Show Me the Money: RFID-based Article-to-Fixture Predictions for Fashion Retail Stores

Matthias Wölbitsch
Detego GmbH
Graz, Austria
m.woelbitsch@detego.com

Thomas Hasler
Detego GmbH
Graz, Austria
t.hasler@detego.com

Denis Helic
Graz University of Technology
Graz, Austria
dhelic@tugraz.at

Simon Walk
Detego GmbH
Graz, Austria
s.walk@detego.com

***Abstract*—Over the course of recent years, Radio Frequency Identification (RFID) technology has been applied in several different business domains to improve a diverse array of practical applications. For example, handheld RFID readers combined with passive RFID tags are used to perform fast and accurate stocktakes for fashion retailers. However, while this approach enables efficient inventory management, automatic localization of RFID-tagged goods in stores is still an open problem. To tackle this problem, we equip fixtures (e.g., shelves, tables, ...) with reference RFID tags and use data we collect during typical RFID-based stocktakes to map articles to fixtures. Knowing the location of goods within a store enables the implementation of several practical applications, such as automated Money Mapping (e.g., creating a heat map of sales across fixtures) or visual merchandising evaluations (e.g., monitoring sales of fixtures before, during, and after the implementation visual merchandising strategies). Specifically, we conduct (i) controlled lab experiments and (ii) a case-study in two fashion retail stores to evaluate our presented approaches for article-to-fixture predictions. The approaches are based on calculating distances between read event time series of article and reference tags using dynamic time warping, and clustering of read events using DBSCAN. We find that we can use read events collected during RFID-based stocktakes to assign articles to fixtures with an accuracy of more than 90% in several of our experiments. Hence, in this paper we present an exploratory venture into novel and practical RFID-based applications, beyond the scope of stock management.**

***Index Terms*—RFID, Money Mapping, DBSCAN, DTW**

I. Introduction

Radio Frequency Identification (RFID) technology is widely adopted in the fashion retail industry for inventory management along the supply chain [1]–[3]. RFID tagging of fashion goods with passive RFID tags during the production enables retailers to keep track of individual items during distribution, as well as inside their brick-and-mortar stores. A typical RFID setup in fashion retail stores consists of a handheld reader allowing store staff to perform daily RFID-enabled stocktakes. This drastically improves inventory accuracy [4], as stocktakes are performed quickly in short and regular intervals.

Problem. However, although RFID technology is already used by many retailers to improve their business processes, it is rarely used for the localization of goods inside stores. While there exist several RFID-based localization approaches [5]–[8], most of them require additional hardware, infrastructure, or parameter tuning, which is not feasible for most retailers due to the associated efforts, costs, and required expertise. Nevertheless, knowing the location of goods is beneficial for several real-world applications. For example, linking sales figures of articles to their placement within a store (i.e., *Money Mapping*) allows visual merchandising departments to better understand the behavior of their customers [9]. In fact, many store owners today evaluate their visual merchandising strategies by manually performing Money Mapping. However, this is a tedious and time consuming task—especially over time—due to the constant changes that take place in a fashion store, and is usually only executed for a subset of fixtures (e.g., for individual shelves or tables).

Approach. Therefore, we set out to tackle the problem of automatically determining locations of articles in fashion retail stores while keeping the administrative and financial overhead at a minimum. To that end, we use the recorded signal strengths of RFID read events during commonly performed stocktakes, in combination with strategically placed passive RFID reference tags, to infer locations of articles on fixture-level, based on temporal clustering and time series similarities. Hence, we follow a data-driven approach that requires no additional hardware other than passive RFID tags, nor adaptations to the underlying RFID-based stock taking process. The only overhead introduced by our approach is the onetime effort of placing and mapping reference tags to fixtures in a store.

Contributions. First, we present a methodology for article-to-fixture predictions, solely relying on read events of passive RFID tags collected using a single handheld reader. Second, we evaluate our proposed approach in a controlled laboratory environment, as well as in two real-world fashion retail stores, where we achieve article-to-fixture prediction accuracies of more than 90%. Third, we publish our large-scale data set[1], so that other researchers can reproduce the experiments and improve on our reported findings.

In this paper we present a novel methodology to leverage RFID-based technology not only for accurate and fast stocktakes, but also to help fashion retailers to keep track of article locations. We do this solely based on data collected during commonly performed RFID-based stocktakes, and outline how to use this additionally obtained information for practical applications such as automated Money Mapping.

978-1-7281-5576-0/20/$31.00 ©2020 IEEE

[1]https://github.com/detegoDS/show_me_the_money_dataset

II. Related Work

Kerfoot et al. [10] conduct a study on how visual merchandising (e.g., the placement and presentation of articles on the sales floor) affects the shopping behavior of customers in retail stores. They find that factors, such as the fixture design and materials they are made of can influence customers' willingness to buy. Lea-Greenwood [11] discusses how modern technology and tools help fashion retailers to design and implement visual merchandising strategies across their stores. Moreover, Rizzi and Volpi [9] outline the impact of RFID-enabled visual merchandising and how it can affect sales with respect to the occupied shelf space. However, they mainly focus on the financial impact and how it can be measured, while our work focuses on the RFID-based implementation that enables the automated assignment of goods to fixtures.

Other RFID-based localization approaches include work from Want and Katabi [12], who propose *PinIt*, an RFID-based approach to determine the position of goods based on a synthetic aperture radar generated by a moving RFID antenna. They compile multipath profiles based on the collected data and use dynamic time warping to approximate locations of RFID tags. Luo and Shin [13] introduce *FINDS*, a framework to detect misplaced RFID tags in smart shelf environments. They use static RFID antennas that measure the phase of tags and improve their measurements using stochastic optimization methods and density-based clustering (i.e., DBSCAN) to detect outliers (i.e., misplaced tags). Hasler et al. [14] use signal strength data collected using a handheld RFID reader to estimate the relative distances between tags by transforming read events into a two-dimensional space using multidimensional scaling. This enables them to identify misplaced RFID tags by inspecting their neighborhoods.

Moreover, Liu et al. [15] propose *TagBooth*, which is a method to mine shopping data (i.e., interactions of customers with goods) based on RFID read events. Specifically, they leverage recorded signal strength and phase information to infer customer actions (e.g., picking up articles from a fixture). Similarly, Zhao et al. [16] determine interesting, correlated, and popular articles based on signal strength patterns found in RFID read events. Furthermore, they are able to identify hot zones in stores that are frequently visited by customers. *ShopMiner* [17] and *CBID* [18] are other frameworks with the goal to obtain a better understanding of customer behavior in RFID-equipped retail environments.

Other customer-facing applications based on RFID technology in retail stores include digital personal shopping assistants [19], and smart fitting rooms that provide additional information about the articles a customer brings into the dressing room [20]. These technologies also enable retailers to provide product recommendations for their customers in brick-and-mortar stores [21]. Furthermore, RFID-based data-driven methods can be used, for example, to identify counterfeit articles in the supply chain [22], or to determine articles that are frequently missed during stocktakes [23].

III. Stocktake Routine & Practical Impact

Stock taking refers to the process of recording and counting all items located in a store (i.e., the inventory). Using RFID technology, the inventory of a store, which usually consists of several thousand individual items, can be recorded within a few minutes, while traditional manual stock taking methods take several days, and often require the temporary closing of the store in comparison.

RFID-based stocktakes are usually performed on a daily basis in the morning, right before the store opens. Store staff use an off-the-shelf handheld UHF RFID reader to record items, which are tagged with passive RFID tags. To that end, a member of the staff walks alongside the fixtures in an undetermined path and scans the items typically at a short but varying distance. In general, RFID-based stocktakes allow retailers to maintain stock accuracies (i.e., the difference between the expected and actual stock) of well beyond 90%, allowing them to know exactly which items are available in a given store. This builds the foundation for many state-of-the-art retail technologies, such as *Click&Collect*, where customers order online and collect the purchased articles in the store.

In this paper, we set out to further leverage the data that is already collected during daily RFID-based stocktakes to evaluate and implement an additional practical application without changing the underlying process, nor requiring extensive RFID parameter tuning by store staff. Specifically, we use the collected handheld reader data (i.e., timestamp and signal strength) to determine the placement of articles on fixtures. When enriching this information with other data streams, such as sales data, we can automatically identify "hot- and cold-fixtures", simply relying on the existing and unmodified daily stocktake routine, which is a powerful tool to monitor visual merchandising strategies. Further, additional use-cases, such as the compilation of smart picking lists (i.e., lists of articles that should be collected, ordered by their location in the store), can also be calculated based on the presented methodology, which we leave open for future work.

IV. Methodology

The first step in our methodology is to place reference RFID tags on fixtures. Next, we perform regular RFID-based stocktakes (see Section III) using a handheld reader, configured either to use Session 0 or Session 1, where we collect RFID read events that consist of the Received Signal Strength Indication (RSSI), the timestamp of the read event (millisecond resolution), and the unique identifier encoded on the RFID tags. Hence, we obtain a time-ordered sequence of RSSI measurements for all RFID tags (i.e., reference tags and item tags). Next, we aggregate reads of individual items belonging to the same article to overcome sparsity in the data set by leveraging commonly used "article stacks" (e.g., a stack of a given T-shirt in multiple sizes), while also aggregating reads of reference tags located on the same fixture. Next, we approximate distances between articles and fixtures using two different approaches, and finally transform distances into probabilities, which we use to predict article-to-fixture assignments. Moreover, we extend our approaches to leverage historic stocktake data to further improve prediction performance over time.

A. Fixture Tagging

For our work we enumerate all fixtures and place reference RFID tags on them. To that end, we use passive RFID tags that are already used in a store to, for example, replace other malfunctioning RFID tags. Note that the definition of what constitutes a fixture in a store setting is fluid and depends on local circumstances. For example, multiple smaller adjacent fixtures, can be combined to form one larger "logical" fixture that identifies an area in a store (e.g., shelves above a rail clothing rack). Ultimately, the store layout in combination with the required information needs define the trade-off between keeping the number of fixtures manageable (i.e., managing the mapping between fixtures and RFID tags), while still providing enough information about article locations to generate new insights for the retailer. Additionally, RFID technology itself imposes limits on the potential granularity of fixtures due to the noisy nature of the reading results. For example, small fixtures located very close to each other, which also contain only one or two reference RFID tags, are difficult to distinguish. We leave the problem of identifying the optimal number and type of reference RFID tags per fixture and the identification of the best fixture size open for future work.

B. Read Event Aggregation

Fashion retailer usually place multiple items of an article on the same fixture so that customers have a selection of different sizes in one location (i.e., article stacks). As the number of RFID reads per individual item is limited, and the recorded signal strength can be very noisy, we leverage that related items are already physically close to each other. Specifically, we perform an aggregation step where read events of items belonging to the same article are combined, thus allowing us to obtain a more dense series of reads on an article-level. Similarly, reads of reference tags that are located on the same fixture are also grouped together to achieve the same effect.

C. Distance Approximation Approaches

We have implemented two different approaches to tackle the problem of identifying the most probable fixture for an article. To that end, we calculate the "distance" between each article read event sequence and all other sequences that represent fixtures. Note that distance is an arbitrary measure that depends on the used fixture prediction approach.

Parameter Estimation. Both presented approaches depend on a set of parameters that we obtain using hyperparameter optimization, which is commonly performed practice in the machine learning domain. To that end, we conduct 5 stocktakes each for Session 0 and Session 1 in a laboratory setting, where locations of articles are known beforehand. We then apply our approaches to the collected stocktake data using a wide range of different parameter value combinations, to determine the best performing configuration. Note that we use the selected parameters for all our (real-world) experiments to evaluate their applicability in practical scenarios, where any additional parameter tuning would not be feasible.

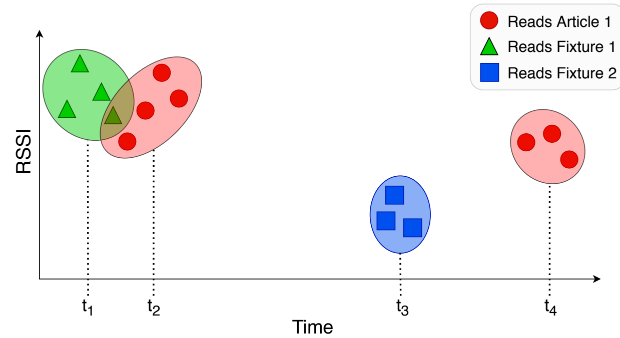


Fig. 1: **DBSCAN-based Approach Illustration.** We show clusters of read events obtained by DBSCAN for an article (●) and two fixtures (▲, ■) (i.e., reference tags). We use the minimum time difference between cluster centroids belonging to the fixtures and the article to determine the distance between them. For example, the distance between Article 1 and Fixture 2 is $d_{1,2} = |t_4 - t_3|$, as $|t_4 - t_3| < |t_3 - t_2|$. However, we ultimately assign Article 1 to Fixture 1 in this example based on the overall smallest distance $d_{1,1} = |t_2 - t_1|$.

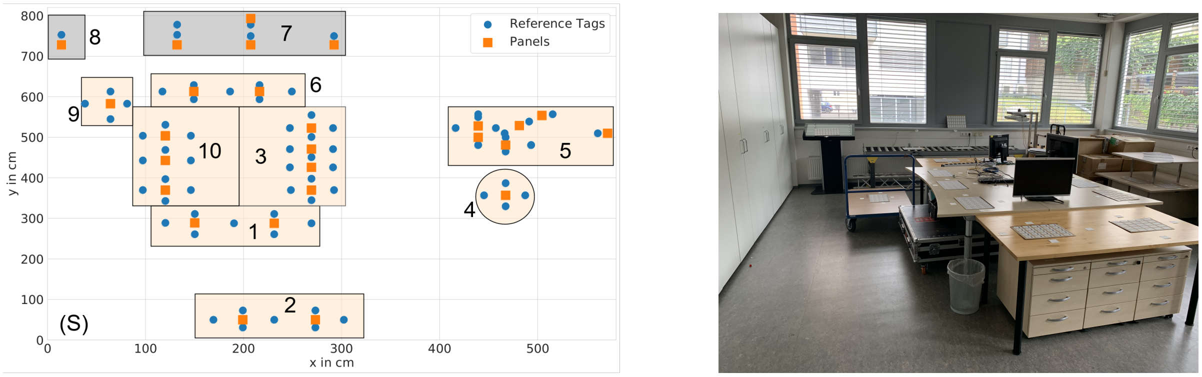
DBSCAN-based Approach. Our first approach is based on Density-based Spatial Clustering of Applications with Noise (DBSCAN) [24]. Essentially, with DBSCAN, we try to identify the shortest distance in time between clusters of read events for articles and fixtures. Specifically, DBSCAN groups timestamps of reads, which are close to each other based on Euclidean distance and a minimum number of points in their neighborhood, allowing us to identify one or more clusters per read event sequence, while simultaneously filtering noise.

For this algorithm we perform MinMax scaling on the time stamps of the read events such that the first read event happens at $t = 0$ and the last one at $t = 1$. Furthermore, we perform MinMax scaling on the signal strengths and remove all read events with a low RSSI value before clustering to reduce the influence of noisy reads. Specifically, we remove read events with an RSSI value smaller than the 0.8 quantile for Session 0 and the 0.77 quantile for Session 1.

For DBSCAN, ϵ defines the maximum Euclidean distance between two read events to be considered neighbors. We set this parameter to $\epsilon = 0.085$ for Session 0 and $\epsilon = 0.068$ for Session 1. Further, we define that a dense region (i.e., cluster) must contain at least 8 reads for Session 0 and 7 for Session 1, which is the second important parameter for DBSCAN.

Finally, the distance between a read event sequence of an article and a fixture is the minimum distance in time between the respective cluster centroids (see Figure 1).

DTW-based Approach. Our second approach is based on the Dynamic Time Warping (DTW) [25] algorithm, which can be used to determine the distance between two time series with different characteristics. Therefore, the first step is to convert



(a) Schematic of Laboratory Setup

(b) View from (S) in Schematic

Fig. 2: **Laboratory Setup.** Rectangles and squares represent fixtures (enumerated with the large numbers in Figure 2a). Orange markers represent panels (i.e., article stacks) and blue markers represent individual reference tags. Note that panels and reference tags on Fixture 5 are overlapping in our schematic, as it is a multi-layered presentation table (see top right of Figure 2b). For our lab setup, we have placed panels with 24 to 36 RFID tags and multiple reference tags on 10 fixtures. (S) marks the starting point for all of our performed stocktakes.

the irregular sequence of read events, consisting of timestamps and measured signal strengths (i.e., RSSI values), to a time series. To that end, we resample the read events for articles and fixtures to obtain equally spaced series of RSSI values that start and end at the same time. For our work, we resample read events with a resolution of 0.2 seconds for Session 0 and 0.1 for Session 1 and sum the RSSI values within a time window.

Note that due to the definition of the RSSI, the corresponding values in our setup are in the range of -80 to -20 dBm. For simplicity sake, we transform the RSSI in the positive domain (i.e., add 100 to the original value). This also conveniently allows us to set the signal strength in the resampled time series to 0 whenever no read event occurred within a time window. Furthermore, we discard read events with an RSSI value smaller than the 0.4 quantile for Session 0 and 0.5 quantile for Session 1 to filter noise.

We then use DTW to determine distances between article and fixture time series. To that end, DTW maps each point in the article time series to one or more points in the fixture time series. The actual distance between two time series is determined by the cost originating from this matching process. Note that for DTW we set the maximum window size w, which is the maximum shift allowed to match two points, to $w = 9$ for Session 0 and $w = 12$ for Session 1.

D. Article-to-Fixture Prediction

Next, we assign articles to fixtures based on distances approximated using the previously discussed approaches. To that end, we transform distances into probabilities via inverse distance weighting. More formally, the probability $p_{i,j}$ that article i is located on fixture j is given by

$$p_{i,j} = \frac{\frac{1}{d_{i,j}^2}}{\sum_k^N \frac{1}{d_{i,k}^2}}, \tag{1}$$

where $d_{i,j}$ is the distance between article i and fixture j and N is the total number of fixtures. Note that we use squared distances to penalize larger distances obtained by our approaches. Finally, we assign article i to the fixture with the largest probability.

E. Leveraging Historic Information

The quality of the collected data not only depends on the underlying RFID technology, but also how thorough stocktakes are performed. As a result, recorded stocktake data may contain noise, a limited number of read events, or other characteristics that have a negative impact on our results.

To tackle this issue, we leverage historic information from previous stocktakes. Due to the fact that article placements on fixtures are relatively stable (i.e., extensive rearrangements typically only happen when new fashion lines are introduced or seasons change) and RFID-based stocktakes are performed frequently (i.e., daily), we can assume that whenever we are confident about a fixture assignment for an article (i.e., high probability for the fixture), the article will most likely also be at the same fixture after the next stocktake. Hence, to compensate for stocktakes that contain many noisy reads we propose a Bayesian approach in which we take evidence from previous assignments into account when updating fixture assignment probabilities.

To that end, we calculate a certainty measure of the fixture assignment for each article. We define certainty as $1 - H/H_{\max}$, where H denotes the entropy of the fixture probability distribution, and $H_{\max} = \log_2 N$ is the maximum entropy (i.e., the entropy of the uniform distribution). As a result, if all fixtures are equally likely, the certainty of the assignment would be 0, while a distribution with an assignment probability for only a single fixture would lead to a certainty of 1. We scale the probability distributions of the current and the previous stocktake by their respective certainty scores and combine the two distributions by adding the individual fixture probabilities. Finally, we normalize the newly created distribution such that the entries sum up to 1 and assign the article to the fixture with the highest probability.

F. Evaluation Metrics

We evaluate our proposed methodology using two different metrics. First, we measure the accuracy of our computed assignments (i.e., the fraction of correctly assigned articles to fixtures). Second, as fixtures are often located very close to each other, we are also interested on how accurate our models are with respect to the distance to the actual fixture. For example, based on the local properties in stores, we expect that some articles will be be assigned to neighboring fixtures (i.e., "off-by-one" errors). To verify this intuition, we compile a distance matrix for the fixtures based on a regular two-dimensional grid that we laid over the floor plan. We use the Chebyshev distance metric as distance measure between two fixtures on the grid as it is easy to interpret. For example, a distance error of 1 using this metric corresponds to the incorrect assignment to a fixture on the neighboring square on the grid independent of the direction.

V. Experimental Setup

To evaluate and test our article-to-fixture prediction approaches, we perform multiple experiments on two different settings. First, we conduct a controlled laboratory experiment to determine the feasibility of our methods. Second, to evaluate our methodology in a real-world scenario, we visit two brick-and-mortar stores of a large international fashion retailer and conduct additional experiments.

A. RFID Setup

For all experiments we use a single off-the-shelf RFID handheld reader (i.e., Zebra RFD8500) and passive RFID tags that are commonly used in the fashion retail domain. We configure the reader to use Session 0 in *AB flip* mode to collect as many read events as possible. However, we also conduct stocktakes with Session 1, which reduces the number of reads per tag, as in contrast to Session 0, the tag does not respond to the reader for a longer period of time. This allows for higher throughput during stocktakes, but increases the difficulty of predicting fixtures for articles. Furthermore, we set the power level of the reader to its maximum (30 dBm; corresponding to 1W) and connect the RFID handheld device via Bluetooth to an Android smartphone, running the software that collects and stores the read events from the RFID reader.

B. Ground Truth

For each setting (i.e., in the laboratory and in both stores) we compile a ground truth, which maps both item and reference RFID tags to fixtures. We use this ground truth to evaluate the performance of our proposed methods.

C. Controlled Lab Experiments

We conduct our first experiment in a controlled laboratory setting (see Figure 2), where fixtures are located in a large room that also contains worktables, toolboxes, racks, and a conveyor belt. Specifically, we use a total of 10 fixtures, which consist of wooden tables (fixtures 1 to 6, and 10), a conveyor belt (fixture 7) as well as a standing desk touch-screen (fixture 8) and an open cart (fixture 9).

We place a total of 27 carton panels on the fixtures, each containing 24 to 36 RFID tags in a regular pattern. These panels simulate items belonging to the same article, mimicking commonly used "article stacks" in multiple sizes and colors on the same fixture. In total, we distribute 916 items (i.e., RFID tags) and 74 reference tags across 10 fixtures (see Figure 2a). Note that we manually select the required number of reference tags placed on a fixture based on the size and location of the fixture itself.

We conduct a total of 14 stocktakes in Session 0 and 19 in Session 1, all using the same starting position (see (S) in Figure 2a). However, for each experiment we vary the walking path and walking speed, while reading the RFID tags.

D. Real-world Experiments

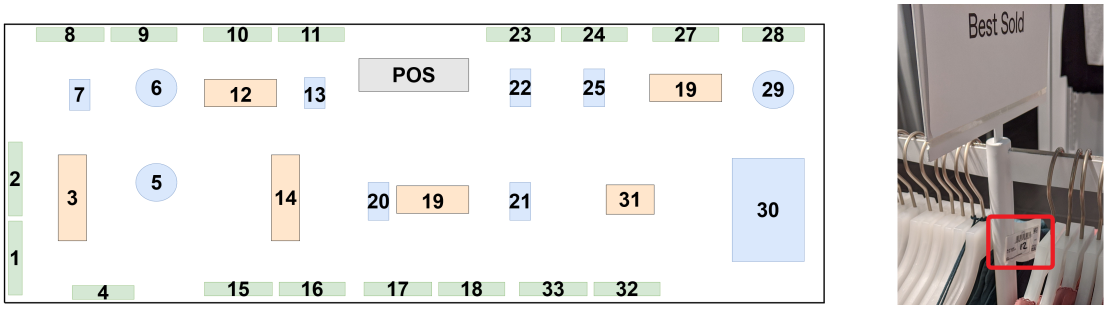
In addition to the stocktakes we perform in the laboratory, we also conduct stocktakes on the sales floors of two brick-and-mortar stores of a large international fashion retailer (see Figure 3). We select the two stores based on their different inventory sizes, shop areas, number of fixtures as well as the materials fixtures are made of. Store A is equipped with several fixtures that are made of metal, which can negatively affect reading performance during stocktakes due to signal reflections, while Store B is mainly equipped with wooden fixtures. The fixtures in the stores cover tables (on which items were placed in stacks), shelves (that can also include clothing rails), and clothing racks (that often contain a large quantity of individual articles).

Overall we observe 197 different articles (1,739 items) on 33 fixtures in Store A, and 200 different articles (1,977 items) on 23 fixtures in Store B. Despite fewer items on the sales floor, Store A has a larger number of items per square meter, which is also reflected in the larger number of fixtures compared to the other store (cf. Table I).

Similarly to the laboratory setup, we determine the number of reference tags per fixture in brick-and-mortar stores according to fixture sizes and types. For example, we put reference tags on the table surfaces near stacks of articles. In general, larger fixtures are tagged with more reference tags than smaller

TABLE I: **Data Set Properties.** For each experiment we list the number of articles and items, as well as the number of reference tags used in the setup. Furthermore, we state the number of fixtures (e.g., shelves, tables, ...). Note that not all fixtures were equipped with the same number of reference tags. Finally, we list the number of conducted stocktakes and their average duration in parenthesis.

Experiment	Articles (Items)	Ref. Tags	Fixtures	stocktakes
Lab	27 (916)	74	10	33 (01:43)
Store A	197 (1,739)	156	33	11 (07:29)
Store B	200 (1,977)	98	23	11 (10:06)



(a) Schematic Store A

(b) Ref. Tag Placement

Fig. 3: **Store Setup.** Rectangles and circles on the schematics in Figure 3a depict fixtures located on the sales floor of Store A, which has a size of about 180 square meters. Overall there are 33 fixtures located on the sales floor, which are either tables (brown), shelves (green), or clothing racks (blue). Furthermore, the point of sale (POS) is also shown. Moreover, Figure 3b shows an example for the placement of a reference tag on a fixture in this store.

ones. For fixtures with metallic rails we place reference tags on plastic hangers, which we evenly distribute along the rails.

In total, three participants conduct a total of 11 stocktakes (8 in Session 0, 3 in Session 1) in both stores. Note that unlike in the laboratory setting, we do not use a fixed starting position for the stocktakes and heavily vary walking paths to mimic realistic stocktakes as close as possible.

VI. Results & Discussion

A. Controlled Lab Environment

First, we evaluate our proposed methods to predict article-to-fixture assignments in a controlled laboratory setup (cf. Lab in Table II). In this setting, we place 26 different articles (i.e., panels containing RFID tags) on 10 fixtures, that are located within a few meters. Assigning articles to fixtures in this scenario at random leads to an accuracy of about 10%, which represents a baseline for our experiments. Our DBSCAN-based approach outperforms this baseline with more than 78% accuracy on average for stocktakes conducted in both Session 0 and Session 1. We find even better results using our DTW-based approach, which achieves an average accuracy of 89.9% for Session 0 and 90.9% for Session 1.

When leveraging historic information (i.e., article-to-fixture assignments computed in previous stocktakes) we find an improvement of accuracy for both our approaches. In the case of our DBSCAN-based method by 9.9% (7.1%) and for DTW-based method by 2.6% (2.2%) for Session 0 (Session 1). Note that we do not evaluate our results in this setup with respect to the distance for wrong article-to-fixture assignments due to the small number of fixtures.

Discussion. While the laboratory setting provides stable conditions for our experiments, the setup differs from a real-world environment in several dimensions (e.g., in terms of the number of articles, fixture types, . . .). Nevertheless, it also exhibits realistic properties, such as reflecting and absorbing surfaces, and representative distances between fixtures. Therefore, the high accuracies we achieve for both our approaches are promising.

Moreover, in the laboratory setting we find a similar performance for both sessions although we have on average 28.5% fewer read events for stocktakes in Session 1 than in Session 0. This indicates that the number of RFID reads events seems not to be an important factor in this setting to achieve good results.

When comparing the two approaches directly with each other, we find that the DTW-based approach outperforms the clustering-based approach. In general, performance for the time-series-based method is also more stable across different experiments, which is evident in the lower standard deviation of accuracy across multiple stocktakes. However, the observed gain in performance when leveraging historic stocktake information is larger for the DBSCAN-based method, which suggests that this additional input stabilizes prediction results to a larger extent. We can also observe this in the reduced standard deviation of achieved accuracies across all stocktakes.

B. Real-world Case Study

Next, we evaluate our methods in two real-world environments (cf. Store A and Store B in Table II). To that end, we conduct stocktakes on the sales floors of two fashion retail stores with different characteristics. As the number of fixtures is larger at these sites, the baseline, which randomly assigns articles to fixtures, performs worse compared to the controlled laboratory environment with accuracies between 2% and 6%.

In Store A we achieve a better average accuracy using our clustering-based approach compared to the lab setting in Session 0 with 81.5%, and 88.3% when leveraging historic information. However, for Session 1 stocktakes in Store A, we are not able to match the accuracy achieved in the lab using the same approach with 74.9% and 78.5%. While the DTW-based approach achieves the overall best performance in the lab setting, accuracy is low for stocktakes in Store A. For example, accuracy for Session 1 is only 65.7%, and slightly decreases to 64.0% when leveraging historic stocktakes.

When comparing the accuracies of our proposed methods between the two stores, we find that we can, in general, achieve a better performance across all approaches and sessions in Store B. Especially the DBSCAN-based approach that uses

TABLE II: **Experimental Results.** In the table we report the performance of our methods to predict fixtures on which articles are located based on stocktake data. We do this with respect to accuracy (i.e., the percentage of correctly assigned articles) as well as in terms of distance to the correct fixture for wrong predictions (i.e., the error). Hence, the distance error indicates how close the predicted fixture is to the correct one on the regular grid laid over the area in which the fixtures are located. We report averages (and standard deviation) for both metrics over the conducted stocktakes in Session 0 and Session 1 in three different setups. The accuracies we report were calculated using hyperparameters that we determined based on selected stocktakes conducted in the laboratory environment.

Session	Experiment	DBSCAN		DBSCAN w/ History		DTW		DTW w/ History	
		Accuracy [%]	Error	Accuracy [%]	Error	Accuracy [%]	Error	Accuracy [%]	Error
0	Lab	78.60 (17.84)	-	88.48 (6.79)	-	88.89 (6.78)	-	91.45 (3.74)	-
	Store A	81.51 (5.57)	1.16 (0.09)	88.34 (6.92)	1.21 (0.20)	70.51 (4.54)	2.03 (0.32)	71.72 (4.54)	2.62 (0.94)
	Store B	88.31 (6.51)	1.65 (0.46)	93.18 (3.65)	1.75 (0.52)	80.82 (5.36)	2.24 (0.70)	83.89 (2.43)	2.13 (0.75)
1	Lab	78.21 (12.20)	-	85.36 (8.99)	-	90.86 (6.39)	-	93.09 (4.41)	-
	Store A	74.89 (5.54)	1.46 (0.54)	78.51 (2.31)	1.69 (0.53)	65.74 (3.97)	3.19 (1.51)	64.04 (2.37)	4.57 (0.19)
	Store B	78.44 (7.83)	1.74 (0.10)	83.63 (12.57)	1.74 (0.17)	74.85 (5.39)	1.84 (0.13)	74.65 (5.09)	1.92 (0.02)

historic information achieves an average accuracy of 93.2% for Session 0 stocktakes at this site.

In contrast to the stocktakes we conduct in the laboratory, we also investigate the error with respect to the distance to the correct fixture for wrong article-to-fixture predictions. We find that these errors are generally very small for the DBSCAN-based approach across all settings (see Table II). Note that the minimum possible distance error is 1, which corresponds to the assignment of an article to the fixture that is located right next to the correct one. The average distance error for wrong predictions for stocktakes conducted in Session 0 in Store A is as low as 1.16 using this approach. Nevertheless, we also find that especially the DTW-based approach not only struggles in terms of accuracy in Store A, but also with larger distance errors for wrongly assigned fixtures.

Discussion. In contrast to the laboratory setting, we find that our time-series-based approach is outperformed by the clustering-based approach in a real-world environment. The margin between the two methods is also often very large. For example, for Session 0 stocktakes in Store B, where we can achieve our overall best average accuracy with more than 93%, the DTW-based approach only achieves an average accuracy of 83.9%. We hypothesize that the differences in performance between the two approaches in the real-world and laboratory setting is related to the number of read events collected during stocktakes. The number of items that are recorded in stores is much larger, which is beneficial for the clustering-based approach, as more stable and distinct clusters can be formed. In contrast to our expectations, the DTW-based approach can not take advantage of the additional data. By investigating different hyperparameter settings based on stocktakes performed in stores, we find that the DTW-based approach is also able to achieve similar performance as the DBSCAN-based approach in both stores. However, we find that the DTW-based approach requires a more careful selection of its hyperparameters, dependent on the store environment. In general, a larger resample window and lower threshold for the removal of reads with smaller RSSI values seems to be beneficial for this method. Nevertheless, for many retailers it is not feasible to generate a ground truth for their stores to fine-tune parameters due to the associated substantial effort that has to be made. Hence, our goal is to build a model that works for a variety of store setups. Our DBSCAN-based approach appears to be better suited to tackle this requirement as we can achieve good performance across three different sites.

Moreover, we can probably attribute the difference in performance between the stores to two factors. First, the two store environments differ substantially. Store A has more fixtures located on a smaller sales floor, and at the same time fewer items per fixture, compared to Store B. Hence, in Store A, we have less data to discriminate between a larger number of fixtures. The two stores also differ in the materials that fixtures are made of. Many fixtures in Store A are made from metal, which increases the potential for reflections of RFID signals, and leads to more noisy data. Second, we also experiment with the placement of reference tags on the fixtures between the two stores. For Store A, we place reference tags directly on the fixtures (e.g., on shelf frames), while for Store B we put reference tags on plastic hangers and wooden surfaces. Therefore, the direct placement of reference tags on fixtures that are made of metal in Store A is another potential factor for lower accuracies at this site. Note that the placement of reference tags on tables was consistent across both stores. These two factors render Store A a more challenging environment. Nevertheless, by leveraging historic article-to-fixture assignments with the DBSCAN-based approach we are still able to achieve an average accuracy of 88.3% in Store A, which is a promising result for many real-world applications.

Moreover, the distance error of 1.21 in the same store is encouraging as well, as the majority of wrong article-to-fixture predictions are assignments to one of the neighboring fixtures. This highlights that our proposed approach is not only able to accurately assign articles to the fixture that they are actually located on, but it is also not far off in terms of distance in case of a wrong prediction. This is an important trait for many real-world use cases as, for example, store staff can usually find articles that they want to retrieve in the near proximity of the predicted location in the few cases the prediction is wrong.

VII. Conclusions & Future Work

In this paper we present a methodology to determine article-to-fixture assignments in fashion retail stores by leveraging read events collected during ordinary RFID-based stocktakes. Specifically, we match RFID reads of strategically placed reference tags with reads of items that are located in a store. To that end, we propose two approaches to predict article-to-fixture assignments based on (i) clustering of RFID read events, and (ii) the similarity of time series generated based on the same data. Moreover, we also leverage historic stocktake data to increase the accuracy of computed article-to-fixture assignments. Using our methods we are not only able to assign articles to fixtures with an accuracy of more than 93% in real-world store environments, but also limit the error in terms of distance to the correct fixture for wrong assignments.

Therefore, only limited additional effort is required to gain information about the location of goods within a store without adapting or affecting the underlying stocktake process that is performed in retail stores on a daily basis. This allows fashion retailers to use this additionally obtained information in conjunction with other data streams (e.g., sales data), for example, to infer valuable insights about their customers or measure the effectiveness of visual merchandising strategies. Up until now, such evaluations required substantial manual effort by store staff, which often restricted their scope, and therefore their accuracy.

For future work, we plan to further improve the accuracy of our proposed models in real-world environments. To that end, we want to investigate the influence of the number and location of reference tags on the accuracy. In addition to that, we want to experiment with different prediction algorithms and include additional features into our models.

So far, we also only focused on the assignment of articles to fixtures in our work, which requires the aggregation of read events of items belonging to the same article. However, we are also interested in applying our approach on the items themselves, so that we can determine the fixtures on which items are located as well. This would open up additional use cases in retail stores, such as the detection of misplaced items.

The methodology we present in this paper highlights the potential in leveraging RFID-based stocktake data for additional use cases (e.g., Money Mapping), to further improve operational processes in fashion retail stores. Additionally, by publishing the data set[2] we collected during our experiments in two fashion retail stores as well as in a laboratory setting, we want to enable other researchers to develop methods for various applications in RFID-based retail environments.

[2] https://github.com/detegoDS/show_me_the_money_dataset

References

[1] K. Moon and E. Ngai, "The adoption of rfid in fashion retailing: a business value-added framework," *Industrial Management & Data Systems*, 2008.

[2] S. Garrido Azevedo and H. Carvalho, "Contribution of rfid technology to better management of fashion supply chains," *International Journal of Retail & Distribution Management*, 2012.

[3] G. Cilloni, R. Leporati, A. Rizzi, and G. Romagnoli, "State of the art of item-level rfid deployments in fashion and apparel retail," *International Journal of RF Technologies*, 2019.

[4] B. Hardgrave, J. Aloysius, and S. Goyal, "Does rfid improve inventory accuracy? a preliminary analysis," *International Journal of RF Technologies*, 2009.

[5] L. Ni, Y. Liu, Y. C. Lau, and A. Patil, "Landmarc: Indoor location sensing using active rfid," in *Proceedings of the 1st IEEE International Conference on Pervasive Computing and Communications*, 2003.

[6] C. Alippi, D. Cogliati, and G. Vanini, "A statistical approach to localize passive rfids," in *Proceedings of the 2006 IEEE International Symposium on Circuits and Systems*, 2006.

[7] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with rao-blackwellized particle filters," *IEEE Transactions on Robotics*, 2007.

[8] F. Xiao, Z. Wang, N. Ye, R. Wang, and X.-Y. Li, "One more tag enables fine-grained rfid localization and tracking," *IEEE/ACM Transactions on Networking*, 2017.

[9] A. Rizzi and A. Volpi, "Rfid-enabled visual merchandising in apparel retail," *International Journal of RF Technologies*, 2017.

[10] S. Kerfoot, B. Davies, and P. Ward, "Visual merchandising and the creation of discernible retail brands," *International Journal of Retail & Distribution Management*, 2003.

[11] G. Lea-Greenwood, "Visual merchandising: A neglected area in uk fashion marketing?" *International Journal of Retail & Distribution Management*, 1998.

[12] J. Wang and D. Katabi, "Dude, where's my card?: Rfid positioning that works with multipath and non-line of sight," in *ACM SIGCOMM Computer Communication Review*, 2013.

[13] J. Luo and K. Shin, "Detecting misplaced rfid tags on static shelved items," in *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, 2019.

[14] T. Hasler, M. Wölbitsch, M. Goller, and S. Walk, "Estimating relative tag locations based on time-differences in read events," in *Proceedings of the 2019 IEEE International Conference on RFID*, 2019.

[15] T. Liu, L. Yang, X.-Y. Li, H. Huang, and Y. Liu, "Tagbooth: Deep shopping data acquisition powered by rfid tags," in *Proceedings of the 2015 IEEE conference on Computer Communications*, 2015.

[16] J. Zhao, L. Wang, D.-a. Li, Y. Li, B. Yang, B. Zhu, and R. Bai, "Mining shopping data with passive tags via velocity analysis," *EURASIP Journal on Wireless Communications and Networking*, 2018.

[17] L. Shangguan, Z. Zhou, X. Zheng, L. Yang, Y. Liu, and J. Han, "Shopminer: Mining customer shopping behavior in physical clothing stores with cots rfid devices," in *Proceedings of the 13th ACM conference on Embedded Networked Sensor Systems*, 2015.

[18] J. Han, H. Ding, C. Qian, W. Xi, Z. Wang, Z. Jiang, L. Shangguan, and J. Zhao, "Cbid: A customer behavior identification system using passive tags," in *Proceedings of the 22nd IEEE International Conference on Network Protocols*, 2014.

[19] E. Ngai, K. Ka-Leung Moon, J. Nga-Kwok Liu, K. F. Tsang, R. Law, F. F. C. Suk, and I. C. L. Wong, "Extending crm in the retail industry: An rfid-based personal shopping assistant system," *Communications of the Association for Information Systems*, 2008.

[20] J. Al-Kassab, P. Blome, G. Wolfram, F. Thiesse, and E. Fleisch, "Rfid in the apparel retail industry: A case study from galeria kaufhof," in *Unique Radio Innovation for the 21st century*, 2011.

[21] M. Wölbitsch, S. Walk, M. Goller, and D. Helic, "Beggars can't be choosers: Augmenting sparse data for embedding-based product recommendations in retail stores," in *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, 2019.

[22] T. Staake, F. Thiesse, and E. Fleisch, "Extending the epc network: The potential of rfid in anti-counterfeiting," in *Proceedings of the 2005 ACM Symposium on Applied Computing*, 2005.

[23] M. Wölbitsch, T. Hasler, M. Goller, C. Gütl, S. Walk, and D. Helic, "Rfid in the wild-analyzing stocktake data to determine detection probabilities of products," in *Proceedings of the 6th International Conference on Internet of Things: Systems, Management and Security*, 2019.

[24] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 1996.

[25] J. Kruskal and M. Liberman, "The symmetric time-warping problem: from continuous to discrete," in *Time Warps, String Edits, and Macromolecules - The Theory and Practice of Sequence Comparison*, 1999.