors by deleting the events of a certain EPC that occurred within a certain time span in other (adjacent) fitting rooms. Therefore, we first ordered the fitting room events first by fitting room, then by EPC, and finally by time using a simple SQL statement. We iterated through the resulting list by copying *real* fitting room read events into a newly created table called *filtered_fitting_room_read_events_table* (i.e., customer try-ons).

As for Cases III and IV, the events of the corresponding EPCs were counted only once, even if they were read for hours. Using a basic SQL statement, we therefore ordered the table *FITTING_ROOM_READ_EVENTS* by fitting room number first, then by the EPC, and finally by time. For each fitting room, we iterated through the list by EPC and calculated the time from the first RFID read event, until the last event recorded by the particular fitting room reader. If it was read longer than 30 minutes, we just copied it once into the table of filtered fitting room events, using another script (see Appendix C).

The resulting table included a complete noise-free list of individual try-ons. Moreover, the solution provider adjusted the fitting room readers in a way that the noise was minimized. For the evaluation of our filtering and aggregation procedures, we conducted several tests in the department store by taking RFID-tagged items through the store and passing by the fitting rooms while logging the routes. Finally, the combination of the hardware adjustments and the filtering and aggregation algorithms provided good results. We then joined this table with the master data tables using the keys *EAN* and *Size*. This table served as the basis for all fitting-room and try-on analyses.

5.7 Merchandise Aggregation

For various analyses it was necessary to assign fitting room visits to individual customers. Since the customers, of course, weren't RFID-tagged themselves, we had to find a different, though anonymized way to match customer visits to tried-on items. Using a basic SQL statement, the fitting room reads were therefore sorted by fitting room, then by the read's *eventTime*. As a result, we received a table with the try-ons in all fitting rooms, sorted by fitting room (*ReadPoint*) in chronological order (see Table 7). Considering the fact that a fitting room is visited by only one person at a time, and that we worked on filtered data (see 5.6. Try-Ons), we could aggregate consecutive fitting room visits to one customer. An example is given in the first four lines of Table 7.

Table 7: Filtered fitting room visits, ordered by fitting room, then by time

EPC	EAN	eventTime	ReadPoint	Cluster	Sold	Customer
EPC1	456789123	Mo, Aug. 11, 2008 02:00.17 pm	Fitting Room 1	Cluster1	0	1
EPC2	123456789	Mo, Aug. 11, 2008 02:00.20 pm	Fitting Room 1	Cluster1	1	1
EPC3	789123456	Mo, Aug. 11, 2008 04:00.30 pm	Fitting Room 1	Cluster1	1	2
EPC4	456123789	Mo, Aug. 11, 2008 04:00.40 pm	Fitting Room 1	Cluster1	1	2
...	Fitting Room 1	...	0	3
...
EPC1	...	Mo, Aug. 11, 2008 10:00.01 am	Fitting Room 2	Cluster1	0	20
EPC17	...	Mo, Aug. 11, 2008 11:00.32 am	Fitting Room 2	Cluster1	0	21
...
EPC20	...	Mo, Aug. 11, 2008 09:15.22 am	Fitting Room 20	Cluster5	1	32

The first two lines show two EPCs being read at 2:00pm in fitting room number 1, and the next two lines show two different EPCs being read at 4:00pm, i.e., two hours later in the same fitting room. Although it cannot be excluded that the same customer tried on items 2 hours later in the same fitting room, these findings lead to the conclusion that the readings belong to two different customer visits, one at 2:00pm, and another one at 4:00pm. Based on this assumption, we designed a script that iterated through the table and added the last column with an aggregation number for each customer (see Appendix D).

With a threshold of 10.5 minutes, i.e., each consecutive read in that fitting room within 10.5 minutes from the last read, was attributed to the visiting customer. The threshold of 10.5 minutes was chosen because of the technical implementation of the RFID installation and because it provided the best results.

5.8 Fitting Room Aggregation

For all analyses that assessed the performance of merchandise items, and the correlation of try-ons and sales events, respectively, it was necessary to combine the filtered try-on read events from the fitting rooms with the POS data. Due to the reasons described in subchapter 5.2., we could not rely on RFID checkout data only but had to include the conventional barcode-based POS data. We did so by analyzing all consecutive sales of items that were tried on in the fitting rooms. Therefore, we matched the EAN code and the size of the tried-on item with the sales events at the POS within the next hour of the try-on event (common denominator). We plotted the resulting data in Figure 5.

The figure shows that the majority of customers who have an intention to buy their tried-on items do so within 30 minutes after they have left the fitting room. Knowing this, we could add to every tried-on item a flag whether the item has been purchased after the try-on event or not, by joining the fitting room data and the barcode POS data using the keys *EAN* and *size*, and the *eventTime* of both events. All consecutive purchases conducted later could not be attributed to the initial try-on event anymore. After describing our several filter mechanisms, we continue (according to Figure 2) with the performance metrics to analyze, how the RFID Data can be used to improve the store management.

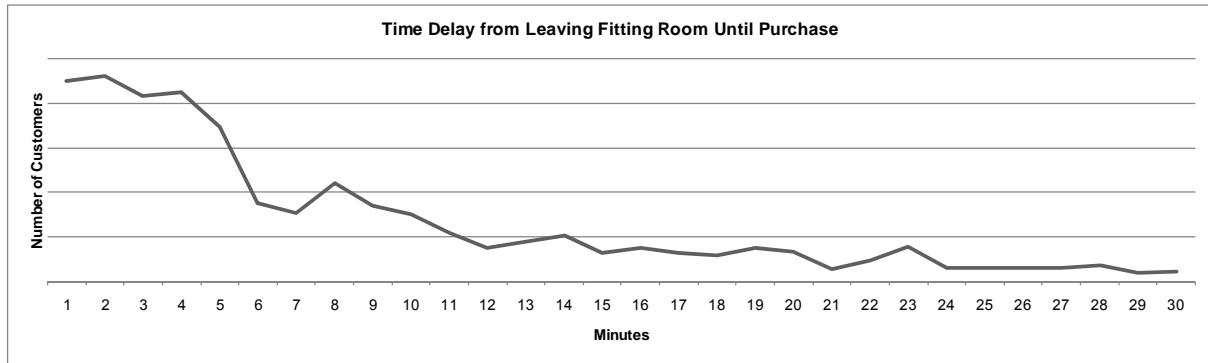


Figure 5: Time delay from leaving fitting room until purchase

6 Key Performance Indicators and Data analyses

This chapter finally exhibits, how we used and analyzed the filtered data from the described data sources to create different performance metrics as Infrastructure Performance Management, Inventory Management, In-Store Logistics Management, Layout Management and Category Management. We show how this valuable information provided can be used to support decision makers in retail organizations and to improve store management processes.

6.1 Infrastructure Performance Management

This sub-chapter presents three infrastructure performance management analyses that were conducted during the trial. The first analysis describes the process of assigning merchandise items to fixtures, the second analysis assesses the quality of the RFID POS data as compared to the barcode-based POS data, and the third analysis provides an overview of the number of merchandise items that arrived in the department store without having been tagged at the DC.

6.1.1 Assignment of Merchandise to Fixtures

The data analysis revealed that after the third month of the trial, the employees hadn't conducted the process on a regular base. The reasons lay in the resource- and time-intense manual process, whose benefits were not transparent to the employees. Therefore, only the data of the first three trial months could be used for further data analyses.

6.1.2 Error Ratio of Bulk Reading at Point of Sales

The RFID-based checkout was performed in parallel to the barcode-based EAN checkout, as the RFID checkout readers were part of the trial and connected to the RFID-based merchandise information system (MIS) only. This system was kept separate from the barcode MIS. Consequently, the data quality of the RFID MIS depended on the reliability and the discipline of the employees at the checkout: items that were not read by the RFID reader remained in the RFID MIS although they were sold (false negative); items that were mistakenly read by the checkout counter while the reader was activated were marked as sold, although they weren't (false positive).

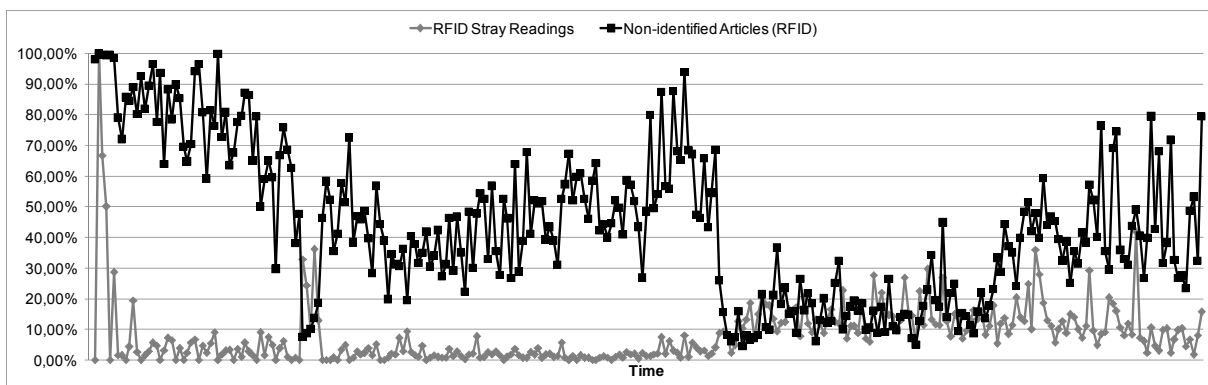


Figure 6: Percentage of stray readings and non-identified articles during the trial

Moreover, in the middle of the trial, the management decided to permanently activate the RFID reader underneath the checkout table. Figure 6 illustrates the increase in the read rate but also of stray readings during the trial, illustrated by the rapid decrease of the *non-identified Articles (RFID)* in the figure. It revealed that RFID data can be incomplete. Consequently, RFID POS read rates were far too bad. POS data had to be included for all analyses relying on sales data (see sub-chapter 5.8).

6.1.3 Proportion of Unlabeled Merchandise Delivered

In Figure 7 we provide an overview of the distribution of labeling activities between the DC and the department store. It shows that only the initial labeling activities were dominated by the department store's labeling printers in the department store, since existing items on the sales floor and in the storage had to be tagged retroactively. Afterwards, the majority of labeling activities was conducted in the DC. However, besides the above mentioned cases like broken tag, missing tag, or retagging of returns, around 5% of the items arrived without a tag in the department store and had to be tagged locally. In what concerns returns, the employees could only re-insert items into the RFID MIS that still had their EPC tag attached. For items returning without their RFID tag, new ones with new EPC numbers had to be printed. This analysis revealed that RFID data can be incomplete, not only due to missing reads at the read points, but also due to the fact that the items didn't have an RFID tag attached when they left the DC, or when they came back as customer returns. Consequently, all EPCs with incomplete traces had to be excluded from the product life cycle and lead time analyses.

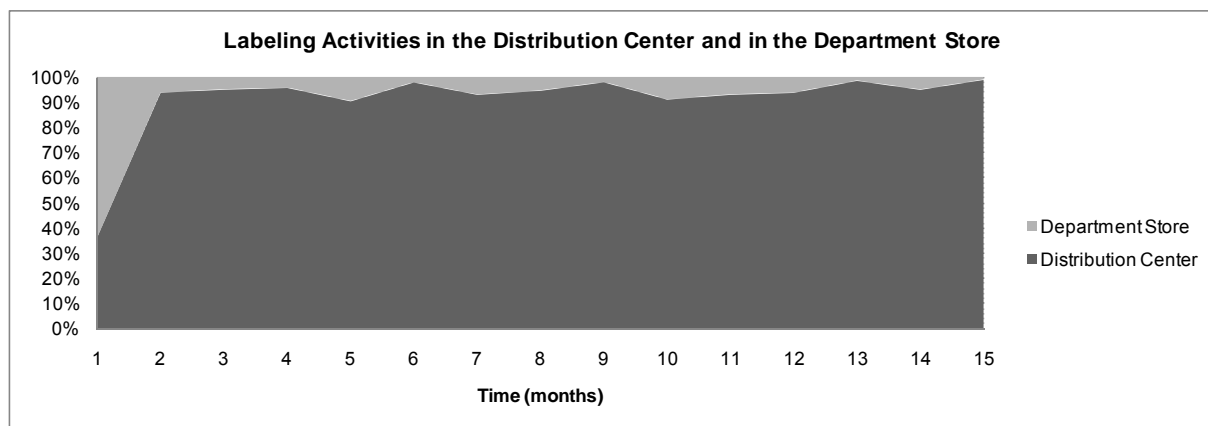


Figure 7: Distribution of labeling activities between the DC and the department store

6.2 Inventory Management

The real-time management of correct inventory levels is one of the main challenges of the retail industry. Incorrect inventory levels can lead to out-of-shelf but in store, or out of stock situations, which lead to lost sales opportunities and a decrease of customer satisfaction [35], [44]. Hence, the objective of an efficient inventory management is the reduction of capital costs through just-in-time ordering and the automatic replenishment using optimal triggers. This sub-chapter presents two inventory management analyses, based on the presented RFID infrastructure and the within gathered data.

6.2.1 Out of Shelf but in Store – Shelf Replenishment

Using trace histories, the front store and back store inventory of items can be calculated in real-time, and retrospectively for any given time and date. A list can be generated for items on EAN level, for each day of the year and plotted as a graph (see Figure 8), for example. By providing more visibility regarding the distribution of the department's inventory stock between front store and back store, this analysis supports the retailer in reducing the number of out-of-shelf but in stock situations, which is responsible for missed sales opportunities and decreased customer satisfaction. Moreover, this analysis can help to detect promotion execution errors (i.e., promotional articles, which are put on the sales floor too early, see I. in Figure 8) and storage inefficiencies (i.e., low inventory levels on the front store, but high inventory levels in the back store, see II and III. in Figure 8).

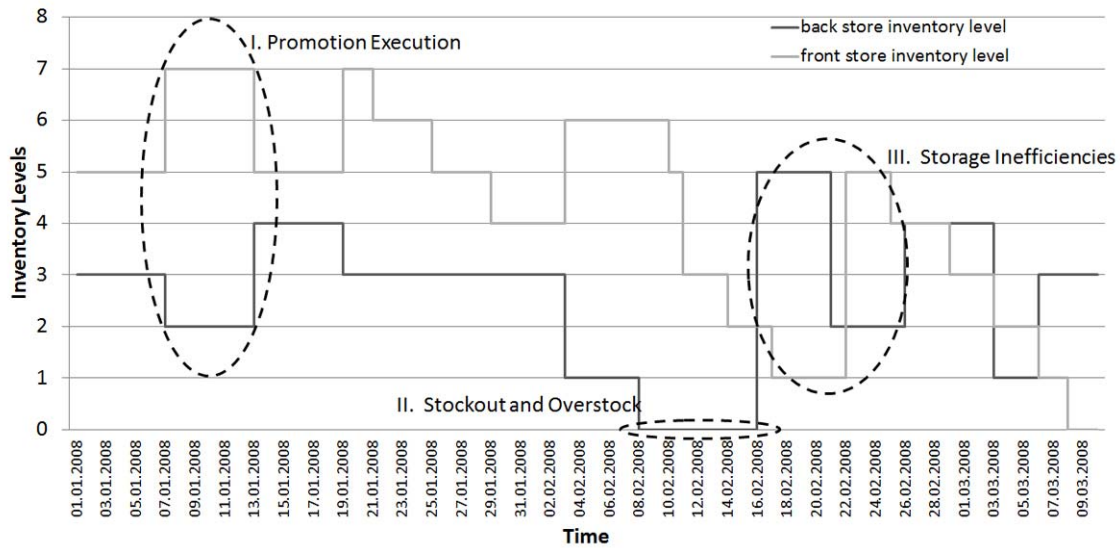


Figure 8: Percentage of stray readings and non-identified articles during the trial

6.2.2 Out of Stock – Store Replenishment

In the grocery industry, for example, 11% of the stock-outs can be traced back to faulty store forecasting practices and 21% to store ordering processes of the department stores. Another 9% of the stock-outs can be explained by the order handling at the DC [14]. During the trial, no Out of Shelf but in store situation has been observed. Therefore, the analysis has been extended to Out of Stock and Not in Store (OOS-NOS) situations. Such a case is exemplified in the case of an all seasons black suit with a retail price of EUR 79. Within the observation time, the article was unavailable and not replenished during 124 weekdays. In that time period, no sales can be generated with this specific item although it is probably marked as *in stock* but physically not available. RFID identifies such constellations and could update the stock level and trigger the replenishment. The inventory log of the observed merchandise in the department store is visualized in Figure 9.

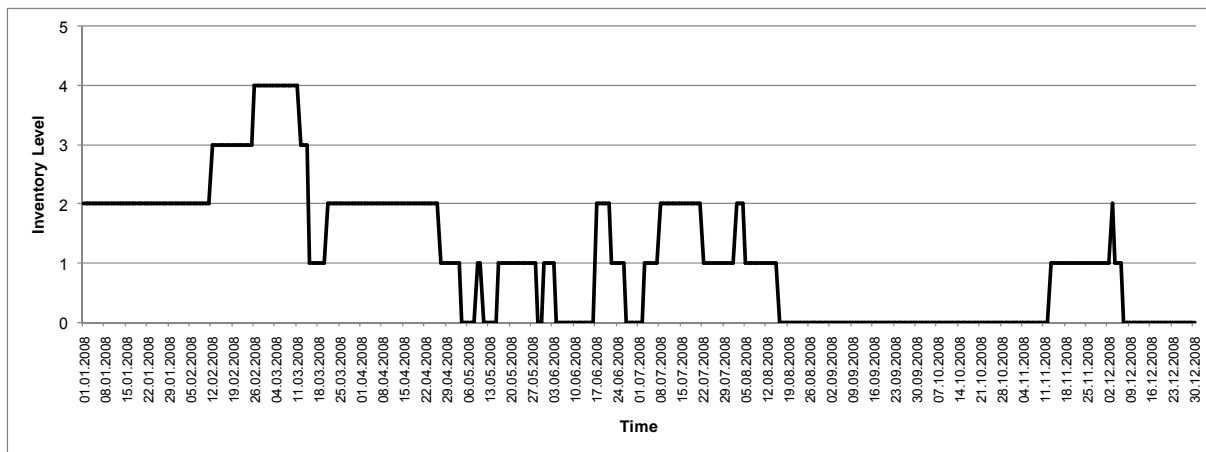


Figure 9: Exemplary inventory log for a NOS suit

6.3 In-Store Logistics Management

The retailer under study was interested in using RFID to assess its in-store logistics management and processes. The question was whether RFID can help to detect process inefficiencies that were previously invisible, to improve them accordingly, and to supervise the proper functioning afterwards. The analyses taken into account include the lead-time analysis, the analysis of front store and back store movement of items, and the misplaced merchandise analysis.

6.3.1 Lead Time Analysis – Process Inefficiencies

Regarding process execution, the retailer was interested in assessing the efficiency of its in-store processes. The RFID event data was analyzed for patterns, such as loops in the storage, products that were taken very often to the fitting room, or products that moved between different floors.

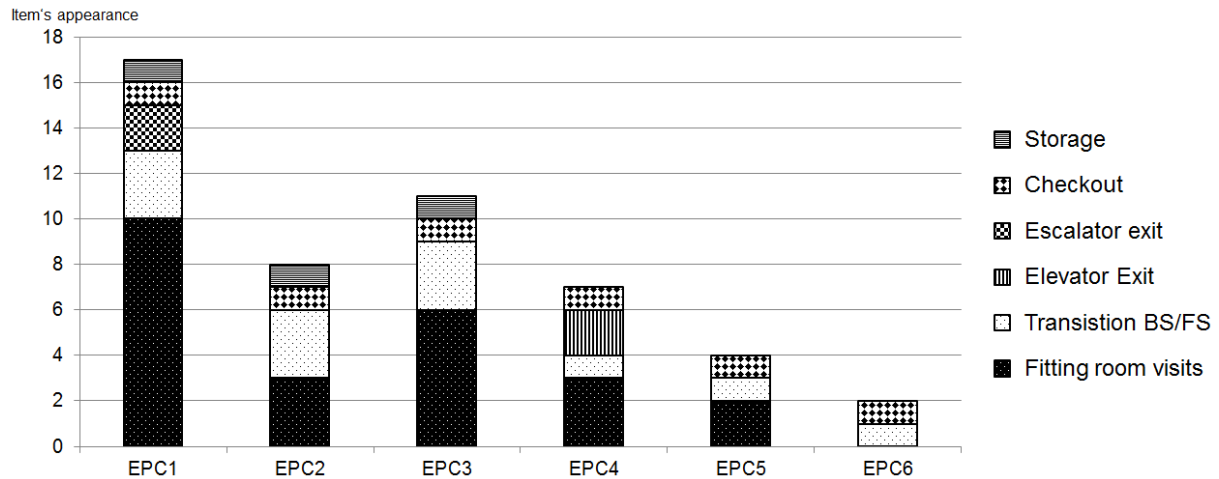


Figure 10: Process cycles (Comparison between different Items/EPCs)

For that purpose, trace histories were aggregated in order to count the number of an item's appearances at different read points, such as fitting room readers, readers between front store and back store, elevator and escalator exits, storage, and checkouts (see Figure 10). This analysis provides the retailer with a higher process visibility and enables him to detect previously unobserved inefficiencies.

6.3.2 Front Store / Back Store Movements

The retailer was also interested in how far RFID can detect process inefficiencies in in-store logistics processes that were previously intransparent. This analysis is used to detect *loops* between back and front store, i.e., items that were initially put on the sales floor and weren't sold during one season, for example, put back in the storage and then again on the sales floor. An over average number of reads on the transition between back and front store, either on the (average) category level, or on the item level, helps the retailer to optimize its in-store processes.

6.3.3 Misplaced Merchandise

Customers frequently misplace merchandise while trying them on or by carrying them around the store [11]. Misplaced merchandise artificially lowers the merchandise availability since the items requested are not found where expected. This reflects an Out-of-Shelf but on sales floor situation. The so-called *phantom stock-out* situations were analyzed by [42] at Borders Group Inc., a books and music superstore chain. They thereby demonstrated that an increase in misplaced products at the store is associated significantly with a decrease in sales.

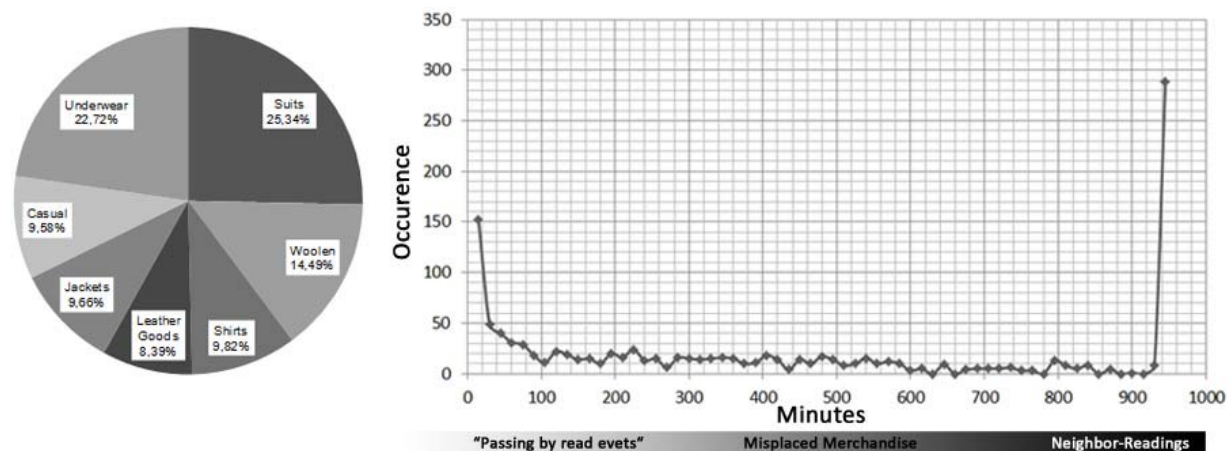


Figure 11: Occurrence of misplaced merchandise on the smart shelves (left: by category, right: Duration and occurrence of misplaced situations)

The retailer under study was interested in the number and the duration of misplacements on the sales floor. For this analysis we counted the occurrences of misplaced situations on the smart shelves and their duration. Figure 11 illustrates the occurrences and the durations of misplaced situations on one smart shelf on the front store.

Misplaced items with only one entry in the database stemmed most likely from customers or employees, which passed by the smart shelf with the item (e.g., after a fitting room visit), or from misplacements with short durations (*passing by read events*). Items that had so many entries that it covered the whole opening times of the store were likely reads of the shelf from neighboring departments due to its closeness to these departments (*neighbor-readings*). Items that had been read for durations between half an hour until several hours were likely to be misplaced items.

6.4 Layout Management

The design of the store's interior and its layout, the so-called visual merchandising or layout management influences the customer's purchasing behavior to a great extent [23]. Hence, it is considered, besides customer service and store management, as a central part of the store management [26]. It includes the optimal and appropriate use and design of fixtures, displays, color, lighting, music, scent, ceilings, and floor. Therefore, the retailer's goal was to assess in how far RFID technology can improve the layout management in department stores. Based on the underlying infrastructure, the following two questions will be discussed: how can RFID improve the sales floor allocation between sales floor and fitting room space (number of fitting rooms on the sales floor and number of fitting rooms within one cluster), and how can RFID help to identify the catchment areas of the fitting room clusters on the sales floor (yielding to the positioning of the fitting rooms and the cross-selling potential).

6.4.1 Fitting Room Usage Distribution

In order to analyze the utilization of fitting rooms, try-on data are counted and grouped into clusters (see Figure 12). This analysis shows, that 55% of all try-ons in Cluster 2 are conducted in FR7. Further on FR10 is almost not used (5% usage). With the information provided by this analysis, the retailer can reassess the trade-off between the total number of fitting rooms on the shop floor and within one cluster – with a surface of 16 square foot each – against more sales floor space. Besides the overall utilization of the clusters, the analysis could also point out the utilization during peak times, thus helping the retailer to either eliminate some of the fitting rooms in the clusters in favor of sales floor surface, or in favor of fewer, but bigger fitting rooms respectively.

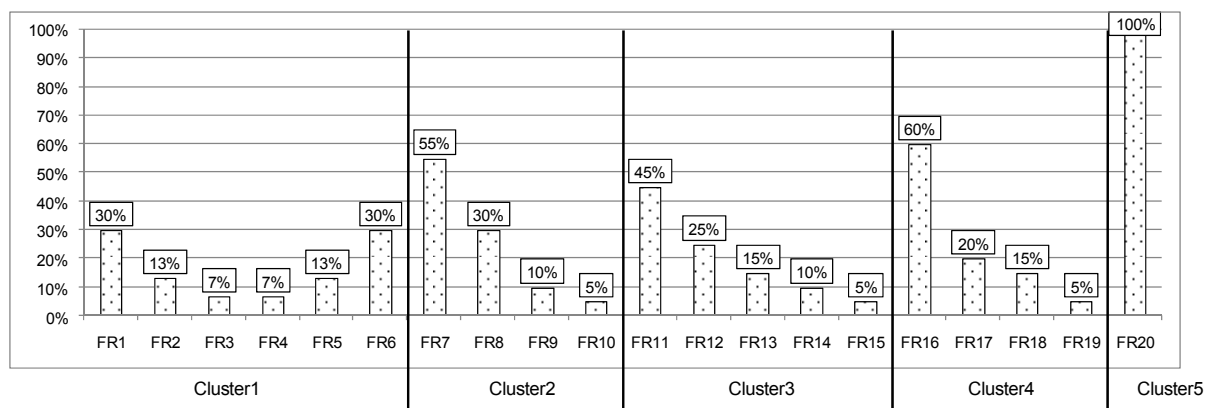


Figure 12: Utilization of the fitting room clusters

6.4.2 Catchment Area of Fitting Room Clusters

Figure 13 provides an overview of the sales floor, subdivided into shop-in-shops of article groups and brands. Besides matching of items to fixtures, fixtures are also matched to their shop. Together with the information on the previously described try-on events, this information can be used in order to visualize the *catchment areas* of fitting rooms. In the example in Figure 13, grey shading indicates the number of products that were taken to one of the fitting room clusters marked in black. The graphical visualization helps the retailer to optimize his sales floor layout and the positioning of the fitting rooms, and gives indications about movement patterns of customers. The results can also be used to optimize product placements on the sales floor (e.g., complementary articles) or to influence customers' shopping paths.

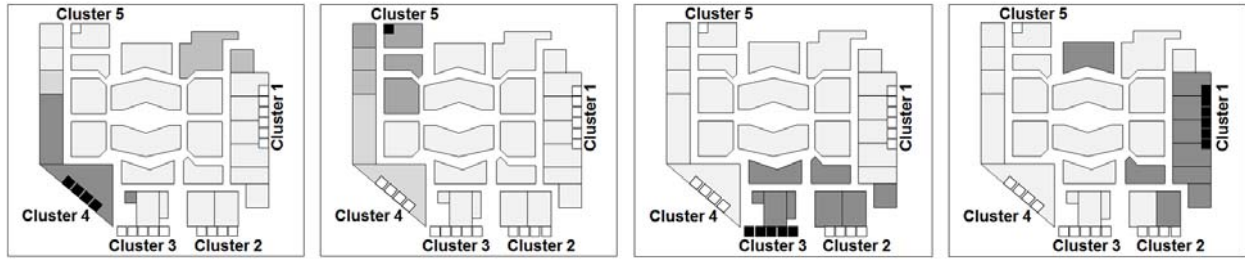


Figure 13: Catchment areas of the fitting room clusters

6.4.3 Time-Dependent Visits of Fitting Room Clusters

Based on the same data, the retailer can, for example, compare try-ons over time with the consecutive sales events in order to optimize the number of employees on the sales floor. The chart depicted in Figure 14 illustrates the number of try-ons for a given article group during the opening hours of one day, the number of sales events, and the ratio of both values (represented as bars). The data indicate that the ratio of try-ons and sales events changes significantly over time. In this case, the retailer under study could not profit proportionally from the rush of interested visitors, trying on articles in the early evening times. The comparison of the try-on/sales ratio at 3 pm with that at 5 pm shows that an increase in try-ons by more than 100% could not be translated into a corresponding sales increase. This might be a consequence of insufficient staff on the sales floor during these hours, resulting in lost sales, especially for articles that require extensive customer counseling, such as suits, for example.

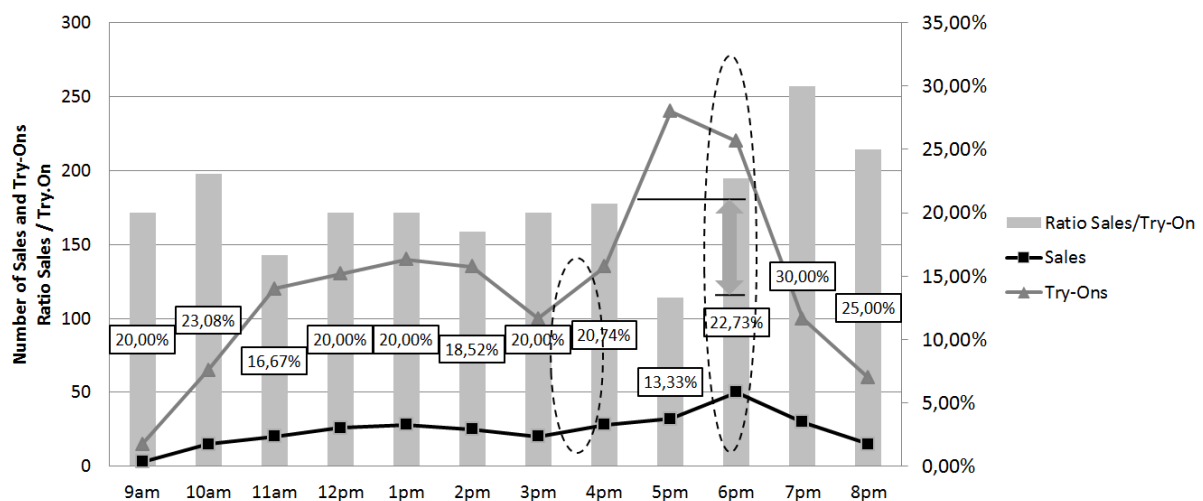


Figure 14: Sales probability (Ratio try-on / Sales ratio) over business hours

6.5 Category Management

According to [8], the product assortment strategies of retailers have undergone drastic changes in the past decade, from large assortments in the early 1990ies to the current emphasis on streamlined, efficient assortments. The broad assortments were based under the assumption that larger product assortments better meet consumer needs and that they increase the probability that consumers will find their ideal product and offer flexibility for variety seekers. However, the tremendous increase in stock keeping units (SKUs) resulted in higher inventory costs for retailers and more stock-outs. Later studies in groceries revealed that the reduction of SKUs per category showed no significant negative impact [8]. In some cases, reductions even resulted in an increase in sales [16]. The experiments by [7] showed that no change was revealed in the consumer's assortment perception but an increase in customer experience. This is especially likely if a few products of the category drive the majority of sales. Furthermore, the customer perception is influenced by how the assortment is organized, displayed and presented on the sales floor [8]. The following analyses aim at assessing the value of RFID data for the category management by providing a toolset to support the evaluation of the assortment and the merchandise performance, as well as the shop layout and the shelf space allocation on the sales floor.

6.5.1 Complements and Substitutes

In order to optimize category management and to learn about customer's fitting room visits within the menswear department store, the retailer under study was interested in the article groups that the customer's perceived as

substitutes or complements. For this purpose, try-on data were joined with themselves to investigate which items from two particular product categories were combined during fitting room visits (see Figure 15). This information helps the retailer under study to gain an overview of the customer perception of its products in order to improve the category management. This analysis can moreover be conducted on the level of brands, within one article group, in order to analyze which brands are perceived as substitutes.

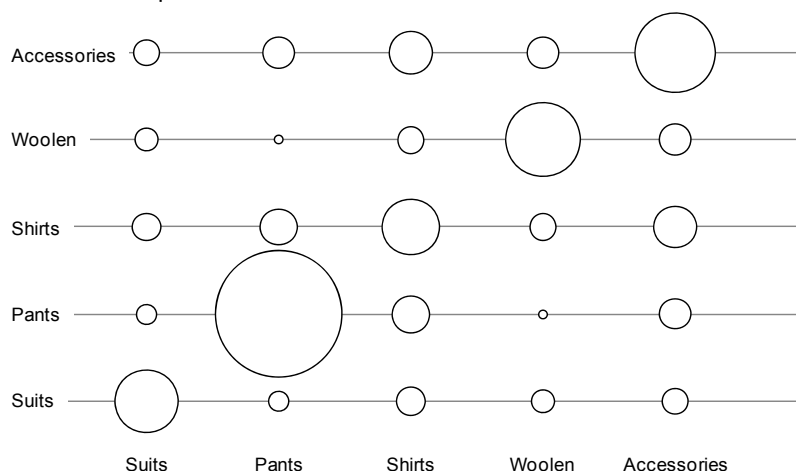


Figure 15: Complements and substitutes on the category level

6.5.2 Correlation of Merchandise Attributes and Try-Ons

The combination of try-on data from the fitting room readers and the master data from the retailer under study, allowed for correlations between different kinds of master data entries and try-ons. For example, the retailer correlated the try-on data with the items' prices, sizes, colors, brands, and all other entries from the master data in order to optimize his assortment by gathering knowledge about customer behavior.

6.5.3 Correlation of Try-Ons and Sales (Category-Level)

The try-on and sales data can also be correlated on the category level. The resulting data allow for various analyses of the attractiveness and performance of entire categories. Figure 16 illustrates an example on the category level between the number of try-on events, compared to the number of sales events. The analysis on the category level aims at answering the question which categories are typically tried-on before they are sold. Knowing the try-on frequency, respectively the sales-ratio/try-on ratio for each category shall provide valuable insights for the shop layout and the positioning of the categories within the shop-in-shops.

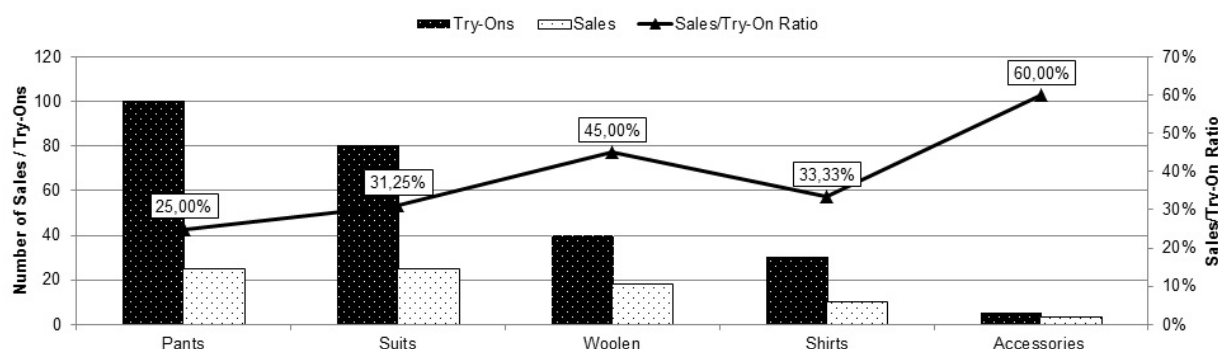


Figure 16: Complements and substitutes on the category level

Therefore articles with a higher try-on demand could be positioned closer to the fitting rooms or the other way around if cross-/up-selling shall be increased. Articles that are hardly tried on could also be positioned closer to the checkout in order to increase impulse buying. It is expected that a demand driven placement of products on the sales floor increases the try-on frequency. And since every try-on is a sales opportunity, it is also expected to lead to an increase in sales.

6.5.4 Correlation of Try-Ons and Sales (Item-Level)

With the means of RFID's unique identifiers, the previously described analysis of the correlation of try-ons and sales can also be conducted on the item-level. On this level of detail, it provides the retailer with information about individual defects of items. Items that are tried on less than average, as compared to the same article group (e.g.,

brown corduroy pants of brand A) might have visible defects or might be dirty. As a consequence, customers do not try them on in the first place. Or items that are tried on more often than the average and not bought can give indications about non-visible defects, e.g., defective zippers or buttons. The retailer can then remove these items from the sales floor and replace them.

6.6 Procedures, Analyses and Underlying Assumptions

Table 8 provides an overview of the utilized procedures, the underlying assumptions, the used countermeasures (or justifications of the assumptions), as well as the thereby affected analyses.

Table 8: Transforming EPC events to trace histories

Procedure	Assumption	Countermeasure / Justification	Affected Analysis
5.4. Trace Histories	- Infrastructure has a read rate of 100% at all locations = all trace histories are complete	- Selection of complete traces only - No sales events occurred during times when inventory = 0 (e.g., Esprit NOS suits) - Sales events of items that had only back store inventory can be assigned to staff sales - From sales floor staff	- Out of Shelf but in Store (Shelf Replenishment) - Out of Stock – Store Replenishment - Lead Time Analysis – Process Inefficiencies - Front Store / Back Store Movements
5.5. Misplacement Aggregation	- Misplacements are Non-Gardeur products on the Smart Shelf	- From sales floor staff	- Misplaced Merchandise
5.6. Try-Ons	- Customers do not stay longer than 30 minutes in the fitting room - Error cases I-IV have been successfully filtered	- From sales floor staff - From results of data analyses and their evaluation - Manual data screening	- Fitting Room Usage Distribution - Catchment Area of Fitting Room Clusters - Time-Dependent Visits of Fitting Room Clusters
5.7. Merchandise Aggregation	- See above	- See above	- Complements and Substitutes - Correlations of Try-Ons and Merchandise Attributes
5.8. Fitting Room Aggregation	- Customers pay their tried-on item(s) 30 minutes after they have left the fitting room	- From the data (plotted) - From sales floor staff	- Correlation of Try-Ons and Sales (Category-level) - Correlation of Try-Ons and Sales (Item-level)

7 Conclusions

Despite the successive uptake of RFID, little research has been conducted so far to identify the necessary procedures to leverage the value of the RFID data streams that are being generated. As a consequence, practitioners and researchers alike still face a variety of fundamental challenges in the pursuit of leveraging the true economic value of massive RFID data sets. RFID is evolving from a base technology for the automatic identification of objects into a rich source of fine-granular information on physical events in the real world, on which several data-driven applications for internal operations, management decisions, and customer services can build upon. However, the availability of this novel information source comes with a variety of challenges induced not only by the rapid growth of data quantities, but also by the suboptimal quality of raw data and the heterogeneity of related data in other internal and external information systems. Models and methods are hence needed that make this complexity manageable even by less IT-intensive organizations, who strive to translate raw data into business value.

The purpose of the present study was to present procedures developed under real-world conditions for the treatment of RFID data that other researchers as well as practitioners can build upon. Our case example from apparel retail has shown that manifold insights can be gained from RFID data, starting from the basic information on the performance of the data collection system itself to several management-related reports on inventories, store processes, and the behavior of customers. The presented performance indicators and visualizations show the potential of RFID as a tool for better informing retail managers, who then make decisions in a different way. For example, store layout decisions might be influenced by detailed analyses of product and customer movements on

the sales floor. Store managers may decide differently on reordering from the distribution centre if they had more precise information on inventories, misplacements, and shrinkage in their store at a certain point in time. Knowledge about the ratio between try-ons and sales might change purchasing decisions. Similarly, suppliers could use the very same information in order to react faster to changes in customer demand and optimize their assortments accordingly. On an operational level, sales staff could benefit from more detailed information on the location and availability of products in order to offer better advisory services to their customer.

The present study offers opportunities for further research in several directions. First of all, we see great potential in the development of analytical methods and models using RFID and related technologies. On the one hand, the noisy nature of RFID data streams may be an appropriate field of application for advanced analytical techniques from data mining and machine learning research, which may be used to detect inaccurate or missing tag reads (i.e., false-negatives and false-positives). On the other hand, further procedures for the evaluation of the collected data from a managerial perspective will be needed, which go beyond the apparel retail industry considered in our study. A second opportunity lies in the processing of not only identification data but also sensor and location data. A number of technology providers have started to combine RFID with sensor functionality or use RFID for building real-time location systems, which generate data types that go beyond traditional RFID systems. The utilization of such data in supply chain management and the corresponding decision models is still a rather young research area. A third opportunity refers to the privacy issues that may arise from the use of RFID for the observation of consumer behavior in stores. Even if customers themselves are not identified via RFID, the extensive collection and combination of data from RFID-equipped goods, customer loyalty cards, or POS systems may pose a potential risk to the privacy rights of individuals. Privacy impact research will hence be necessary to evaluate such risks and proposes strategies for mitigating them and avoiding negative customer reactions. Not least, we see much potential in research that not only proposes ways for better informing managers with the help of RFID data and related technologies, but also investigates how such information is ultimately used. While the value of the processing techniques proposed in our study may seem intuitively clear even to laypeople, the way such information influences managerial decision-making processes is still unknown. In particular, it is not known to what extent better information actually leads to better decisions, which makes any attempt to quantify the business value of RFID data extremely complex. Therefore, empirical research that investigates the impact of RFID on management processes will be necessary to better justify the substantial investments in RFID systems in retail and other industries.

References

- [1] H. Baars, H.-G. Kemper, H. Lasi, and M. Siegel, Combining RFID technology and business intelligence for supply chain optimization – Scenarios for retail logistics, in Proceedings of the 41st Hawaii International Conference on System Sciences, HI, USA, 2008, pp. 73.
- [2] H. Baars and X. Sun, Multidimensional analysis of RFID data in logistics, in Proceedings of the 42nd Hawaii International Conference on System Sciences, HI, USA, 2009, pp. 1-10.
- [3] Y. Bai, F. Wang, and P. Liu, Efficiently filtering RFID data streams, in Proceedings of the 1st International VLDB Workshop on Clean Databases, CleanDB, Seoul, Korea, 2006, pp. 50-57.
- [4] M. Bhattacharya, C.-H. Chu, and T. Mullen, A comparative analysis of RFID adoption in retail and manufacturing sectors, presented at the IEEE International Conference on RFID The Venetian, Las Vegas, Nevada, USA, April 16-18, 2008, pp. 241-249.
- [5] E. Bottani, Reengineering, simulation and data analysis of an RFID system, Journal of Theoretical and Applied Electronic Commerce Research, vol. 3, no. 1, pp. 13-29, 2008.
- [6] E. Bottani, M. Bertolini, R. Montanari, and A. Volpi, RFID enabled business intelligence modules for supply chain optimization, International Journal of RF Technologies: Research and Applications, vol. 1, no. 4, pp. 253-278, 2009.
- [7] S. M. Broniarczyk, W. D. Hoyer, and L. McAlister, Consumers perceptions of the assortment offered in a grocery category: The impact of item reduction, Journal of Marketing Research, vol. 35, no. 2, pp. 166-176, 1998.
- [8] S. M. Broniarczyk and W. D. Hoyer, Retail assortment: More ≠ better, in Retailing in the 21st Century. Current and Future Trends (M. Krafft and M. K. Mantrala, Eds.). Berlin: Springer, 2006, pp. 225-238.
- [9] S. A. Brown, Revolution at the Checkout Counter: The Explosion of the Bar Code. Cambridge, MA: Harvard University Press, 1997.
- [10] Business Week, The supermarket scanner that failed, Business Week, March 22, pp. 52B-52E, 1976.
- [11] G. Chappell, D. Durdan, G. Gilbert, L. Ginsburg, J. Smith, and J. Tobolski. (2003, February) Auto-ID in the box: The value of auto-ID technology in retail stores. Accenture. [Online]. Available: <http://www.autoidlabs.org/uploads/media/ACN-AUTOID-BC006.pdf>.
- [12] M. Christopher, R. Lawson, and H. Peck, Creating agile supply chains in the fashion industry, International Journal of Retail & Distribution Management, vol. 32, no. 8, pp. 367-376, 2004.
- [13] C. Condea, F. Thiesse, and E. Fleisch, RFID-enabled shelf replenishment with backroom monitoring in retail stores, Decision Support Systems, vol. 52, no. 4, pp. 839-849, 2012.
- [14] D. Corsten and T. Gruen, On shelf availability: An examination of the extent, the causes, and the efforts to address retail out-of-stocks, in Consumer Driven Electronic Transformation: Applying New Technologies to Enthuse Consumers and Transform the Supply Chain (G. J. Doukidis and A. P. Vrechopoulos, Eds.). Berlin: Springer, 2005, pp. 131-149.

- [15] D. Delen, B. C. Hardgrave, and R. Sharda, RFID for better supply-chain management through enhanced information visibility, *Production and Operations Management*, vol. 16, no. 5, pp. 613-624, 2007.
- [16] X. Dreze, S. Hoch, and M. Purk, Shelf management and space elasticity, *Journal of Retailing*, vol. 70, no. 2, pp. 301-326, 1994.
- [17] J. Fenn. (2009, August) Inside the hype cycle: What's hot and what's not in 2009. Gartner. [Online]. Available: http://my.gartner.com/it/content/1101800/1101817/august12_hype_cycle_final_jfenn.pdf.
- [18] C. Floerkemeier and M. Lampe, RFID middleware design - addressing application requirements and RFID constraints, in *Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence*, Grenoble, France, 2005, pp. 219-224.
- [19] J. L. Gagnon and J. J. Chu, Retail in 2010: A world of extremes, *Strategy & Leadership*, vol. 33, no. 5, pp. 13-23, 2005.
- [20] C. Goebel, H. Krasnova, H. Syllwasschy, and O. Günther, Vertical integration and information sharing - An empirical investigation in the German apparel industry, in *Proceedings of 9th Internationale Tagung Wirtschaftsinformatik*, Vienna, Austria, 2009, pp. 399-408.
- [21] H. Gonzalez, J. Han, X. Li, and D. Klabjan, Warehousing and analyzing massive RFID data sets, in *Proceedings of the 22nd International Conference on Data Engineering*, Atlanta, Georgia, USA, 2006, pp. 83-92.
- [22] A. Haberman, 17 Billion reasons to say thanks, in *Twenty-Five Years Behind Bars* (A. Haberman, ed.). Cambridge, MA: Harvard University Press, 2001, pp. 113-151.
- [23] B. Hardgrave, M. Waller, and R. Miller. (2006, June) RFID's impact on out of stocks: A sales velocity analysis. Information Technology Research Institute, Sam M. Walton College of Business, University of Arkansas. [Online]. Available: <http://rfid.uark.edu/papers/ITRI-WP068-0606.pdf>.
- [24] ICMR. (2012, April) Store layout, design and visual merchandising. [Online]. Available: <http://www.icmrindia.org/courseware/Retail%20Management/Store%20Layout-Design-Visual%20Merchandis.htm>.
- [25] I. Lee and B.-C. Lee, An investment evaluation of supply chain RFID technologies: A normative modeling approach, *International Journal of Production Economics*, vol. 125, no. 2, pp. 313-323, 2010.
- [26] M. Levy and B. Weitz, *Retailing Management*. Boston: McGraw-Hill/Irwin, 2008.
- [27] H. Liu, J. Li, J. Zhang, and H. Deng, RFID data management in manufactory, in *Proceedings of the 2nd International Conference on Anti-Counterfeiting, Security and Identification*, Guiyang, China, 2008, pp. 420-423.
- [28] C. Loebbecke and J. W. Palmer, A real-world pilot of RFID in the fashion industry: Kaufhof and Gerry Weber in Germany, in *Proceedings of the Academy of Management Annual Meeting*, Atlanta, USA, 2006, pp.1-15.
- [29] V. Manthou and M. Vlachopoulou, Bar-code technology for inventory and marketing management systems: A model for its development and implementation, *International Journal of Production Economics*, vol. 71, no. 1-3, pp. 157-164, 2001.
- [30] A. McAfee, A. Sjomán, and V. Dessain, *Zara: IT for Fast Fashion*. Cambridge, MA: Harvard Business Publishing, 2004.
- [31] K. L. Moon and E. W. T. Ngai, The adoption of RFID in fashion retailing: A business value-added framework, *Industrial Management & Data Systems*, vol. 108, no. 5, pp. 596-612, 2008.
- [32] J. E. Nelson, Scanning's silver celebration, in *Twenty-Five Years Behind Bars* (A. Haberman, ed.). Cambridge, MA: Harvard University Press, 2001, pp. 25-33.
- [33] E. W. T. Ngai, K. K. L. Moon, J. N. K. Liu, K. F. Tsang, and R. Law, Extending CRM in the retail industry: An RFID-based personal shopping assistant system, *Communications of the Association for Information Systems*, vol. 23, no. 1, pp. 277-295, 2008.
- [34] E. W. T. Ngai, K. K. L. Moon, F. J. Riggins, and C. Y. Yi, RFID research: An academic literature review (1995-2005) and future research directions, *International Journal of Production Economics*, vol. 112, no. 2, pp. 510-520, 2008.
- [35] Y. Rezik, E. Sahin, and Y. Dallery, Analysis of the impact of the RFID technology on reducing product misplacement errors at retail stores, *International Journal of Production Economics*, vol. 112, no. 1, pp. 264-278, 2008.
- [36] C. Samli, *Up Against the Retail Giants: Targeting Weakness, Gaining an Edge*. Mason, USA: Thomson South-Western, 2004.
- [37] A. Sarac, N. Absi, and S. Dauzère-Pères, A literature review on the impact of RFID technologies on supply chain management, *International Journal of Production Economics*, vol. 128, no. 1, pp. 77-95, 2010.
- [38] C. Sellitto, S. Burgess, and P. Hawking, Information quality attributes associated with RFID-derived benefits in the retail supply Chain, *International Journal of Retail & Distribution Management*, vol. 35, no. 1, pp. 69-87, 2007.
- [39] M. Tajima, Strategic value of RFID in supply chain management, *Journal of Purchasing & Supply Management*, vol. 13, no. 4, pp. 261-273, 2007.
- [40] F. Thiesse, J. Al-Kassab, and E. Fleisch, Understanding the value of integrated RFID systems: A case study from apparel retail, *European Journal of Information Systems*, vol. 18, no. 6, pp. 592-614, 2009.
- [41] F. Thiesse, T. Staake, P. Schmitt, and E. Fleisch, The rise of the 'next-generation barcode', *Supply Chain Management: An International Journal*, vol. 16, no. 5, pp. 328-345, 2011.
- [42] Z. Ton and A. Raman, *The Effect of Product Variety and Inventory Levels on Misplaced Products at Retail Stores: A Longitudinal Study*. Cambridge, MA: Harvard Business School, 2004.
- [43] Y. Tu and S. Piramuthu, A decision support model for filtering RFID read data, in *Proceedings of the 16th International Conference on Advanced Computing and Communications*, Chennai, India, 2008, pp. 221-224.

- [44] F. Uhrich, U. Sander, F. Resatsch, J. M. Leimeister, and H. Krcmar, RFID in retailing and customer relationship management, *Communications of the Association for Information Systems*, vol. 23, no. 1, pp. 219-234, 2008.
- [45] S. F. Wamba, L. A. Lefebvre, Y. Bendavid, and E. Lefebvre, Exploring the impact of RFID technology and the EPC network on mobile B2B eCommerce: A case study in the retail industry, *International Journal of Production Economics*, vol. 112, no. 2, pp. 614-629, 2008.
- [46] F. Wang and P. Liu, Temporal management of RFID data, in *Proceedings of the 31st International Conference on Very Large Data Bases*, Trondheim, Norway, 2005, pp. 1128-1139.
- [47] D. C. Wyld, RFID 101: The next big thing for management, *Management Research News*, vol. 29, no. 4, pp. 154-173, 2006.
- [48] P. J. Zelbst, K. W. Green, V. E. Sower, and P. M. Reyes, Impact of RFID on manufacturing effectiveness and efficiency, *International Journal of Operations & Production Management*, vol. 32, no. 3, pp. 329-350, 2012.

Appendix A: Script 1

Line	Script	Description
1:	Function misplacement_aggregation(RFIDeventTable c1)	
2:	begin	
3:	//c1: table with RFID read events	
4:	for row1 in c1 loop	
5:	if (epc_old=row1.epc and row1.eventTime < eventTime_old +20/1440) then	<i>Has the item already been seen within the past 20 minutes?</i>
6:	counter := counter + 1;	<i>Yes → increase counter</i>
7:	else	
8:	insert into misplaced_articles_sm1 values(epc_old, ean_old, department_old, articlegroup_old, start, eventtime_old, counter*15);	<i>No → store EPC in the list of misplaced items (including the number of detections)</i>
9:	counter := 0;	
10:	start := row1."eventTime";	
11:	end if;	
12:	eventTime_old := row1."eventTime";	<i>Reset tag information</i>
13:	epc_old := row1.epc;	
14:	ean_old := row1.ean;	
15:	department_old := row1.department;	
16:	articlegroup_old := row1.articlegroup;	
17:	end for;	
18:	end;	

Appendix B: Script 2

Line	Script	Underlying Objective
1:	Function filtering_script(RFIDfitting_room_read_events <i>table1</i>)	
2:	begin	
3:	for row1 in <i>table1</i> loop	
4:	if (current EPC number same as proceeding EPC number) then	
5:	if (eventTime of next read event <= current eventTime + 10.5min) then	<i>Has this item already been seen within the following 10.5. minutes?</i>
6:	skip	<i>Yes → do nothing</i>
7:	else if	
8:	insert line into table filtered_fitting_room_read_events_table;	<i>No → add item to the list of detected items</i>
9:	end if;	
10:	end for;	

Appendix C: Script 3

Line	Script	Underlying Objective
1:	Function aggregation_script(RFIDfitting_room_read_events <i>table1</i>)	
2:	begin	
3:	for row1 in table1 loop	
4:	if (current EPC number same as proceeding EPC number) then	
5:	if (difference of “last read event” and “first read event of that day” <=30 min) then	<i>Has this item already been seen within the past 30 minutes?</i>
6:	skip	<i>Yes → do nothing</i>
7:	else if	
8:	insert line into table filtered_fitting_room_read_events_table;	<i>No → add item to the list of detected items</i>
9:	end if;	
10:	end for;	
11:	end;	

Appendix D: Script 4

Line	Script	Underlying Objective
1:	Function	
	merchandise_aggregation_script(RFIDfitting_room_read_events <i>table1</i>)	
2:	begin	
3:	// ts: timestamp variable	
4:	// counter: a counter variable	
5:	// date: a date variable	
6:	ts:=""	<i>timestamp variable</i>
7:	counter:=0;	<i>customer count variable</i>
8:	date:=""	<i>date variable</i>
9:	for row1 in table1 loop	
10:	counter:=0;	
11:	ts:=null;	
12:	select to_number(count(ean)), min(time) into counter, ts from rf_ps where ean=row1.ean and time between row1."eventTime" and row1."eventTime"+35/1440;	<i>Select all detections within the following 35 minutes</i>
13:	if (counter = 0) then	
14:	ts:=null;	
15:	insert into table (values, counter, ts);	
16:	else	
17:	counter := count + 1;	<i>Increase counter</i>
18:	insert into table (values, counter, ts);	<i>Insert item to room visit table</i>
19:	end if ;	
20:	end for ;	
21:	end ;	