Bus Inf Syst Eng 61(1):51–66 (2019)
https://doi.org/10.1007/s12599-018-0566-9

RESEARCH PAPER

Towards Digital Transformation in Fashion Retailing: A Design-Oriented IS Research Study of Automated Checkout Systems

Matthias Hauser · Sebastian A. Günther · Christoph M. Flath · Frédéric Thiesse

Received: 5 March 2018 / Accepted: 21 September 2018 / Published online: 26 November 2018

Abstract Automated checkout systems promise greater sales due to an improved customer experience and cost savings because less store personnel is needed. The present design-oriented IS research study is concerned with an automated checkout solution in fashion retail stores. The implementation of such a cyberphysical system in established retail environments is challenging as architectural constraints, well-established customer processes, and customer expectations regarding privacy and convenience impose limits on system design. To overcome these challenges, the authors design an IT artifact that leverages an RFID sensor infrastructure and software components (data processing and prediction routines) to jointly address the central problems of detecting purchases in a reliable and timely fashion and assigning these purchases to individual shopping baskets. The system is implemented and evaluated in a research laboratory under real-world conditions. The evaluation indicates that shopping baskets can indeed be detected reliably (precision and recall rates greater than 99%) and in an expeditious manner (median detection time of 1.03 s). Moreover, purchase assignment reliability is 100% for most standard scenarios but falls to 42% in the most challenging scenario.

Keywords Design-oriented IS research · Digital innovation · Internet of things · Cyberphysical systems · Retail industry · Radio frequency identification · Machine learning · Automated checkout systems

Accepted after two revisions by the editors of the special issue.

M. Hauser (✉) · Prof. Dr. F. Thiesse
University of Würzburg, Josef-Stangl-Platz 2, 97070 Würzburg, Germany
e-mail: matthias.hauser@uni-wuerzburg.de

Prof. Dr. F. Thiesse
e-mail: frederic.thiesse@uni-wuerzburg.de

S. A. Günther
University of Bamberg, Kapuzinerstraße 16, 96047 Bamberg, Germany
e-mail: sebastian.guenther@uni-bamberg.de

Prof. Dr. C. M. Flath
University of Würzburg, Sanderring 2, 97070 Würzburg, Germany
e-mail: christoph.flath@uni-wuerzburg.de

1 Introduction

Digital innovations manifest themselves in the transformation of processes, content or objects from the physical realm to the digital sphere (Fichman et al. 2014; Yoo et al. 2010). A particularly interesting form of digital innovation is the emerging class of cyberphysical systems, which are expected to greatly enhance the efficiency, functionality, and reliability of previously non-digitized systems (National Science Foundation 2010). Such systems, having progressed beyond speculative visions and early pilot implementations, create previously infeasible processes and establish new business models across various economic sectors (Borgia 2014; Stankovic 2014). In manufacturing, industrial internet applications are increasingly turning shopfloors into smart factories (Lasi et al. 2014; Lee 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 OEMs a run for their money by replacing individually owned conventional cars with fleets of shared, autonomous vehicles (The Economist 2016). Smart grids are reversing the accustomed supply-follows-demand paradigm of power systems to enable a greener and more reliable electricity supply (Amin and Wollenberg 2005; Blumsack and Fernandez 2012; Farhangi 2010). 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 (e.g., automated checkout systems, personal shopping assistants, omnichannel services) 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).

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” and “one of the most powerful tools [they] have in being able to learn about what [their] customers need” (Morrell 2015). The importance of digital innovation in retail is often attributed to the strong competition between traditional brick and mortar stores and online players (Brynjolfsson et al. 2013; Herhausen et al. 2015; Rigby 2011). In this context, competitive pressure on traditional retailers is not only exerted by price, but also by new digital service offerings that have altered customer relationships, customer behavior, and their expectations regarding retail service quality (Grewal et al. 2017; PwC 2015). Cyberphysical systems can help traditional retailers to meet these challenges by providing them with the means to simultaneously increase cost-efficiency and the attractiveness of physical stores (Gregory 2015; Inman and Nikolova 2017; Kourouthanassis and Roussos 2003; Piotrowicz and Cuthbertson 2014). McKinsey projects the economic potential of cyberphysical systems in stationary retail environments to exceed $410 billion per year in 2025 (Manyika et al. 2015).

Recently, various cyberphysical systems in retail stores have been conceptualized. Smart kiosks, for example, allow customers to browse product offerings or order products that are currently unavailable in the store (Herhausen et al. 2015; Shankar et al. 2011). Smart fitting rooms offer additional services (e.g., product recommendations or omnichannel services) based on a customer’s garment selection displaying information on a screen within the cabins (Parada et al. 2015; Senecal and Nantel 2004; Wong et al. 2012). With an economic potential of more than $150 billion per year in 2025 (Manyika et al. 2015), automated checkout systems have emerged as the most significant opportunity among cyberphysical retail systems. Against this backdrop, the present study describes a design-oriented IS research project concerned with the implementation and evaluation of an automated checkout system. Thereby, we seek to expand the existing knowledge base concerning the creation of smart retail environments, which are an ideal use case for the implementation of cyberphysical systems. Within the retail sector, we focus on fashion retailing, which is a sizable sub-segment characterized by high margins and a recent, drastic shift towards innovative, adaptable players (Amed et al. 2018).

Our research seeks to address the two main tasks of reliably and instantaneously detecting products (i.e., garments) and correctly assigning them to individual shopping baskets. Reliable detection is decisive for automated checkout systems because undetected products cause revenue losses for the retailer (Kang and Gershwin 2005). Incorrectly assigning particular items to shopping baskets, on the other hand, results in customer dissatisfaction and interruptions of in-store operations (Hayes and Blackwood 2006). The design of cyberphysical systems is generally considered challenging because their components have to be seamlessly integrated into existing real-world environments (Baheti and Gill 2011; Böhmann et al. 2014; Brandt et al. 2017; Khaitan and McCalley 2015; Kourouthanassis and Roussos 2003). This is particularly problematic in fashion retail environments, which are characterized by a prevalence of immutable physical system components (e.g., architectural constraints, lack of space) and immutable non-physical system components (e.g., established customer behavior patterns, unpredictable customer behavior).

Design-oriented IS research seeks to develop innovative artifacts with a strong focus on utility in practice (Peffers et al. 2018). In keeping with this research practice, our study makes a twofold contribution. First, we introduce an innovative IT artifact that offers clear benefits for retail companies (i.e, a reduction in cashier staff requirements) and their customers (i.e., the elimination of checkout queue times) in an increasingly relevant and widespread field of application (Amed et al. 2018; Manyika et al. 2015). Second, beyond the specific use case, our research demonstrates how machine learning approaches can help mitigate the problem of immutability of the environment. The latter findings, in particular, may be generalized and applied to the design of other cyberphysical systems.

2 Practical Background

Traditional clerk-based checkout systems are labor-intensive and can be a great source of frustration for customers having to wait in line (Manyika et al. 2015). To reduce costs, retailers have started adopting self-service technologies that enable shoppers to detect, 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

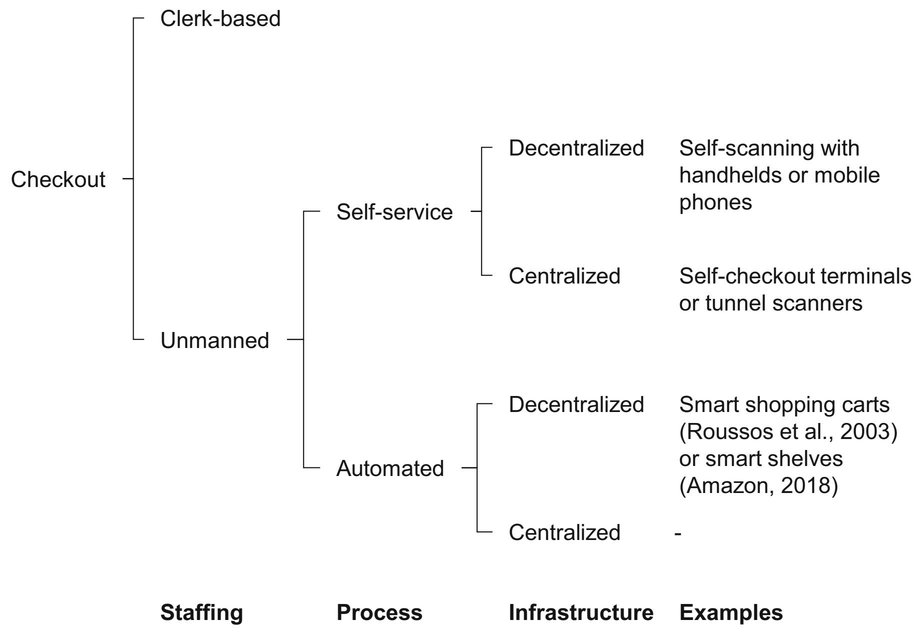

Fig. 1 Differentiation of checkout systems

traditional checkout process with respect to the customer experience, potentially creating new challenges as many customers consider the service frustrating, irritating, and alienating (Meuter et al. 2000).[1]

Self-service checkout systems can be roughly categorized into (1) centralized systems at store exits and (2) 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° 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 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 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').

The literature on automated checkout systems is sparse. To the best of our knowledge, only two systems from the literature address the aforementioned challenges. The first system (MyGrocer) relies on shopping carts equipped with RFID readers that 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 detect products taken from or returned to shelves, keep track of the products chosen by customers in virtual shopping carts, and

[1] 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, and customer-based failures (e.g., forgetting one's personal identification number).

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 regarding the Amazon Go system suggests that it stores the inventory locations of all products within stores and mainly relies on cameras to detect products that customers take from or return to particular inventory locations.[2] In addition to the cameras, additional sensors (e.g., pressure sensors, infrared sensors, light curtains, and RFID readers) and customer information (e.g., purchase history) can be utilized to identify and assign purchases.

3 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.

3.1 Requirements Analysis

The present study was conducted in the course of a research project undertaken by multiple research institutions and two leading European fashion retailers. 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 retailing, 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). Lastly, 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 of fashion store environments).
2. *Customers in fashion retail stores usually leave unwanted garments in the changing room* We consider this to be another immutable business process 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 et al. 2017; Litfin and Wolfram 2006).
4. *Major fashion retailers have implemented item-level RFID tagging of products*[3] Fashion retailers and suppliers first adopted RFID at case-level mainly for inventory management purposes (Hardgrave et al. 2013). Item-level tagging has, however, moved out of the research environment and into mainstream commerce (Barthel et al. 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 Sect. 2. The first observation rules out automated checkout systems based on smart shopping carts (i.e., systems such as MyGrocer). The second observation rules out automated checkout systems that rely on shelf activity to track purchases (i.e., systems such as Amazon Go).

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

[2] 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 et al. 2015) provide insights into the implementation of this cyberphysical retail system.

[3] RFID identifies products at the item level without a direct line of sight. Furthermore, it facilitates the simultaneous bulk detection of multiple objects.

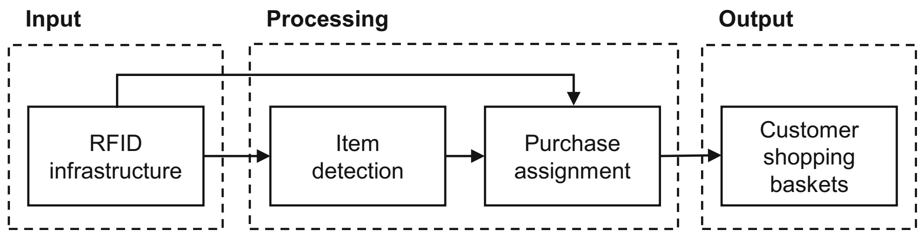

Fig. 2 Architecture of the automated checkout artifact

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: For example, Chaves et al. (2010) present a model for the automatic detection of misplaced garments in retail stores. Parada et al. (2015) present a system that detects products taken from smart shelves based on the analysis of low-level RFID data. Similarly, Li et al. (2015) introduce a system able to distinguish between different touch events (e.g., browsing through RFID-tagged garments, selection of garment of interest).

3.2 Research Methodology

We aim at creating an artifact that reliably and instantaneously detects items that are leaving the store and correctly assigns them to individual shopping baskets. Our design process follows the guidelines put forward by Hevner et al. (2004):

- *Problem Relevance* There are many possible applications for automatic detection systems. The gross economic potential of automated checkout systems is projected to exceed \$150 billion per year in 2025 (Manyika et al. 2015). Adoption reduces waiting times and thus increases customer shopping satisfaction, as well as cutting costs. Systems described in the literature can not be applied in fashion retail environments because they were developed for supermarket settings which differ significantly from fashion retail environments with respect to in-store processes and the suitability of specific technologies.
- *Design as an Artifact* The proposed automated checkout artifact combines hardware (RFID readers and antennas) and software components (data processing and prediction routines) to ensure (1) the reliable and timely detection of items and (2) the correct assignment of these items to shopping baskets.
- *Design Evaluation* We evaluate the artifact using a comprehensive experimental study in the laboratory. Our setup takes into account the limited process control in fashion retail stores by considering, for example, multiple typical customer movement patterns, different numbers of people, and different movement speeds.
- *Research Contribution* Our research contributes to the understanding of the design of cyberphysical systems and provides prescriptive knowledge regarding the design of automated checkout systems. In addition, our research demonstrates how machine learning approaches can help mitigate the problem of environmental immutability.
- *Research Rigor* Our software components leverage state-of-the-art supervised and unsupervised machine learning techniques to implement a reliable automated detection system. By relying on separate training and test data sets, our artifact evaluation incorporates best practices established in data science.
- *Design as a Search Process* Our design artifact is based on existing models and research contributions (e.g., Hauser et al. 2015; Keller et al. 2014; Ma et al. 2018). Moreover, the findings may be generalized and applied to the design of other cyberphysical systems in retail environments and beyond.
- *Research Communication* Our research informs both technical and managerial audiences. The data mining models primarily address audiences with a more technical focus. In addition, we want to encourage decision-makers to leverage the potential of low-level RFID data with data analytics techniques.

3.3 System Architecture and Infrastructure

The architecture of our automated checkout artifact combines hardware and software components (see Fig. 2). The hardware consists of two RFID reader installations, a ceiling-mounted system that helps track items, 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; the second assigns items leaving the store (identified by the first component) to individual shopping baskets. These shopping baskets are the output of the artifact.

Figure 3 depicts the infrastructure with the two parallel RFID readers from Impinj, a manufacturer of RFID devices and software. The gate-mounted system features four far-field antennas (Impinj Inc. 2017a), while the ceiling-

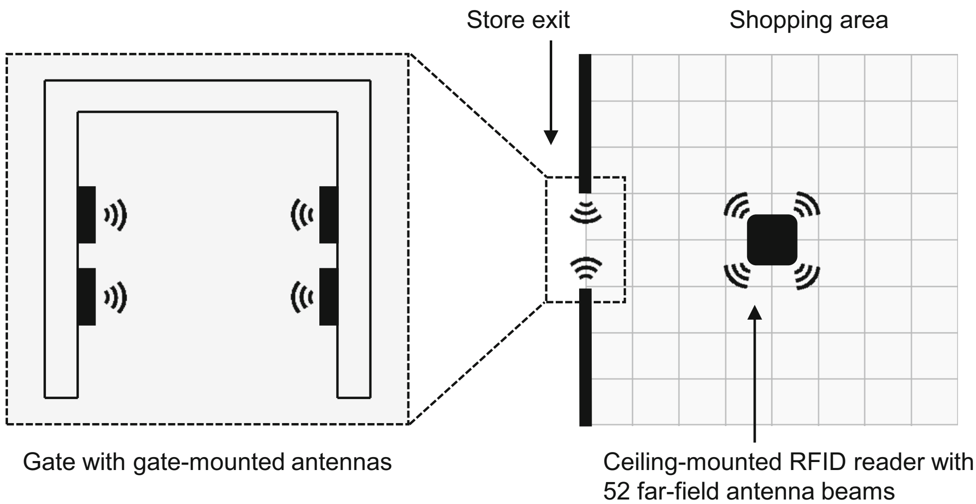

Fig. 3 Infrastructure with two parallel RFID reader installations

mounted system boasts an array of 52 far-field antenna beams mounted in one housing (Impinj Inc. 2017b).

3.4 Item Detection Approach

The item detection software component has to reliably distinguish between items that pass through a transition area and others (e.g., static items near the RFID reader). 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. Advanced data analytics techniques provide an avenue by which to reduce the shortcomings of solely hardware-based RFID solutions with respect to detection quality. To apply machine learning algorithms, the RFID data streams first need to be split into chunks to enable continuous evaluation in real time. 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.

3.4.1 Data Preprocessing

Table 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 readers' antennas. Here, *EPC* stands for Electronic Product Code and is the unique identifier of the RFID tag, *RSSI* stands for Radio Signal Strength Indication indicates the radio signal's power, *phase angle* is the current state of the backscattered sinusoidal wave, and *antenna* is the ID of the antenna that read the tag.

Prior research has usually considered aggregates for single runs and the classification is thus performed after a tag has moved through a transition area (e.g., Hauser et al. 2015; Keller et al. 2014; Ma et al. 2018). In contrast, 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. Similar to Parlak and Marsic (2013), we 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 2 s) that is updated at regular time intervals (Jeffery et al. 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 2 s offer sufficient information to reliably classify the events. To facilitate real-time evaluation, we apply window shifts every 250 ms.

Table 1 Representative low-level RFID data excerpt

Reader	EPC	Timestamp	Antenna	RSSI	Phase angle
Ceiling	3032…7D	1,453,989,765.31	15	− 59.0	3.50
Ceiling	3032…D1	1,453,989,765.31	15	− 56.0	2.91
Gate	3032…7D	1,453,989,765.34	4	− 69.0	2.72
Ceiling	3032…7D	1,453,989,765.34	17	− 56.0	3.07

3.4.2 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. Several authors stress the fact that feature generation is a key phase of any data mining project (Domingos 2012; Halevy et al. 2009). 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.

Prior research leveraging data analytics techniques for the improvement of RFID-based transition detection systems has almost exclusively focused on systems in controlled environments such as production or logistics facilities (e.g., Buffi et al. 2017; Keller et al. 2014; Ma et al. 2018). In these environments, companies can instruct their employees how to behave in the proximity of RFID readers, which is clearly not possible when dealing with customers. For this reason, we focus on the development of features that facilitate the reliable identification of multiple moving objects. We engineered 184 different features for training of the classification models. 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 et al. 2014; Ma et al. 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. 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. A complete list of the features considered in our classification models is provided in "Appendix" (available online via https://www.springerlink.com).

3.4.3 Modelling

We approach the classification problem using a set of standard algorithms: logistic regression (LogReg) (Menard 2018), artificial neural networks (ANN) (Bishop 2006), support vector machines (SVM) (Chang and Lin 2011), and gradient tree boosting (XGBoost) (Chen and Guestrin 2016). We perform hyper-parameter optimization of the classification models considering, for example, numbers of hidden layers and nodes or maximum number of constructed trees (Witten et al. 2016).

Every 250 ms, 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 tags in question have moved through the gate or not. To detect whether an item has moved through the gate, the models have to classify at least one of the associated two-second windows as having moved 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 must not classify any of the two-second windows associated with detections of 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.

3.5 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 in front of the gate 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.

3.5.1 Item Path Determination

We rely on state-of-the-art indoor localization techniques to infer item paths. To this end, we 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). We again consider a sliding window approach with window shifts every 250 ms 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

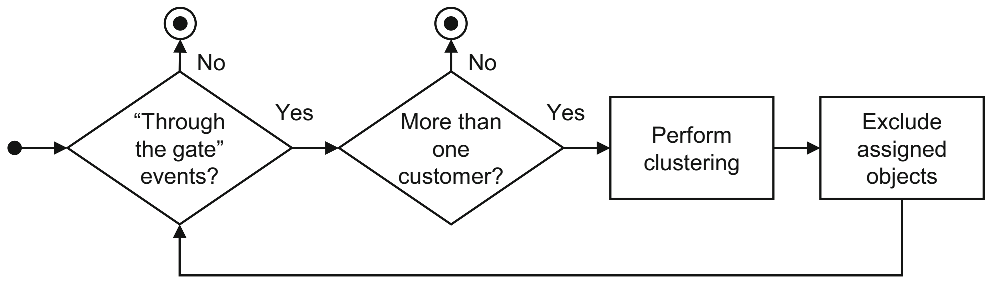

Fig. 4 Visualization of the process for the assignment of products to shopping baskets

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.

We developed 174 features for the training of the classifiers that help localize tags within reading range of the 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 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 this second classification model is again provided in "Appendix".

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 and again perform hyper-parameter optimization. To determine item paths, we concatenate the most probable locations of individual items over time.

3.5.2 Assignment Process

To assign items to individual shopping baskets, the artifact needs to identify the correct customer associated with the items that are currently leaving the store. This task can be tackled by grouping the items within the antennas' reading range (i.e., the shopping area in front of the gate) such that items in the same group are regarded as belonging 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 4 illustrates the assignment process. The process is triggered every time the first software component detects an item being moved through the gate. The assignment component then has to determine all the other items that also belong to this shopping basket. This is achieved by analyzing the paths of all items within the antennas' reading range. 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 shopping basket. Otherwise, we use clustering techniques to determine the items that form a group 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 but exclude items that are already assigned to customer shopping baskets.

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 can be distinguished from neighboring clusters (Rousseeuw 1987).

4 Evaluation

We collected large data sets in a retail research laboratory for instantiation and evaluation of the automated checkout

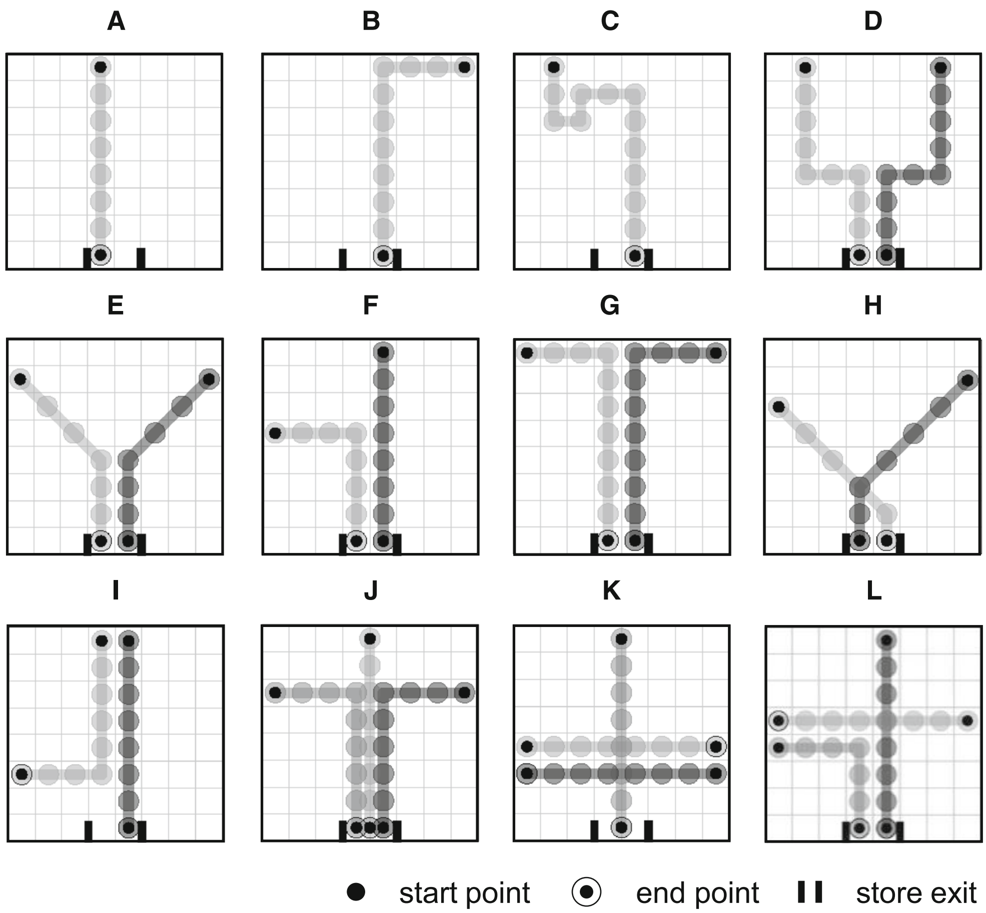

Fig. 5 Test setting with typical customer movement patterns

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 data-mining model (see Sect. 3.5.1).

4.1 Evaluation Setting

We set up an experimental shopping area in a retail research laboratory for the evaluation of the automated checkout artifact. The dimensions of our experimental shopping area were 4.8 m by 4.8 m.[4] For the collection of training data for the indoor localization model, we divided this area into 64 grid fields of equal size.

[4] 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.

The artifact design necessitates the collection of (1) RFID raw data fingerprints at different locations within the shopping area for training and testing of the indoor localization model and (2) 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 2 min of 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.

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). Figure 5 illustrates the

Table 2 Experimental design (numbers in table fields indicate numbers of repetitions per test)

People	Tags	Speed	Movement patterns											
			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

customer movement paths that we considered in our analysis. Error sources that we identified during our experiments are (1) customers with tagged objects who walk in close proximity to the gate and (2) 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. 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. Our experimental design includes 18 tests in total, each of which was repeated 50 times. Table 2 provides a complete overview of the experimental design. The data set comprises 1500 runs with a total of 1,431,347 individual tag readings.

4.2 Evaluation Results

The artifact evaluation is based on the tests with typical movement paths in retail stores (i.e., the second data set). To ensure representative results, we performed fivefold 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.

4.2.1 *Reliability of Detection*

In our tests, 4350 items (1300 customer shopping baskets) were carried through the 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 Fig. 5). We base our evaluation of the model's reliability on the criteria of *balanced accuracy*, *precision*, and *recall*. 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.

Table 3 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

The performance indicators for the four types of classifiers are summarized in Table 3. With the exception of the logistic regression model (LogReg), 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. 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 Fig. 5).

Recall values below 100% at item level (see Table 3) do not necessarily imply that some items might not get assigned to customers' 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 4 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

Table 4 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

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.

4.2.2 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 6 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

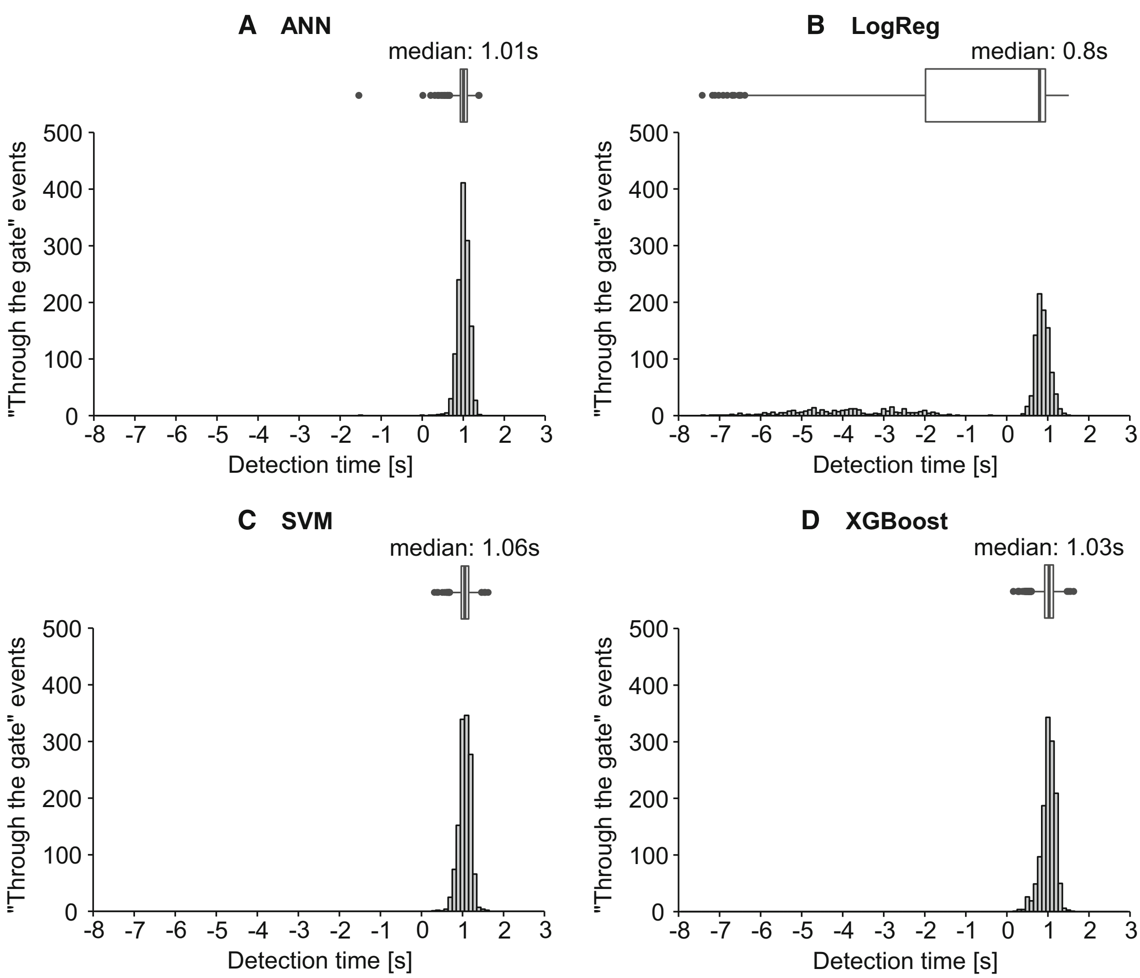

Fig. 6 Detection time histograms and boxplots with 2.5 and 97.5 percentiles

Table 5 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

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.

4.2.3 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 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 movement pattern I and movement pattern 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.

5 Discussion

The present study aimed to design an automated checkout system for fashion retail stores that reliably and instantaneously detects items leaving a store and correctly assigns them to individual shopping baskets. We find that while most purchases were correctly assigned to customers, 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 a situation could easily arise when friends are shopping together, which highlights the limitations of the pilot implementation. To solve this issue, various model improvements could be considered to bolster detection reliability: Probabilistic models may be able to improve the accuracy of item paths (Hauser et al. 2017). Furthermore, the integration of additional data sources can improve the assignment process. One possibility is the integration of information from additional sensor systems or the inclusion of other data sources (e.g., customer purchase history, sales data, and garment characteristics). This approach is in line with Lee (2008), who suggests that in such cases "the next level of abstraction [...] must compensate with robustness." In addition, expanding the monitored area through additional hardware (i.e., the installation of more ceiling-based RFID systems) would make it possible to more accurately distinguish item movement paths.

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. To further boost predictive power, ensemble methods and alternative algorithmic approaches (e.g., deep learning) may help create a more reliable system. 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. Only then can retailers move towards more advanced, customer-oriented smart service offerings. In addition, further tests of the automated checkout system should include consideration of the payment initialization process. This process differs depending on the utilized wireless payment technology. Candidate technologies include Bluetooth Low Energy and Near Field Communication. Leveraging these technologies would require that customers register upon entering a retail store to ensure that they have a compatible device for wireless payment.

Going forward, we want to enhance the generalizability of the proposed automated checkout artifact and extend our approach to form an entire system, i.e., a pervasive retail store, instead of considering individual system components. From the perspective of an entire service system, automated checkout is only a small building block. Future research on the design of cyberphysical systems for retail environments should establish integrated smart environments instead of individual system components. Thereby, the investment in costly technology is not made to augment a single business process, but should rather fuel a transformation of store environments as well as integration in the context of omnichannel retailing.

6 Conclusion

Leading scholars in the field of design science research have called for more research on the actual design of novel and useful artifacts (Baskerville et al. 2018; Peffers et al. 2018). A particularly interesting form of design artifacts are cyberphysical systems, which are expected to greatly enhance previously non-digitized systems by providing "new use that was previously inaccessible" through tight integration and coordination between physical and digital resources (Brandt et al. 2017). Applications of such systems can be found in different areas such as manufacturing (Lasi et al. 2014), personal transportation (The Economist 2016), power delivery (Amin and Wollenberg 2005), healthcare (Lee and Sokolsky 2010), and retail (Kourouthanassis and Roussos 2003). Specific challenges in the design of cyberphysical systems include the consistent, reliable detection and interpretation of events on the physical level, which is critical for the quality and efficiency of the digital services based on them. The design of cyberphysical systems is considered challenging because many of their characteristics cannot be freely designed, but are limited by the environment in which the artifact is to be embedded (Brandt et al. 2017; Khaitan and McCalley 2015).

Automated checkout is a particularly suitable showcase for our design-oriented IS research study, as it features an environment with immutable physical system components (e.g., architectural constraints, lack of space) and immutable non-physical system components (e.g., established customer behavior patterns, unpredictable customer behavior). Ours is the first automated checkout system specifically developed for fashion stores. Existing systems were developed for supermarket settings and are not applicable in the fashion retailing domain because they either (1) rely on shopping carts or baskets, (2) use camera systems (which is problematic in key areas of fashion stores), or (3) 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 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. To this end, we successfully leveraged machine learning techniques to mitigate problems arising from immutable components of the environment in which the system is to be embedded.

Apart from presenting prescriptive knowledge on the design of an innovative IT artifact, our research also provides an example of how data analytics enables the establishment of new internal processes which in turn may result in innovative service offerings. Interestingly, our artifact offers capabilities that can be applied in instances beyond the intended checkout use case. First, automated detection systems that can be implemented in environments with limited process control offer various opportunities for additional use cases in, for example, article surveillance systems or fitting rooms that detect items within them in order to provide customers with additional information. 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 questions such as "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 system are key for the successful introduction of novel cyberphysical systems. Therefore, we conclude that pilot implementations relying on a generic system infrastructure provide businesses with the opportunity to identify related business cases.

Appendix: Feature descriptions

Table 6 describes the features used in the item detection model (see Sect. 3.4); Table 7 the features used in the localization model (see Sect. 3.5.1).

Table 6 Item detection model features

Features	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–F164	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
F165–F166	Mean RSSI measurement of the R420 antennas on the right and on the left gate side
F167–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 (i.e., height of Gaussian curve peak, position of center of peak 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 all other item detection model features, this feature does not only take into account the measurements of a particular tag but the measurements of all tags)

Table 7 Localization 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	Time difference between the first and the last xArray reading
F174	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)

References

Amazon (2018) Amazon go: frequently asked questions. https://www.amazon.com/b?node=16008589011. Accessed 15 Apr 2018

Amed I, Berg A, Kappelmark S, Hedrich S, Andersson J, Drageset M, Young R (2018) The state of fashion 2018. Technical report, The Business of Fashion and McKinsey & Company

Amin SM, Wollenberg BF (2005) Toward a smart grid: power delivery for the 21st century. IEEE Power Energ Mag 3(5):34–41

Baheti R, Gill H (2011) Cyber-physical systems. Impact Control Technol 12:161–166

Barthel R, Hudson-Smith A, de Jode M (2014) Future retail environments. Technical report

Baskerville R, Baiyere A, Gregor S, Hevner A, Rossi M (2018) Design science research contributions: finding a balance between artifact and theory. J Assoc Inf Syst 19(5):358–376

Bishop CM (2006) Pattern recognition and machine learning. Information science and statistics. Springer, Heidelberg

Blumsack S, Fernandez A (2012) Ready or not, here comes the smart grid!. Energy 37(1):61–68

Böhmann T, Leimeister JM, Möslein K (2014) Service systems engineering: a field for future information systems. Res Bus Inf Syst Eng 6(2):73–79

Borgia E (2014) The internet of things vision: key features, applications and open issues. Comput Commun 54:1–31

Brandt T, Feuerriegel S, Neumann D (2017) Modeling interferences in information systems design for cyberphysical systems: insights from a smart grid application. Europ J Inf Syst. 27(2):207–220

Brynjolfsson E, Hu YJ, Rahman MS (2013) Competing in the age of omnichannel retailing. MIT Sloan Manag Rev 54(4):23

Buffi A, D'Andrea E, Lazzerini B, Nepa P (2017) UHF-RFID smart gate: tag action classifier by artificial neural networks. In: 2017 IEEE international conference on RFID technology application (RFID-TA), pp 45–50. https://doi.org/10.1109/RFID-TA.2017.8098900

Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol 2(3):27:1–27:27

Chaves LWF, Buchmann E, Böhm K (2010) Finding misplaced items in retail by clustering RFID data. In: Proceedings of the 13th international conference on extending database technology, ACM, New York, NY, USA, EDBT '10, pp 501–512. https://doi.org/10.1145/1739041.1739102

Chen T, Guestrin C (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

Domingos P (2012) A few useful things to know about machine learning. Commun ACM 55(10):78–87

Farhangi H (2010) The path of the smart grid. IEEE Power Energ Mag 8(1):18–28

Fichman RG, Dos Santos BL, Zheng ZE (2014) Digital innovation as a fundamental and powerful concept in the information systems curriculum. MIS Q 38(2):329–353

Gregory J (2015) The internet of things, revolutionizing the retail industry. Technical report

Grewal D, Roggeveen AL, Nordfält J (2017) The future of retailing. J Retail 93(1):1–6

Halevy A, Norvig P, Pereira F (2009) The unreasonable effectiveness of data. IEEE Intell Syst 24(2):8–12

Hardgrave B, Aloysius J, Goyal S (2013) RFID-enabled visibility and retail inventory record inaccuracy: experiments in the field. Prod Oper Manag 22(4):843–856

Hauser M, Zügner D, Flath CM, Thiesse F (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

Hauser M, Griebel M, Thiesse F (2017) A hidden markov model for distinguishing between RFID-tagged objects in adjacent areas. In: IEEE international conference on RFID

Hayes R, Blackwood R (2006) Evaluating the effects of EAS on product sales and loss: results of a large-scale field experiment. Secur J 19(4):262–276

Herhausen D, Binder J, Schoegel M, Herrmann A (2015) Integrating bricks with clicks: retailer-level and channel-level outcomes of online-offline channel integration. J Retail 91(2):309–325

Hevner AR, March ST, Park J, Ram S (2004) Design science in information systems research. MIS Q 28(1):75–105

Impinj Inc. (2017a) SpeedwayR installation and operations guide version 5.12.0

Impinj Inc. (2017b) xArray/xSpan installation and operations guide version 5.12.0

Inman JJ, Nikolova H (2017) Shopper-facing retail technology: a retailer adoption decision framework incorporating shopper attitudes and privacy concerns. J Retail 93(1):7–28

Jeffery SR, Garofalakis M, Franklin MJ (2006) Adaptive cleaning for RFID data streams. In: Proceedings of the 32nd international conference on very large data bases, VLDB Endowment, pp 163–174

Kang Y, Gershwin SB (2005) Information inaccuracy in inventory systems: stock loss and stockout. IIIE Trans 37(9):843–859

Keller T, Thiesse F, Fleisch E (2014) Classification models for RFID-based real-time detection of process events in the supply chain: an empirical study. ACM Trans Manag Inf Syst 5(4):1–30

Khaitan SK, McCalley JD (2015) Design techniques and applications of cyberphysical systems: a survey. IEEE Syst J 9(2):350–365

Kourouthanassis P, Roussos G (2003) Developing consumer-friendly pervasive retail systems. IEEE Pervasive Comput 2(2):32–39. https://doi.org/10.1109/MPRV.2003.1203751

Kumar D, Kornfield EM, Prater AC, Boyapati S, Ren X, Yuan C (2015) Detecting item interaction and movement. US20150019391A1

Lasi H, Fettke P, Kemper HG, Feld T, Hoffmann M (2014) Industry 4.0. Bus Inf Syst Eng 6(4):239–242

Lee EA (2008) Cyber physical systems: design challenges. In: 2008 11th IEEE international symposium on object oriented real-time distributed computing (ISORC). IEEE, pp 363–369

Lee I, Sokolsky O (2010) Medical cyber physical systems. In: 2010 47th ACM/IEEE design automation conference (DAC). IEEE, pp 743–748

Lee J, Bagheri B, Kao HA (2015) A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manuf Lett 3:18–23

Li H, Ye C, Sample AP (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, Wolfram G (2006) New automated checkout systems. In: Krafft M, Mantrala MK (eds) Retailing in the 21st century: current and future trends. Springer, Berlin, pp 143–157

Liu H, Darabi H, Banerjee P, Liu J (2007) Survey of wireless indoor positioning techniques and systems. IEEE Trans Syst Man Cybern Part C (Appl Rev) 37(6):1067–1080. https://doi.org/10.1109/TSMCC.2007.905750

Ma H, Wang Y, Wang K (2018) Automatic detection of false positive RFID readings using machine learning algorithms. Expert Syst Appl 91:442–451

Manyika J, Chui M, Bisson P, Woetzel J, Dobbs R, Bughin J, Aharon D (2015) The internet of things: mapping the value beyond the hype. Technical report, McKinsey Global Institute

Menard S (2018) Applied logistic regression analysis, vol 106. SAGE, Washington

Meuter ML, Ostrom AL, Roundtree RI, Bitner MJ (2000) Self-service technologies: understanding customer satisfaction with technology-based service encounters. J Market 64(3):50–64

Morrell L (2015) Getting to the roots of what digital innovation means. Internet retailing. http://internetretailing.net/issue/digital-innovation-report-october-2015/getting-to-the-roots-of-what-digital-innovation-means. Accessed 14 Nov 2018

National Science Foundation (2010) Cyber-physical systems (CPS). https://www.nsf.gov/pubs/2010/nsf10515/nsf10515.htm. Accessed 14 Nov 2018

Orel FD, Kara A (2014) Supermarket self-checkout service quality, customer satisfaction, and loyalty: empirical evidence from an emerging market. J Retail Consum Serv 21(2):118–129. https://doi.org/10.1016/j.jretconser.2013.07.002

Parada R, Melià-Seguí J, Morenza-Cinos M, Carreras A, Pous R (2015) Using RFID to detect interactions in ambient assisted living environments. IEEE Intell Syst 30(4):16–22

Parlak S, Marsic I (2013) Detecting object motion using passive RFID: a trauma resuscitation case study. IEEE Trans Instrum Meas 62(9):2430–2437

Peffers K, Tuunanen T, Niehaves B (2018) Design science research genres: introduction to the special issue on exemplars and criteria for applicable design science research. Europ J Inf Syst 27(2):129–139. https://doi.org/10.1080/0960085X.2018.1458066

Piotrowicz W, Cuthbertson R (2014) Introduction to the special issue information technology in retail: toward omnichannel retailing. Int J Electron Commer 18(4):5–16

Puerini GL, Kumar D, Kessel S (2015) Transitioning items from a materials handling facility. US Patent App. 14/495818

PwC (2015) Total retail 2015: retailers and the age of disruption. Technical report

Reynolds AP, Richards G, de la Iglesia B, Rayward-Smith VJ (2006) Clustering rules: a comparison of partitioning and hierarchical clustering algorithms. J Math Model Algorithm 5(4):475–504. https://doi.org/10.1007/s10852-005-9022-1

Rigby D (2011) The future of shopping. Harv Bus Rev 89(12):65–76

Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20:53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Roussos G, Kourouthanasis P, Spinellis D, Gryazin E, Pryzbliski M, Kalpogiannis G, Giaglis G (2003) Systems architecture for pervasive retail. In: Proceedings of the 2003 ACM symposium on applied computing, ACM, New York, NY, USA, SAC '03, pp 631–636. https://doi.org/10.1145/952532.952656

Senecal S, Nantel J (2004) The influence of online product recommendations on consumers' online choices. J Retail 80(2):159–169

Shankar V, Inman JJ, Mantrala M, Kelley E, Rizley R (2011) Innovations in shopper marketing: current insights and future research issues. J Retail 87:29–42

Stankovic JA (2014) Research directions for the internet of things. IEEE Internet Thing J 1(1):3–9

The Economist (2016) The future of personal transport: the driverless, car-sharing road ahead. https://www.economist.com/business/2016/01/09/the-driverless-car-sharing-road-ahead. Accessed 14 Nov 2018

Witten IH, Frank E, Hall MA, Pal CJ (2016) Data mining: practical machine learning tools and techniques. Morgan Kaufmann, Burlington

Wong WK, Leung S, Guo Z, Zeng X, Mok P (2012) Intelligent product cross-selling system with radio frequency identification technology for retailing. Int J Prod Econ 135(1):308–319

Yoo Y, Henfridsson O, Lyytinen K (2010) Research commentary-the new organizing logic of digital innovation: an agenda for information systems research. Inf Syst Res 21(4):724–735