Abstract—In this paper, we investigate the role of Twitter for German online retailers based on an empirical study. We analyze Tweets of the best selling German online retailers with qualitative and quantitative methods. Based on our results, we derive a conceptual model that can be used to classify different interaction strategies for online retailers. We identify four different interaction strategies. Three strategies are based on social interactions either on Twitter or by redirecting users to other social networks. The fourth strategy is applied to promote products of the retailers’ online store. Another result is that accounts which apply a user interaction focused strategy inside of Twitter through User Mentions, have a significantly higher number of followers compared to the accounts that use other strategies.

Keywords—E-commerce; Twitter; Study; Social Web.

I. INTRODUCTION

E-commerce plays an ever increasing role in the German economy. With the rise of the social web, the importance of social interactions with customers has also increased for online retailers. Traditional tools to reach customers include for example discussion boards, weblogs or newsfeeds. But in recent years, more and more online retailers have turned to the microblogging platform Twitter in order to reach out to new or existing customers.

A lot of research is conducted to understand the usage of Twitter. For example, the study of Java et al. investigated the user intentions and community structures of Twitter [1]. But to our best knowledge, there is no study that explicitly investigated the Twitter usage of online retailers. In this paper, we contribute knowledge to this area of research by presenting an empirical study that investigates the Twitter usage of German B2C online retailers.

In our study, we applied qualitative and quantitative methods to answer the following research questions:

- **RQ1**: How many online retailers are actually using Twitter and how active are they?
- **RQ2**: How can the Twitter interaction strategy of online retailers be classified?

The results to research question RQ1 will show the acceptance of Twitter as a tool to reach new and existing customers. Secondly, we answer research question RQ2 by analyzing Tweets in regard to different Twitter entities, e.g., URLs or User Mentions. We will present a conceptual model that can be applied to categorize the used Twitter interaction strategies of the conducted online stores.

The remainder of this paper is organized as follows. In Section II, we present the theoretical background. We then explain the design of our empirical study in Section III. After that, we show the results of the study in Section IV and discuss them in Section V. Finally, the article concludes with an outlook for future research.

II. THEORETICAL BACKGROUND

The main research areas addressed in this paper comprise the social web and e-commerce. The combination of social media and e-commerce are often denoted as Social Commerce [2], [3]. Most of the studies in social commerce investigate the customers’ perspective to the platform Twitter. In this study, we will focus on the retailers’ perspective instead.

Twitter is the most popular microblogging service in the web. Every day, about 500 million Tweets are published on this platform. Reasons for this success can be found in its simplicity, scalability, ubiquity, and interactivity. Due to its publish/subscribe capabilities, more and more users shift from traditional newsfeeds based on RSS and Atom to Twitter.

This publish/subscribe capability is Twitter’s fundamental pattern: Users subscribe either to other users or to Hashtags (HTs). The use of HTs is a communication convention, that enables authors to post a message either to a community or to add a content information [4]. Other conventions are User Mentions (UMs) (denoted by an @-sign) and Retweets [5]. From the Twitter API’s view, Hashtags, User Mentions and URLs are treated as special entities. The use of those entities are examined by [6] in a large scale study.

Previous works have investigated that Twitter is used for various communication purposes [7], [4]. Twitter entries can be classified in five different genres [8]: (a) Personal Updates, (b) Directed Dialog, (c) Real-time Sharing, (d) Business Broadcasting and (e) Information Seeking. These communication genres are discussed below.

The genre of personal updates contains mostly “daily chatter” [1], which means information about, what a user is thinking or doing at the moment. According to [1] this is the most common use of Twitter. According to [8] those posts are mainly issued by sparsely connected users.

A directed dialog is a message that is directed to a certain user by the use of an @-sign (User Mention). Measures to detect UMs are not exact, since there are other conventions that use the @-sign as well, but in Twitter traffic the vast majority of occurrences are UMs [9]. According to [10], p.21-25, UMs are an indicator about the interactivity of online relationships in Twitter.

In regard to business broadcasting, it is interesting to classify and measure the influence of a user. Mainly based
on the ratio between followers and friends, Twitter users can be categorized as broadcasters, acquaintances, miscreants and evangelists [11]. The number of posts as a function of the number of friends is increasing without any saturation [12]. Nonetheless, the classification of miscreants/spammers is controversial, as this group would also include users, that use Twitter primarily as newsfeed. The measurement of influence was refined by [13] and adopted to take Twitter’s characteristic features into consideration. The indegree influence is the number of followers of a user. This is the size of audience a user can reach, without directing a post via HTs to further channels. A Retweet influence indicates the ability to generate content, that is redistributed by the followers. The User Mention influence indicates the number of User Mentions containing the name of the examined user, which indicates the ability to engage the audience in a conversation.

The information seeking behavior is examined in [14]. They gathered and analyzed data for three different question types. They found out that the most popular type of questions are rhetorical questions with an overall very low response rate.

Another interesting aspect about the communication on Twitter are URLs that are embedded in Tweets. URLs are by default shortened by the platform’s own shortening service http://t.co [15]. The targeted URLs can be categorized as self-links, social media links and other external links. For our study, the former two contain interesting information. A self-link points to the own website of an online retailer and indicates the promotion of a product. Social media URLs direct users to discussions on other social networks, e.g., Facebook. This indicates a more community-centric activity than links to product pages.

### III. STUDY DESIGN

In this section, we will present the underlying design of our empirical study.

#### A. Data Collection

To acquire a sample of e-commerce related communications, we used a list of the 115 best-selling online-retailers in Germany [16]. We matched their Twitter accounts by querying search engines. If we were not able to find an account, we tried to collect the information manually by examining the online retailer’s website. We then evaluated manually, whether their Twitter account is targeting the German market. For this purpose we checked the profiles’ descriptions and timelines.

Based on the found Twitter accounts, we collected two different datasets. First, we retrieved common account information for the Twitter account, e.g., the lifetime of the account, number of followers, number of friends, number of favorites and the number of status updates issued by the account. Second, for each account we collected the timeline consisting of last 100 status updates. We used the Twitter REST API to access this data.

For the analysis of Tweets, we queried the REST API for the last 100 status updates issued by the acquired accounts (see Section IV-A). From this sample we extracted a smaller subset of 200 Tweets from 10 different retailers, which was used to identify genres by a manually performed content analysis. The data collection was carried out on 14th of February, 2014.

#### B. Analysis

First, we analyzed the data set containing the common information of the accounts. The lifetime of an account in days is defined as $L$. To calculate the Tweet rate $R_T$ per lifetime as an indicator for broadcasting activity, we use the total number of Tweets since the creation of the account ($T$) and relate it to $L$ (see (1)).

$$R_T = \frac{T}{L} \quad (1)$$

In order to analyze the links to other users, the number of followers $f_{in}$, the number of friends $f_{out}$, and the listed count $l_{in}$ can be used. As stated in Section II, those values are considered as in- and outdegree measures. To reflect the lifetime of an account, we relate those values to $L$ and, thus, define indegree rate $R_{in}$ and $R_{out}$ in (2) and (3).

$$R_{in} = \frac{f_{in} + l_{in}}{L} \quad (2)$$

$$R_{out} = \frac{f_{out}}{L} \quad (3)$$

We used the timeline data set to analyze the data set quantitatively to derive information about the communication strategy implemented by a retailer. We were particularly interested in distinguishing between the purposes to engage users in a dialog or to promote products directly. User Mentions and URLs are indicators for these purposes. A UM indicates the attempt to start a conversation with a customer or to reply to a post issued by a user. Depending on the target, a URL serves as an indicator for both categories: A link to a site of the retailers’ online shop follows the purpose of promoting a product, whereas a link to the companies blog or a social network is clearly a community management activity. To enrich the coarse grained strategical categories, we analyzed the sample of Tweets qualitatively to identify e-commerce related Tweet genres, that represent common interaction activities between retailers and consumers.

#### IV. RESULTS

In this section, we will present the results of our study. First, we present a common analysis of the investigated accounts. Second, we describe the found interaction strategies of the investigated accounts and illustrate those strategies with Tweet genres.

##### A. Analysis of accounts

First, we have identified the accounts that are managed by the online retailers. Figure 1 shows the results of the identification and retrieval process. As bar (a) shows, we were able to identify 88 accounts for our list of 115 retailers. The removal of accounts not targeting German users is represented by bar (b). We excluded 11 accounts from further retrieval. Bar (c) visualizes the retrieval of account information, which depend on the accounts settings. We successfully retrieved the information for 76 shops. As bar (d) shows, 59 of those accounts were classified as being active in the last month.
Most of the account names reflect the shop’s name, which can consist of the company’s name or a brand name. Besides that, many ids also include regional attribute, which can be a country name, a region name or a country code. The latter could also reflect the shop’s URL.

The variables lifetime, status updates per day, indegree and outdegree derived from the profile information are summarized in Table I. We calculated the .25, .50 and .75 quantiles and added Geary’s skewness indicator. The mean account lifetime is slightly above 4 years, while the values are slightly left-skewed. Since its creation, an account issued on average about seven Tweets per day, whereas the data is right-skewed. The majority of accounts issued less than two Tweets per day. The mean increase of indegree is about 1.6 followers per day, while the data is right-skewed. For 75% of the retailers, this value is at about 1.5 or less. The outdegree value for 75% of the accounts is at about 0.6 or less.

B. Twitter interaction strategies

To determine the Twitter interaction strategies of the conducted accounts, we first define the following sets:

- \( A \): All Twitter accounts whereas each element \( a \) represents an online retailer from the sample.
- \( T \): All observed Tweets.
- \( T_M \): All Tweets that address other Twitter users, i.e., Tweets that contain at least one User Mention.
- \( T_S \): All Tweets that contain at least one URL that is referring to another social network, e.g., Facebook.
- \( T_F \): All Tweets that contain at least one URL that is referring to the online store, e.g., URLs that link to certain products or special offers.

Based on the Tweets that contain User Mentions (\( T_M \)), URLs to social networks (\( T_S \)) and URLs to the online store of the account owner (\( T_F \)), we derived our conceptual model as depicted in Figure 2. As shown in the figure, we identified four different strategies, which will be explained below.

The first strategy \( S_1 \) is characterized by a frequent communication with other Twitter users. Accounts that apply this strategy (\( A_{S_1} \)) make use of User Mentions in at least two-thirds of their Tweets (see (4)).

\[
A_{S_1} = \left\{ a | a \in A; \frac{T_M}{T_a} \geq 0.66 \right\} \tag{4}
\]

Strategy \( S_2 \) is applied by accounts that intend to direct Twitter users either to the weblog of the company or to a website of another social network (e.g., Facebook) to continue communication there. We assigned accounts to this strategy when at least two-thirds of their Tweets contain URLs to other social networks or to a company weblog (see (5)).

\[
A_{S_2} = \left\{ a \big| a \in A; \frac{T_S}{T_a} \geq 0.66 \right\} \tag{5}
\]

Strategy \( S_3 \) is categorized by accounts that use strategy \( S_1 \) and \( S_2 \) moderately, i.e., accounts that make moderate use of User Mentions and moderate use of URLs to other social networks or weblogs. Moderate use means that User Mentions and social network URLs are present in at least one third, but no more than two-thirds of the Tweets (see (6)).

\[
A_{S_3} = \left\{ a \mid a \in A; 0.33 \leq \frac{T_M}{T_a} < 0.66 \right\} \bigcap \left\{ a \mid a \in A; 0.33 \leq \frac{T_S}{T_a} < 0.66 \right\} \tag{6}
\]

The last strategy \( S_4 \) is based on the URLs in Tweets that refer to the website of the online retailer. An account is using this strategy, when at least two-thirds of the Tweets contain URLs to the retailers online store, i.e., URLs for promotional purposes (see (7)).

\[
A_{S_4} = \left\{ a \mid a \in A; \frac{T_F}{T_a} \geq 0.66 \right\} \tag{7}
\]

Finally, all accounts that did not fit into the four above strategies were classified as having no clear strategy (see (8)).

\[
A_{S_4} = A \setminus (A_{S_1} \cup A_{S_2} \cup A_{S_3} \cup A_{S_4}) \tag{8}
\]

A short summary of the above mentioned strategies, their relative occurrences and their average number of followers are displayed in Table II. As shown in the table, strategies \( S_1 \) to \( S_3 \) can be classified as an interactive strategy, while strategy \( S_4 \) can be best described as a promotional strategy.

Based on these interaction strategies, we have qualitatively identified Tweet genres, that represent various communication purposes in the e-commerce. To perform that task, we have analyzed the timeline containing the last 20 Tweets of 10 accounts \((n=200)\). We illustrate each genre with synthetic...
TABLE I. VARIABLES CHARACTERIZING THE RETAILERS’ TWITTER PROFILES.

<table>
<thead>
<tr>
<th></th>
<th>min</th>
<th>Q25</th>
<th>Q75</th>
<th>mean</th>
<th>Q25</th>
<th>max</th>
<th>skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lifetime in years (L)</td>
<td>0.893</td>
<td>3.748</td>
<td>4.563</td>
<td>4.139</td>
<td>4.822</td>
<td>5.777</td>
<td>-1.350</td>
</tr>
<tr>
<td>Tweet rate (R_T)</td>
<td>0.057</td>
<td>0.418</td>
<td>1.208</td>
<td>7.711</td>
<td>1.894</td>
<td>367.300</td>
<td>7.345</td>
</tr>
<tr>
<td>Indegree rate (R_in)</td>
<td>0.055</td>
<td>0.482</td>
<td>0.987</td>
<td>1.575</td>
<td>1.471</td>
<td>8.258</td>
<td>2.209</td>
</tr>
<tr>
<td>Outdegree rate (R_out)</td>
<td>0.004</td>
<td>0.068</td>
<td>0.198</td>
<td>0.620</td>
<td>0.629</td>
<td>8.008</td>
<td>4.314</td>
</tr>
</tbody>
</table>

Figure 2. Conceptual model derived from our exploratory research.

TABLE II. SUMMARY OF THE IDENTIFIED TWITTER INTERACTION STRATEGIES.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Description</th>
<th>Occurrences</th>
<th>Ø followers</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>High interaction with other Twitter users through User Mentions</td>
<td>19%</td>
<td>6048</td>
</tr>
<tr>
<td>S2</td>
<td>Redirecting Twitter users to other social network sites (e.g., Facebook)</td>
<td>24%</td>
<td>1175</td>
</tr>
<tr>
<td>S3</td>
<td>Moderate interaction through User Mentions and moderate redirection to social networks</td>
<td>5%</td>
<td>3097</td>
</tr>
<tr>
<td>S4</td>
<td>Posting URLs to promote products of the own online store</td>
<td>21%</td>
<td>2398</td>
</tr>
<tr>
<td>Sx</td>
<td>No clear strategy</td>
<td>31%</td>
<td>1994</td>
</tr>
</tbody>
</table>

TABLE III. GENRES AND SYNTHETIC EXAMPLES FROM OUR QUALITATIVE ANALYSIS.

<table>
<thead>
<tr>
<th>Genre</th>
<th>Example(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>News</td>
<td>“Today would be Mozart’s 258th birthday”</td>
</tr>
<tr>
<td>Review</td>
<td>“@user1: very fast delivery by @retailerA”</td>
</tr>
<tr>
<td></td>
<td>“RT @user2: testing guitar manufactured by AESP <a href="http://shortenzer/URL">http://shortenzer/URL</a> @retailerA”</td>
</tr>
<tr>
<td></td>
<td>“Test: Steinberg Cubase 7.5, Digitale Audio Workstation <a href="http://shortenzer/URL%E2%80%9D">http://shortenzer/URL”</a></td>
</tr>
<tr>
<td>Promotion</td>
<td>“under #offer: TOUCHLET X7Gs Tablet-PC for only EUR 89,90! <a href="http://shortenzer/URL%E2%80%9D">http://shortenzer/URL”</a></td>
</tr>
<tr>
<td>Service</td>
<td>“@user3 Our customer service help@retailerC sure has an idea. They will contact you.”</td>
</tr>
<tr>
<td></td>
<td>“@user4 We’ve ordered 20 pieces in November, but we still don’t have a delivery date”</td>
</tr>
<tr>
<td>Dialog</td>
<td>“@user5 What exactly are you planning? #renovate #cm”</td>
</tr>
<tr>
<td></td>
<td>“@user5 Wow, that’s quite a project! How many squaremeters? #cm”</td>
</tr>
<tr>
<td></td>
<td>“@user5 You’ll certainly cope with it! #cm”</td>
</tr>
<tr>
<td>Recruiting</td>
<td>“@retailerD Do you need a job? Apply now! <a href="http://shortenzer/URL%E2%80%9D">http://shortenzer/URL”</a></td>
</tr>
<tr>
<td>Request for collaboration</td>
<td>“To Celebrate #NationalPuzzleDay we are giving away 2100 of puzzles for chance to win just RT&amp;Follow ends 6pm 31/01/14”</td>
</tr>
<tr>
<td>Agent</td>
<td>“@user5 You’ll certainly cope with it! #cm”</td>
</tr>
<tr>
<td></td>
<td>“@user6 Please send an email with your customer number to <a href="mailto:twitter@retailerE.de">twitter@retailerE.de</a>.”</td>
</tr>
<tr>
<td></td>
<td>“We are looking forward to help you soon as possible. “Tom”</td>
</tr>
<tr>
<td></td>
<td>“@user7 I need to ask a colleague. I’ll get back to you (ar)”</td>
</tr>
</tbody>
</table>

Examples in Table III. Because Twitter’s terms of service prevents republication of Tweets, we needed to create synthetic examples of our sample for publication. Those Tweet genres are explained below.

**News**: A Tweet that contains news that are related to the business domain of the retailer. Both examples for this category are published by a retailer for musical instruments. Both contain news that might be interesting for their customers. Since the Tweet contains neither HTs nor UMs, we assume that the Tweet is intended to provide an additional incentive to follow the account.

**Review**: A Tweet transporting a review for a product, service or overall quality of service. The issuer of the Tweet could be a customer or the retailer itself. The Tweet could be self-contained or could link to media located on the retailer’s site or on external media. Examples 3 and 4 are retweeted by the company and contain a review generated by a user.

**Promotion**: The intention of these Tweets are to promote products or special offers. The Tweets might contain an URL linking to product page and the respective price of that product. Both attributes are contained in the corresponding examples.
Service: Tweets that answer questions posted by customers either before or after purchasing a product. Example 7 indicates that the retailer answers to a message of a customer after purchasing a technical device. Example 8 is also an answer to a directed post by a user, but in contrast to the former the user was asking about the availability of an article that was sold-out at this time.

Dialog: An retailer’s agent engages in a dialog in requesting more information, e.g., about a project. The corresponding example shows a dialog that consists of 5 entries and which was initiated by customer with a directed Tweet.

Recruiting: Companies try to promote job offers online. In the corresponding example, a URL linking to a more detailed job description is contained. To increase the audience reached by such an advertisement, usually HTs are added. In example 12, the author misleadingly tags the company’s name, which is not very likely to increase the number of direct followers.

Request for collaboration: In order to enhance the attention and visibility, companies try to motivate users to collaborate in retweeting posts issued by a retailer. In example 13 the company uses a competition to motivate users to retweet.

Tagging agent: Adding an attribute, which identifies the person issuing a Tweet on behalf of a company. The corresponding examples illustrate three different variations that were found in the sample. Examples 14 and 15 use the circumflex sign to mark this attribute, whereas the former uses an abbreviation. In example 16, the issuer is identified by a numeric id, which is annotated with a hash notation. The example shows an abbreviated name in brackets.

V. Discussion

In Section IV-A, we analyzed the existence and activity of online retailers on Twitter. In our sample, about 77% of the retailers maintain a Twitter account and about 78% of these accounts were active at the time of retrieval. According to our data, the majority of the accounts were created since 2010, which means that the adoption in German e-commerce started roughly two years after the dramatic growth in popularity in the U.S. in 2008 [9]. This reflects the slower adoption of the German population [6], [17], [18]. Besides that, we observed that the indegrees of those accounts are higher than the outdegrees. Which means that online retailers do not tend to follow back each of their followers, to improve the reputation of their customers. This characteristic is consistent with the category of broadcasters defined in [11]. Surprisingly, there are also outliers that do follow back all of their followers.

For the identification of different communication strategies, we have analyzed the use of URLs and UMs in Section IV-B. A limitation of this part of our work is, that we analyzed only the locations of the URLs and not the content behind it. We can’t rule out the possibility that, for example, a Facebook URL points to a post, that contains a link to a product promotion. This was not further examined in our study.

We found out that 31% of the accounts did not follow a clear communication strategy. These accounts are not regularly involved in user interactions and are not using Twitter often to promote products. The other 69% of the accounts were following a specific strategy according to our conceptual model (see Figure 2). Although our model allows the occurrence of multiple strategies per account, we did not find accounts that actually applied more than one strategy. This shows that our defined strategies are disjunctive and clearly separated from each other.

We found that Twitter accounts with a frequent user interaction through User Mentions, have a significantly higher number of followers. On the other hand, Twitter accounts that only post URLs to redirect users to other social networks (e.g., Facebook), have the lowest number of followers. These two results leave room for interpretation. One possible explanation is that a direct contact to customers through User Mentions is a motivator for users to follow the online retailers’ account. An interpretation for the low follower count of accounts that only redirect users could be that the targeted users are already connected to the online retailer in another social network (e.g., they are friends in Facebook). Because the interaction happens in another social network, it could be of less value for a customer to follow an online retailer on Twitter.

Besides that, we derived genres of Tweets from a subset of our sample, to describe the coarse grained strategies in detail. Those genres build a tool for classification and analysis of direct interaction between companies in the e-commerce and users on Twitter. Due to the exploratory character of this study and the relatively small data set, we did not quantify the frequency of those genres. Our results could build the conceptual base for such a study.

VI. Conclusion and Future Work

In this paper, we made three contributions. First, we have determined the share of retailers, that maintain an active account on Twitter. Second, we have identified genres and communication patterns used in the e-commerce for direct communication with users. Those categories can be used as an analytic framework for studying microblogging in the Social Commerce. Third, we have derived basic interaction strategies from our data set and created a conceptual model based on these strategies. We also showed the usage of User Mentions and URLs and their purpose inside of Tweets.

In our future work, we primarily aim at working on the limitation of our results, as stated in Section V. An important aspect will be the validation and refinement of our approach to classify the communication strategies as mentioned in Section IV-B. We aim to perform an in-depth content analysis on a subset of our data to evaluate our result. Besides that, we plan to collect and label a larger data set and quantify the use of the Tweet categories defined in this article. Another limitation of our approach lies within the composition of our sample with a focus on the German market. Thus, we plan to perform a similar study with samples focusing on different markets and compare the results to the results of this work.

REFERENCES


