Comparing the Twitter Usage of Online Retailers in Germany and in the UK

Georg Lackermair, Daniel Kailer
Munich University of Applied Sciences
Department of Computer Science and Mathematics
Munich, Germany
Email: {georg.lackermair, dkailer}@hm.edu

Abstract—The usage and acceptance of Twitter microblogging differs from region to region as various works have shown. As this platform is gaining importance as a channel to reach customers in the E-commerce, the question arises whether this differences get apparent in the Twitter usage of online retailers as well. This paper investigates this question by comparing German and UK based online retailers empirically. A data set composed of the top selling companies from both countries is analyzed quantitatively. For this purpose, a conceptual model is presented to classify different interaction strategies for microblogging in the E-commerce domain. There are four different strategies used to distinguish between a more bidirectional, interactive communication and a more unidirectional, promotional communication. The model used distinguishes between direct dialogs inside the borders of Twitters and the redirection of users to other social networks.

Keywords–E-commerce; Twitter; Social Web; Germany; UK.

I. INTRODUCTION

The transformation of the Web from an unidirectional media of linked content towards an interactive communication platform affects the E-commerce massively. Due to its distributed nature, the Web 2.0 extends the range of channels for distributing information to the (potential) customers respectively. Another aspect is the bidirectional flow of communication. This requires companies to take care of information published by users and to react somehow to this.

A major actor in the social web is the microblogging platform Twitter. Every day, about 500 million status messages are published on this platform. Reasons for its success can be found in its simplicity, scalability, ubiquity, and interactivity. Due to its publish/subscribe capabilities, traditional newsfeeds based on Rich Site Summary (RSS) or Atom are shifting gradually to Twitter. This development applies to the E-commerce as well, as within this domain microblogging plays the role of a update notification capability for tools for personalization and direct customer interaction, e.g., discussion boards, weblogs or newsfeeds.

In academia, a considerable amount of research was already conducted to understand the usage of Twitter, in particular the communication conventions, user intentions and the network structure. Several works suggest that the usage of this platform is related to regional characteristics (1)(2). A blog post by The Economist states that Twitter is in Germany less popular than in other countries (3). The data shows that the ratio of Twitter accounts related to the size of population is a multiple times higher in Great Britain than in Germany.

But to the best knowledge, there is no study that explicitly investigated the Twitter usage of online retailers and compares two different samples against each other. This paper examines the following research questions:

1) How many retailers in Germany are using Twitter compared to the UK?
2) Are there differences in how German retailers are using Twitter compared to the UK?

To answer the above questions, an empirical study was conducted. A sample was collected to compare the Twitter usage in the E-commerce in Germany and UK. Question 1 is answered by hand of account related data retrieved from the Twitter profiles. In order to answer question 2, a model is presented to classify a communication strategy based on directed messages and embedded URLs.

This paper is organized as follows: In Section II, the theoretical background is presented. Then, the research design of the empirical study is explained in Section III. After that, the results of the collection and analysis are presented in Section IV and discussed in Section V. Finally, this paper 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* (4)(5). Most of the studies in Social Commerce investigate the customers' perspective to the platform Twitter. This study is focused on the retailers' perspective instead.

A. Conventions

The publish/subscribe capability is Twitter's fundamental pattern. The different possibilities for routing a status message to other users are summarized in Fig 1. Users subscribe either to other users or add another user to a list. Every message issued gets by default submitted to the following users and to all lists. To broaden the audience, one can embed *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 (6). Besides that, messages can be directed to a single user by annotating the user's name with an @-sign, which is called *User Mention*. Retweeting means the redistribution of a Tweet to a user's own audience (7). From the Twitter API's view, Hashtags, User Mentions and URLs are treated as special entities. The use of these entities are examined by (2) in a large scale study.
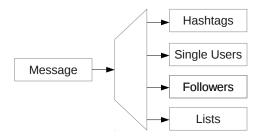
Figure 1: Routing of status messages on Twitter.

B. Purposes

Previous works have investigated that Twitter is used for various communication purposes (8)(6). Twitter entries can be classified in five different genres (9)(10): personal updates, directed dialog, real-time sharing of news and links, business broadcasting, and information seeking. For the study of E-commerce related communication, the broadcasting behaviour and dialogs are interesting. The success of broadcasting activities can be measured in the size of audience, redistribution rate, and the ability to engage users in a dialog (11)(12)(13). Dialogs can be identified by the use of the @-sign, which is often used to direct a message to the user mentioned. Though, this is no exact measure, the vast majority of @-signs are used for UMs in Twitter (14) and, thus, can be used for indicating the interactivity of an online relationship (15, p. 21 - 25).

Related to the E-commerce, by means of an explorative observation of 200 Tweets from 10 online-retailers domain-specific communication practices could be discovered:

- *News:* Published news that are related to the business domain of the retailer, without any direct relation to a product or service offered by the company.

- *Review:* Messages that contain links to a review about a procuct or service offered by the company. The status can be issued by the retailer or by a customer an republished (retweeted) by the company.

- *Promotion:* Tweets that promote products or services directly. Those kind of messages often contain an URL linking to the product page and the respective price of the offer.

- *Dialog:* Directed messages flow betweet the retailer and a customer. The purpose of such dialogs vary: there are examples, where customers request further information about a product or date of delivery via Twitter.

- *Agent:* Retailers mark their messages with an attribute representing the issuing agent personally. This is a communication convention, which is not explicitly supported by Twitter, but used to facilitate directing a message to a certain agent acting on behalf of a company's account. Mainly, the circumflex notation (^Name) is used to annotate an agent.

C. Linking

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 (16). The targeted URLs can be categorized as self-links, social media links and other external links. For the study presented, 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, the design of the empirical study will be presented.

A. Data Collection

To acquire a sample of E-commerce related communications, two lists of the 115 best-selling online-retailers in Germany (17) and the 100 best-selling shops in UK (18) were used. The selection of retailers for further analysis was performed in four consecutive steps:

1) *Find account:* Twitter accounts were matched to the shop sites by querying search engines. In case that the query did not return a valid result, the search process was continued by examining the shopping site manually.

2) *Targeting specific country:* With this step, it was checked, if a given Twitter account is really targeting the respective country (Germany of UK). For this purpose, a manual examination of each profile's description and timeline was performed.

3) *Retrieval of account data:* The profile information was collected, including the accounts' lifetime, the number of connections to other accounts (followers, friends, listed) and the number of statuses issued since creation.

4) *Activity check:* In order to filter out inactive accouts, the last status, issued by the regarding account was retrieved. Then, the time elapsed between the publication of the last Tweet and the retrieval was calculated. If the last status was issued more than 30 days ago, an account was considered as not being active anymore.

After the selection and retrieval of profile information, the timeline consisting of the last 100 status messages were collected for each account passing the preceding process. Twitter's REST API was queried to retrieve the data set. The data collection was performed on 14th of February, 2014 for the German subset and on 25th of February for UK.

B. Account data

First, the profile information was analyzed. The lifetime of an account in days is defined as L. In order to calculate the Tweet rate R_T per lifetime as an indicator for broadcasting activity, the total number of Tweets since the creation of the account (T) was related to L (see (1)).

$$R_T = \frac{T}{L} \tag{1}$$

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, those values were related to L and, thus, they define the indregree rate R_{in} and R_{out} in (2) and (3).

$$R_{in} = \frac{f_{in} + l_{in}}{L} \tag{2}$$

$$R_{out} = \frac{f_{out}}{L} \tag{3}$$

C. Interaction strategies

First, different strategical categories were defined for the study. As described in Section II-B, there are different purposes of Twitter communication. One purpose is the dialog with users, which is called *interactive strategies* subsequently. Two major forms of such strategies can be distinguished: S_1 for communicating inside Twitter and S_2 for distributing links pointing to social networks for dialogs outside Twitter. While S_1 and S_2 were used, when the corresponding attribute dominates, a third form S_3 is introduced for cases, where both attributes dominate only when combined. Thus, S_3 describes an interactive focus, combining both Twitter dialogs and linking to other Web 2.0 sites. A promotional strategy S_4 is indicated by the use of URLs pointing to the shop, owned by the issueing account. S_x is assigned, when none of those attributes dominates in such a manner that one of the other strategies could be assigned.

To determine the Twitter interaction strategies of the conducted accounts, the following definitions are introduced:

- A: All Twitter accounts whereas each element a represents an online retailer from the sample.

- P_M: The fraction of Tweets that address other Twitter users relative to.

- P_S: The fraction of Tweets that contain at least one URL linking to another social network.

- P_P: The fraction of Tweets that contain at least one URL pointing to the online store of the issuer.

A conceptual model was derived, which is based on the Tweets that contain User Mentions (P_M), URLs to social networks (P_S) and URLs to the online store of the account owner (P_P). As shown in Figure 2, 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 make use of User Mentions in at least two-thirds of their Tweets.

Strategy S_2 is applied by accounts that intend to direct Twitter users either to the weblog of the company or to a
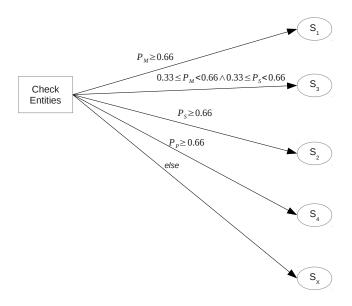

Figure 2: Decision tree for assigning strategical category.

website of another social network (e.g., Facebook) to continue communication there. Accounts were assigned with this strategy when at least two-thirds of their Tweets contain URLs to other social networks or to a company weblog.

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.

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.

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

IV. RESULTS

In this section the results of the study are presented. First, an analysis of the profile information is given, and, second, accounts are assigned with interaction strategies.

A. Analysis of accounts

For each step of the selection and retrieval process described in III-A, the resulting population size of retailers passing the respective step is examined. The passing of this process is summarized in Table I and divided into the two samples Germany and UK. The value $\frac{x_i}{n}$ is the percentage of the number of accounts passing the step i relatively to the sample's n. The fraction $\frac{x_i}{x_{i-1}}$ represents the number of accounts passing the step i relatively to the previous step $i-1$.

It can be noted that both subsets differ from each other strongly in steps 1 and 4. While the corresponding Twitter

TABLE I: RETRIEVAL OF TWITTER ACCOUNTS.

Step i	DE (n = 115)			UK (n = 100)		
	x_i	$\frac{x_i}{n}$	$\frac{x_i}{x_{i-1}}$	x_i	$\frac{x_i}{n}$	$\frac{x_i}{x_{i-1}}$
1. Find account	88	77%	77%	95	95%	95%
2. Targeting	77	70%	88%	77	77%	81%
3. Retrieving	75	65%	97%	73	73%	95%
4. Active Accounts	58	50%	77%	70	70%	96%

TABLE III: COMPARISON OF TWITTER INTERACTION STRATEGIES.

	Strategy	Occurrences	
		DE	UK
S_1	Interactive (Twitter)	19%	80%
S_2	Interactive (other platforms)	24%	0%
S_3	Interactive (mixed)	5%	0%
S_4	Promotional	21%	6%
S_x	No clear strategy	31%	14%

account for 95% of the retailers could be identified for the UK subset, only 77% of the account could be matched for the German subset. While passing 96% of the incoming accounts the activity check for the UK set, only 77% are considered as being active in the German sample.

The variables lifetime, status per day, indegree and out-degree derived from the profile information are summarized in Table II both for Germany and UK. For those values, the .25, .50 and .75 quantiles and Geary's skewness indicator were calculated. For a easier comparison of the location parameters of these values, boxplots visualizing the differences of both samples are depiced in Figure 3. For the sake of readability, the outliers on the right side were cut out, mainly affecting the UK plots.

The German sample is characterized as follows: 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 strongly right-skewed. The majority of accounts issued less than two Tweets per day. The mean increase of indegree is about 1.6 users 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.

The UK sample is characterized as follows: The mean account lifetime is $4^1/2$ years, and the values are slightly left-skewed. Since creation, an account issued on average about 17 statuses per day, whereas the values are right-skewed. The mean increase of indegree is at about 33 users per day, while the values are right-skewed. The average increase of outdegree is slightly above two user per day, which is almost identical to the third quantile.

B. Twitter interaction strategies

In order to analyze interaction strategies, each URL was resolved and classified by the hostname in one of the categories self-link, social media and other for both data sets. The German sample consisted of 5792 Tweets (99.86 per account) with 4661 URLs (79%). 4479 of those links could be resolved (96%) and those pointed to 411 unique hosts. The UK sample consisted of 6997 Tweets (99.96 per account) that contained 2726 URLs (39%), therefrom 2546 (93%) could be resolved. Those URLs were pointing to 431 unique hosts. Both sets differ notably in the overall URL usage in microblogging (UK: 39%, Germany: 80%) which indicates that in the German sample, Twitter is dominantly used to direct users to other content located on the web, whereas among the retailers in the UK, Tweets are more self-contained.

Table III shows the comparison of strategies for Germany and UK. Strategies as described in Section III-C were assigned to each account. 80% of the retailers in UK use one of the interactive strategies (S_1, S_2, S_3), while only 48% of the German retailers are assigned such a strategy. While in the German subset, the values distribute somehow across those three, in the UK sample, the whole category is concentrated on S_1. The promotional strategy S_4 is much more prevalent in the German subset, as well as S_X.

V. DISCUSSION

In Section IV-A, a comparison of the overall activity by Twitter accounts managed by online retail companies is given. The results of the retrieval process distinguish the two samples from each other in two major aspects: The ratio of assigned Twitter accounts is substantially higher in UK compared to Germany, as well as the ratio of active accounts among all accounts. The account lifetime value is used as an indicator for the adoption behaviour among the retailers. The UK sample of accounts is characterized by a slightly longer account lifetime, which indicates an earlier adoption of this technology. Another interesting measure for the adoption is the frequency of usage. According to the number of status messages issued per day, UK retailers are publishing Tweets more frequently. Besides that, a comparison of common influence measures was performed. The data showed that UK retailers are more successful in generating user followers. Another interesting observation is that the UK retailers are also a more likely to follow other accounts.

For the identification of different communication strategies, the use of URLs and UMs in Section IV-B was compared. The collected data shows that among the UK sample, interactive strategies – particularly dialogs inside Twitter – are much more prevalent in the German sample. In return, the ratio of promotional strategies is much higher among German retailers. This is also true for the share of retailers, that could not be assigned a strategical category. Although the model defined in Section III allows the occurrence of multiple strategies per account, there were no accounts that actually applied more than one strategy. This shows that the defined strategies are disjunctive and clearly separated from each other.

But the choosen approach has also several limitations: First, the use of UMs and URLs for the indication of the communication purpose. As a previous work showed, the special communication conventions supported by the Twitter platform are not always used correctly (14). Thus, e.g., a locational "@" can be mistaken as indicator for a directed message by the classification approach. Besides that, only the target of an URL points to was examined and not the content

TABLE II: VARIABLES CHARACTERIZING THE RETAILERS' TWITTER PROFILES.

	min		$Q_{.25}$		$Q_{.50}$		$mean$		$Q_{.75}$		max		$skewness$	
	DE	UK	DE	UK	DE	UK	DE	UK	DE	UK	DE	UK	DE	UK
L	0.893	2.008	3.748	4.334	4.563	4.752	4.139	4.508	4.822	5.055	5.777	6.400	-1.350	-1.0412
R_T	0.057	1.145	0.418	3.632	1.208	5.839	7.711	17.840	1.894	15.510	367.300	458.200	7.345	7.338
R_{in}	0.055	1.576	0.482	11.570	0.987	18.060	1.575	32.980	1.471	35.68	8.258	174.800	2.209	2.094
R_{out}	0.004	0.009	0.068	0.333	0.198	0.804	0.620	2.199	0.629	2.094	8.008	32.940	4.314	4.872

(a) Account lifetime

(b) Status messages per day

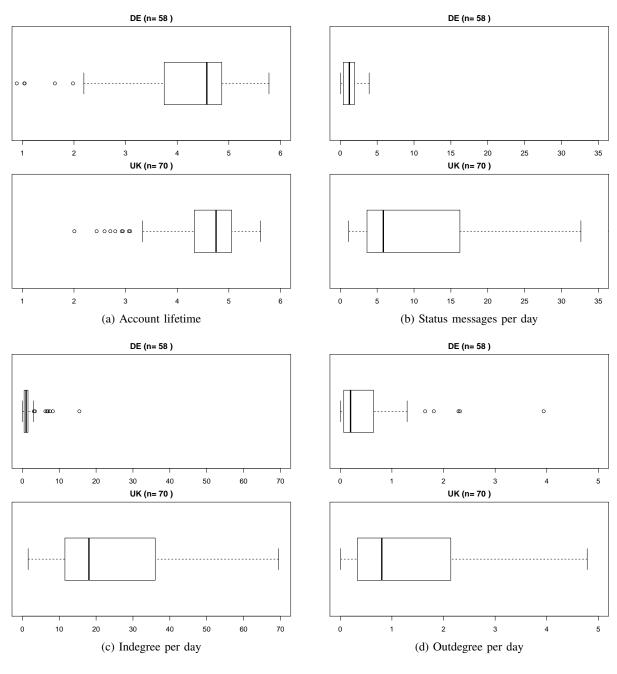

(c) Indegree per day

(d) Outdegree per day

Figure 3: Comparison of account life time and number of status per day.

residing at that location. A link to a post on Facebook linking back to the product site of the shop for example would be falsely classified as a "social link".

VI. CONCLUSION AND FUTURE WORK

This paper contains contributions to the subjects E-commerce and the social web. The comparison of the usage of communication strategies in the Web 2.0 on hand of the example Twitter shows a different adoption between German online retailers and the UK. For studying the usage of Twitter in general this work provides a model for assigning of four basic communication strategies to an examined account. Besides that, an interesting finding is that there are enormous differences in the communication between retailers and customers. The ratio of dialogs on the Twitter platform is much higher among the UK sample. This could reflect the earlier adoption and higher acceptance of Twitter in the UK as other works suggest.

Since there are some limitations, the refinement and improvement of the approach used in this paper will be the next steps. Besides the sole quantitative analysis of communication patterns presented, it is planned to analyze a subset of the data qualitatively to evaluate the precision of the classification process described in Section III-C. Another limitation of this work lies within the composition of the sample's population, consisting of various different business domains. An examination based on a sample narrowed down to single branch will be carried out.

REFERENCES

[1] A. Java, T. Finin, X. Song, and B. Tseng, "Why we twitter: Understanding microblogging usage and communities," 2007, uRL: http://aisl.umbc.edu/get/softcopy/id/1073/1073.pdf [retrieved: 24th April, 2014].

[2] L. Hong, G. Convertino, and E. H. Chi, "Language matters in twitter: A large scale study." in ICWSM, 2011, pp. 518–521.

[3] "Why do germans shun twitter?" 2013, uRL: http://www.economist.com/blogs/babbage/2013/12/social-media [retrieved: 6th March, 2014].

[4] M. Bächle, "Economical perspectives on web 2.0 – open innovation, social commerce and enterprise 2.0 (Ökonomische perspektiven des web 2.0 – open innovation, social commerce und enterprise 2.0)," WIRTSCHAFTSINFORMATIK, vol. 50, no. 2, 2008, pp. 129–132. [Online]. Available: http://dx.doi.org/10.1365/s11576-008-0024-2

[5] A. Richter, M. Koch, and J. Krisch, "Social commerce: An analysis of the change in the e-commere (social commerce: eine analyse des wandels im e-commerce)," 2007.

[6] L. Yang, T. Sun, M. Zhang, and Q. Mei, "We know what @you #tag: Does the dual role affect hashtag adoption?" in Proceedings of the 21st International Conference on World Wide Web, ser. WWW '12. New York, NY, USA: ACM, 2012, pp. 261–270, uRL: http://doi.acm.org/10.1145/2187836.2187872 [retrived: 24th April, 2014].

[7] A. Bifet, G. Holmes, B. Pfahringer, and R. Gavalda, "Detecting sentiment change in twitter streaming data," in Workshop on Applications of Pattern Analysis (WAPA) 2011 Proceedings, 2011, pp. 1–15.

[8] H. Kwak, C. Lee, H. Park, and S. Moon, "What is twitter, a social network or a news media?" in Proceedings of the 19th International Conference on World Wide Web, ser. WWW '10. New York, NY, USA: ACM, 2010, pp. 591–600, uRL: http://doi.acm.org/10.1145/1772690.1772751 [retrieved: 24th April, 2014].

[9] S. Westman and L. Freund, "Information interaction in 140 characters or less: Genres on twitter," in Proceedings of the Third Symposium on Information Interaction in Context, ser. IIiX '10. New York, NY, USA: ACM, 2010, pp. 323–328, uRL: http://doi.acm.org/10.1145/1840784.1840833 [retrieved: 24th April 2014].

[10] S. A. Paul, L. Hong, and E. H. Chi, "Is twitter a good place for asking questions? a characterization study." in ICWSM, 2011, pp. 578–581.

[11] B. Krishnamurthy, P. Gill, and M. Arlitt, "A few chirps about twitter," 2008, uRL: http://www2.research.att.com/~bala/papers/twit.pdf [retrieved: 24th April, 2014].

[12] B. A. Huberman, D. M. Romero, and F. Wu, "Social networks that matter: Twitter under the microscope," 2008, uRL: http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2317/2063 [retrieved: 24th April, 2014].

[13] M. Cha, H. Haddadi, F. Bevevenuto, and K. P. Gummadi, "Measuring user influence in twitter: The million follower fallacy," 2010, uRL: http://snap.stanford.edu/class/cs224w-readings/cha10influence.pdf [retrieved: 24th April, 2014].

[14] C. Honeycutt and S. C. Herring, "Beyond microblogging: Conversation and collaboration via twitter," 2009, uRL: http://ella.slis.indiana.edu/~herring/honeycutt.herring.2009.pdf [retrieved: 24th April, 2014].

[15] H. Edman, "Twittering to the top: A content analysis of corporate tweets to measure organization-public relationships," 2010, uRL: http://etd.lsu.edu/docs/available/etd-04292010-162453/unrestricted/edmanthesis.pdf [retrieved: 24th April, 2014].

[16] About Twitter's link service (http://t.co), uRL: https://support.twitter.com/entries/109623# [retrieved: 24th April, 2014].

[17] D. Kailer, P. Mandl, and A. Schill, "An empirical study on the usage of social media in german b2c-online stores," International journal of advanced Information technology, vol. 3, no. 5, 2013, pp. 1–14.

[18] Top 100 online retailers in the UK 2013, uRL: http://www.digitalstrategyconsulting.com/intelligence/2013/06/top_100_online_retailers_in_the_uk_2013.php [retrieved: 24th April, 2014].

71