SOCIAL BOTS IN A COMMERCIAL CONTEXT – A CASE STUDY ON SOUNDCLOUD

Research in Progress

Ross, Björn, University of Duisburg-Essen, Duisburg, Germany, bjoern.ross@uni-due.de

Brachten, Florian, University of Duisburg-Essen, Duisburg, Germany, florian.brachten@uni-due.de

Stieglitz, Stefan, University of Duisburg-Essen, Duisburg, Germany, stefan.stieglitz@uni-due.de

Wikström, Patrik, Queensland University of Technology, Brisbane, Australia, patrik.wikstrom@qut.edu.au

Moon, Brenda, Queensland University of Technology, Brisbane, Australia, brenda.moon@qut.edu.au

Münch, Felix, Queensland University of Technology, Brisbane, Australia, felixvictor.muench@hdr.qut.edu.au

Bruns, Axel, Queensland University of Technology, Brisbane, Australia, a.bruns@qut.edu.au

Abstract

Recently, automated communication on social media has seen increased attention. Social bots, social media accounts controlled by algorithms that mimic human behaviour, have been found to attempt to influence users in several political contexts. However, their use in a commercial context, e.g. to boost sales of a product by aggressively promoting it with thousands of messages, has so far been neglected. To address this shortcoming, this paper examines the case of the social media music platform SoundCloud. We gathered a dataset of six months of activity, comprising 15,850,069 tracks and 12,125,095 comments. We then calculated a comment uniqueness score for highly active accounts to assess the variability of their comments. First analyses show that some accounts post suspiciously repetitive comments. These accounts also frequently repost existing content, but contribute little original content. An analysis of the commenting network further underlines that these accounts differ clearly from regular users. We conclude that the comment uniqueness metric can be used as an indicator to distinguish bots from humans, and that a considerable proportion of SoundCloud comments are likely to emanate from bots or semi-automated accounts. The implications of these findings and future plans are discussed.

Keywords: social bots, SoundCloud, social media, automated communication, social network analysis.

1 Introduction

Social media platforms have become an important source for customers to inform themselves about products and services they are interested in (Gu, Park, and Konana, 2012). Much research has shown that product information which is shared by other customers has a larger effect on buying decisions than promotions and advertisements initiated by the selling companies (Zhou and Duan, 2016). As a consequence, there is a growing interest in influencing customer reviews and the visibility of products and services on social media platforms. While those are crucial to the success of companies (especially if they generate the majority of revenue on the internet), findings from research on automated communication on social media indicate that a part of the seemingly human-generated content is actually produced by software algorithms, called social bots (Varol et al., 2017). Findings from the field of politics show that bots apparently try to influence the opinion climate on certain topics (Hegelich and Janetzko, 2016). Simultaneously, research on whether social bots try to influence customers of companies is sparse. Given their apparent pervasiveness in political contexts, it seems plausible that they are also being used for commercial purposes, but the extent of their usage and their specific aims are unclear. To address this gap, this paper examines the application of bots in a commercial context using the example of the music industry and more specifically the social network SoundCloud (see also Bruns et al., 2018).
This short paper presents preliminary results on the following research questions:

1. How can bots be differentiated from humans on the music sharing platform SoundCloud?
2. How do these bots behave, especially how do they interact with tracks and human users?

To address these questions, it is first necessary to develop a bot identification method for this platform. We present a simple metric that separates the accounts into two distinct groups and we then apply it on a data set spanning six months of activity from SoundCloud. The analysis of this data reveals that the two groups of accounts also differ in several other behavioural attributes. One of the groups closely resembles what is known about bots, underlining the validity of the proposed identification method. In a network analysis, we also find that the bot-like accounts focus on tracks that were uploaded by specific accounts, suggesting that they might be used to boost the visibility of these accounts. This research is part of a larger study with the overall goal of examining the relevancy of these bots to the music industry, that is, to determine whether or not they are successful in influencing consumers' purchasing and streaming behaviour.

2 Background

2.1 Social bots

Social bots have seen a rise in attention in recent years, especially on Twitter (Stieglitz et al., 2017). Social bots are accounts that are based on 'computer algorithms that automatically produce content and interact with humans on social media, trying to emulate and possibly alter their behavior' (Ferrara et al., 2016, p. 96). The assumption that these accounts try to influence the human users' behaviour is one of the main drivers of research on this topic, as authors try to assess the actual importance of bots. One of the main domains of research has been the use of these bots in a political context (e. g. Abokhodair, Yoo, and McDonald, 2015; Brachten et al., 2017; Hegelich and Janetzko, 2016). While a possible influence in other scenarios may have limited impact, bot influence on politics, especially elections, may in theory alter their outcomes and thus affect the lives of many citizens. Elections with a global impact, such as the 2017 French presidential election (Ferrara, 2017) or the 2016 US presidential election (Bessi and Ferrara, 2016), are among the events researched in the political context, as are elections with a smaller scale or regionally limited importance such as the 2010 US midterm elections (Ratkiewicz et al., 2011) or 2017 German state elections (Brachten et al., 2017). Besides these scenarios, bots have also been researched in political conflicts such as the civil war in Syria (Abokhodair, Yoo, and McDonald, 2015) or the Ukrainian crisis (Hegelich and Janetzko, 2016). All of the studies found at least some use of social bots – and often, their strategies could be observed. The studies showed that bots simulate broad agreement on certain

topics, even though this consensus might not be prevalent in the general population (a strategy often called *astroturfing*), or try to hijack a hashtag by posting unrelated messages with the hashtag, thus misdirecting human users that search using these hashtags (Abokhodair, Yoo, and McDonald, 2015).
While this behaviour could be observed, the actual ability of bots to alter human behaviour has not been shown on a larger scale. There are, however, findings which suggest that human users treat social bots as they would other human users (Edwards et al., 2016), which suggests that these bots have the same potential to influence human behaviour as other human users. Furthermore, it could also be shown that, under certain circumstances, bots that directly address human users regarding a specific topic were able to alter the future behaviour of these users on social media (Munger, 2017). Some research also indicates that the number of sources of a message (i.e. the number of social bots that spread a message) rather than the number of messages from a single source determines the probability for this message to spread throughout a network (Mønsted et al., 2017). The authors used positive, encouraging messages to test their model, and state that the findings might have implications not only in a political context but also for improving marketing and advertising strategies. While the latter domain seems to hold great potential for social bot usage, little to no research in this context has been conducted. As the possibility of altering the behaviour of human users and of spreading messages in favour of a specific topic seems to exist, findings also suggest that user-generated content on social media can be used to point consumers to a particular product (i.e. make the target group aware that the product exists) and, in the form of user recommendations, affect users' purchasing intentions (Aggarwal and Singh, 2013; Bai, Yao, and Dou, 2015). The potential for misuse of automated communication in this context as well as the need for research in this area has become apparent.
A domain which seems suitable for this purpose is the music industry. In contrast to mere online retailers, there are social media music platforms on which a lot of user interactions take place. These platforms are often as unregulated as Twitter, on which most findings on social bots rely (Stieglitz et al., 2017). In addition, previous research has shown that social media activity has significant effects on music sales (Chen, De, and Hu, 2015).

2.2 Social media and the music industry

The digital transformation of the music economy is generally considered to have started in 1999 when peer-to-peer filesharing services emerged as a mainstream (albeit illegal) mechanism for online music distribution (Alderman, 2008; Knopper, 2009). During the years that followed, the digital transformation primarily affected music distribution and dramatically improved the consumers' access to music (Wikström, 2012). While at first analogue broadcast radio remained the most influential tastemaker, social media platforms soon emerged, such as MySpace in 2003 and Facebook, Twitter and YouTube in 2006, and they gradually took over that role (Mjos, 2013; Wikström, 2013). This logic has been further reinforced as music streaming platforms that serve as social media platforms in themselves (i.e. YouTube and SoundCloud) have become the de facto standard for music distribution (ibid.).
As these social media platforms increasingly gain influence as tastemakers in the music economy, the economic incentives also increase to try to influence the communication dynamics on these platforms in one direction or another.

3 Method

3.1 Data

SoundCloud is a music streaming platform headquartered in Berlin, Germany, that was launched in 2007 (Mahroum, 2016). The service allows its users to upload, listen and comment on audio files of all kinds (not only songs but also speeches and podcasts). While the platform has struggled to establish a sustainable business model (Cook, 2017), it has become one of the most widely used social music platforms with 40

	Descriptive statistics[1]			Correlations			
	% Zero[2]	Mean	Max	Downld.	Fav.	Reposts	Comments
Playback count	4.7	1,119.15	116,423,945	.260[3]	.842[1]	.727[1]	.323[4]
Download count	87.2	1.86	92,535		.343[3]	.268[3]	.201[5]
Favouritings count	34.7	20.39	1,447,597			.849[1]	.328[4]
Reposts count	79.5	3.22	127,565				.514[4]
Comment count	85.8	0.96	15,132				

[1]All tracks (n = 15,850,069) [2]Percentage of tracks with value 0
[3]Downloadable tracks only (n = 2,409,814) [4]Commentable tracks only (n = 15,781,151)
[5]Downloadable and commentable tracks only (n = 2,366,729)

Table 1. Descriptive statistics and correlation table for track interactions on SoundCloud

million registered users (July 2013) and 175 million unique monthly listeners (December 2014) (Walker, 2015). While there are other social music platforms, none of these platforms have been able to challenge SoundCloud in terms of user numbers. This pivotal position of SoundCloud in the music economy and the social media features offered by the platform are the main reasons that the platform has been selected as the empirical context for this paper. In combination, they create an exceptional opportunity for studying the prevalence of bots on a large online platform and the bots' capability to influence the behaviour of the human users on the platform.

Social media analytics (Stieglitz et al., 2018) was used to analyse the behaviour of SoundCloud users. We collected the metadata of all publicly accessible tracks uploaded to SoundCloud between 1 January and 30 June 2017, including the users who uploaded them, all comments on these tracks and the authors of the comments. The data was collected through SoundCloud's restricted HTTP API[1], which we gained access to after an application. This API allows the crawling of tracks uploaded at an arbitrary point in the past. Data collection took place from July to September 2017. Each track was crawled at least a month after it was uploaded. To ensure that the metadata of tracks can be compared regardless of when they were uploaded, we only consider the first month of comments in the analysis.

The final dataset comprises a total of:

- **15,850,069 tracks**, including the number of times a track has been played, downloaded, favourited, reposted or commented on. Favouriting refers to publicly marking a track as a favourite, and reposting refers to adding a track to one's profile page to share it with one's followers. Each account can only favourite or repost a track once. Not all tracks permit downloading or commenting. While 15,781,152 tracks (99.6%) in the dataset are commentable, only 2,409,814 tracks (15.2%) can be downloaded. The metrics for these fields were calculated based only on the tracks that allow commenting or downloading, respectively. Descriptive statistics for the track interaction metrics are given in Table 1.
- **12,125,095 comments** on the tracks, including the date and time of each comment, the account making the comment and the text of the comment.
- **1,983,846 users** who wrote the comments, including the number of followers, the number of accounts being followed (at most 2,000), the number of playlists created, the number of tracks reposted and the number of tracks uploaded.

To understand how users' interactions with tracks are related, consider the correlations of these metrics (Table 1). These correlations mean that the data are consistent with our assumption that comments on tracks are associated with other forms of interaction. While the correlations are of course not enough to establish a sequence of events or causation, such as more comments on some tracks by bots leading to more plays by humans, this first result is enough to warrant further investigation.

[1] https://developers.soundcloud.com/docs/api/

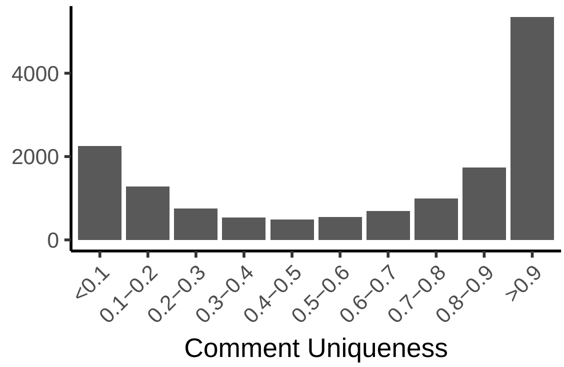

Figure 1. Number of users in each range of comment uniqueness. Only the 14,600 users who have written at least 100 comments are considered.

3.2 Comment uniqueness as a measure of behavioural complexity

It has been suggested that there are active bots on SoundCloud (Jordanous, Allington, and Dueck, 2015). These bots might interact with the tracks in several ways, such as download them, repost them or comment on them. For most forms of interaction, only aggregate counts are available. More detailed information can be examined for the comments, including the date, time, content and author of a comment. A preliminary look at the comments suggested that some of them seemed to be very repetitive. We therefore define a metric that captures how unique the comments posted by a user are. This method is similar to the identification of bots by examining the entropy of posting intervals. A high entropy reflects the behavioural complexity and unpredictability hypothesised to be typical of humans (Chu et al., 2012; Ghosh, Surachawala, and Lerman, 2011). Likewise, we propose an identification method based on the predictability of the texts of users' comments.

To measure the diversity of the comments by a user, we divide the number of distinct comments a user has made by the total number of comments k by the user. We call the result comment uniqueness (CU). Two comments are considered different if they differ by at least one character. Formally, let $C = (c_1, c_2, \ldots c_k)$ denote the tuple of comments posted by a user. Then,

$$CU = \frac{|\{c \in C\}|}{k}.$$

Comment uniqueness takes values between $1/k$ and 1. It will equal 1 if every comment by a user is different from the others. In contrast, it will equal $1/k$ if all of a user's comments are identical, and approach zero the more identical comments are made.

As described above, we assume that automated accounts post repetitive comments whereas contributions by humans are more diverse. This assumption does not hold for bots that generate arbitrary sequences of characters and post them as comments, or generate comments from complex rules. However, we assume that many bots follow simplistic rules such as choosing one out of a number of possible predefined comments (e.g. 'nice', 'great song', 'cool'). This type of activity would result in a very low comment uniqueness.

The complexity of a user's commenting behaviour can only be estimated reliably if the user has posted sufficient comments. For example, three identical comments are arguably not enough to reliably determine whether or not a user will always post repetitive comments. The minimum threshold used in the following is 100. It was deliberately set high, to examine whether there are accounts which can be confidently classified as suspicious. The 14,600 users with at least 100 comments only make up a tiny fraction (0.74 %) of all SoundCloud users, but are together responsible for almost a third (3,781,195 or 31.18 %) of comments.

Figure 1 shows that, contrary to what one might expect, these highly active users are divided into two groups by the comment uniqueness metric while only few occupy the space in between those groups. This

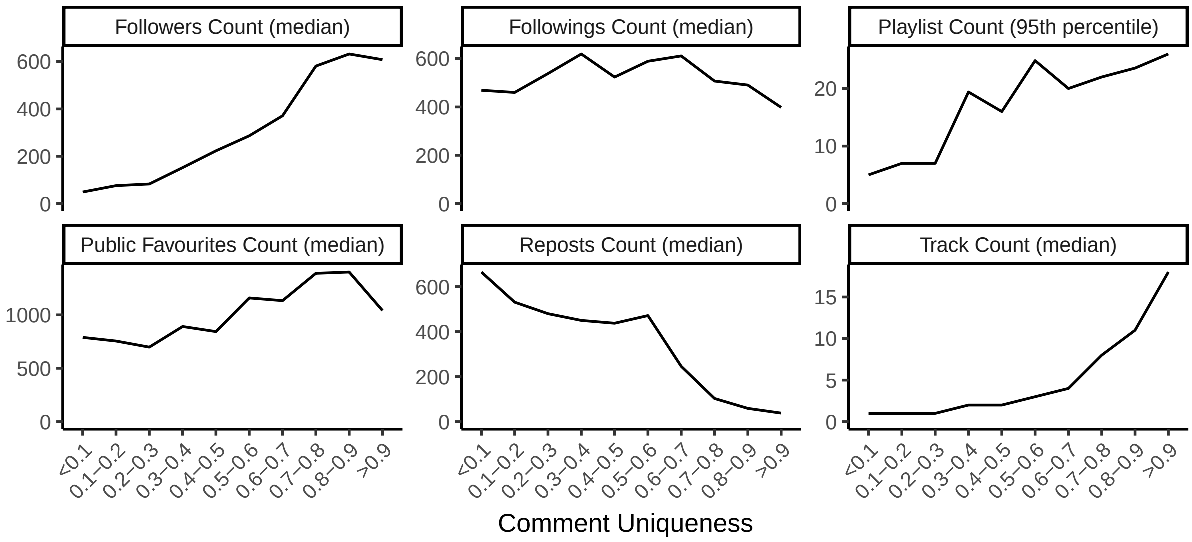

Figure 2. Statistics for users by comment uniqueness. Only users who have written at least 100 comments are considered. The median is used because it is less sensitive to outliers than the mean. With playlist count, the 95th percentile is used because the median is 0 in most groups.

suggests that comment uniqueness indeed offers a useful way of distinguishing between groups of users. For further analyses, such as visualising the user types in a network, a threshold to separate the groups is needed. A simple option is the global comment uniqueness, i.e., the number of distinct comments in the data set divided by the total number of comments ($5{,}706{,}663/12{,}125{,}095 = 0.471$).

4 Preliminary Results

4.1 Behaviour of the two groups

The previous section showed how the highly active users of SoundCloud can be divided into two separate groups based on their commenting behaviour. Not every active SoundCloud user can be clearly assigned one of these categories; instead, the users form a continuum between two extremes. However, the data in Figure 1 showed that most users are much closer to one of these two archetypes than the other. Next, we examine these extremes more closely to see how they differ in other behaviour.

Figure 2 compares comment uniqueness with other metrics of user behaviour. Users with a lower comment uniqueness (that is, those we consider potential bots) tend to have far fewer followers (people who are interested in updates about what they are doing). Between the two extremes, this difference is very pronounced, as the accounts with $CU > 0.9$ have 12.4 times the median number of followers as accounts with $CU > 0.1$ (608/49). They also upload far fewer tracks (with medians of 18 and 1, respectively) and create fewer playlists (although many active users do not create playlists at all). In contrast, they repost 17.5 times as many tracks (665/38). They also follow about the same number of people and publicly favourite about the same number of tracks.

We conclude that the comment uniqueness metric results in a meaningful distinction between two types of highly active users (i.e. those with 100 or more comments in the examined six-month timespan):

- The first type comprises those who actively participate in the community by writing varied comments, uploading tracks, and sometimes creating playlists, and who perhaps as a result have a high number of followers.
- The second type consists of those who are active, but write highly repetitive comments and mainly repost tracks uploaded by others, and who therefore do not contribute to the community as much.

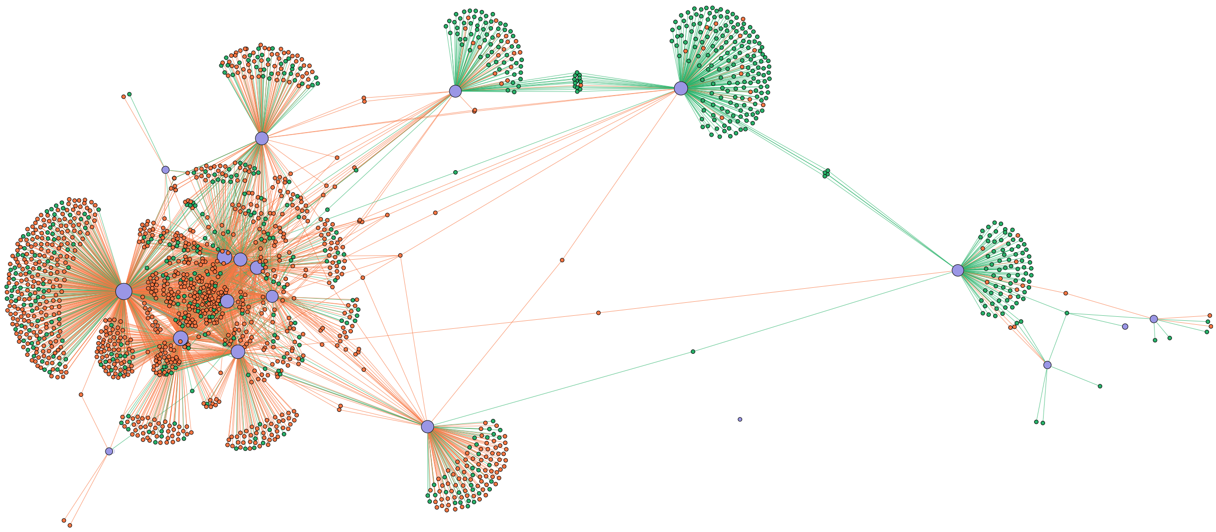

Figure 3. Network visualisation of the 25 most commented tracks on SoundCloud. Nodes represent users, edges comments. The 19 purple nodes are the track authors (some authors uploaded more than one of the 25 most commented tracks). Other users are only included if they made at least 100 comments (across all tracks). Users who post highly repetitive comments (CU < 0.471) are coloured red, users who post diverse comments (CU > 0.471) green.

4.2 Network visualisation

Given that the users can be divided into two groups which differ clearly in several behavioural attributes, the next question is: Do members of these groups interact with the same tracks or do the groups represent separate communities? To address this question, we visualised the network of highly active users and their comments (Figure 3) using the open source network visualisation package Gephi (Bastian, Heymann, Jacomy, et al., 2009). Several observations can be made.

At the centre of the network is a group of eight track authors (artists or publishers, shown as large nodes) whose tracks are frequently commented on by the same users. In this central group, the share of commenting 'bots' (red nodes) is much higher than for the other track authors. In addition, many of the 'bots' commented on tracks by several of the central authors, while many of the humans only commented on one. The other track authors have few or no commenting users in common with each other, or with this central group. In addition, the remaining track authors are mostly commented on by humans.

5 Discussion

5.1 Implications

We identified two archetypes of accounts on SoundCloud. The highly active users on the platform form two groups, which more or less correspond to the archetypes. One of these groups represents about a third of the highly active users, and they aggressively post repetitive comments. To validate the comment uniqueness metric, we examined the behaviour of these accounts in terms of other types of interactions (see Figure 2). Their low number of followers and lack of original content conform with prior findings on bot activity on social media (Ferrara et al., 2016; Ratkiewicz et al., 2011).

Our early results relying on comment uniqueness thus indicate that there is a substantial number of bots active on SoundCloud. These bots do not contribute to the platform in the way humans do, and in the way the platform was intended: by uploading tracks or curating content through the creation of playlists, both of which attract followers. The comment network for the top 25 most commented tracks (Figure 3) shows that the accounts exhibiting bot-like behaviour (shown in red) are highly focused on some of the tracks. It is therefore plausible that they might be attempting to influence the discussion on these tracks to prop them up artificially or to increase the visibility of the commenting accounts themselves. This might help artists gain an unfair advantage over other artists who do not engage in such activities. However, there are very few 'human' accounts active for these tracks, so it is not clear whether the bots are successful in influencing activity by non-bot accounts.

Our findings have important implications for SoundCloud and the wider music industry, since a considerable amount of activity seems to be generated by bots or semi-automated accounts. This raises the question, for example, to what extent success on SoundCloud – in terms of the sheer number of interactions – depends on technical expertise rather than musical creativity. This finding also has implications for other domains, since bots could equally be used in other commercial contexts to promote goods, including other social networks and other types of products. SoundCloud, through its open API and large userbase, simply made it possible to study this phenomenon in a relevant context.

5.2 Limitations and outlook

The presented preliminary findings have several limitations which will be addressed in further research. Although the introduced comment uniqueness measure yields promising results, it currently only identifies one kind of bot behaviour, namely repetitive comments. Furthermore, human-operated accounts with a simple commenting behaviour could lead to false positives, while the comment uniqueness measure is not able to detect bots with more sophisticated, randomised routines either. To do so, we will examine other attributes and factors. Triangulating and complementing our results obtained with the comment uniqueness measure, for example with entropy measures regarding content and timing of comments, is the logical next step.

Furthermore, our network visualisation of comment activities is promising: a more detailed analysis of the comment network, as well as collecting the follower network of a sample of active users will provide us with more possibilities to unveil bot-like behaviour. Our results already point at differences in following behaviour, as the following count seems to be uniform between both groups of accounts while the follower count increases with comment uniqueness. Therefore, we are confident that community detection in combination with density, clustering, and centrality measures will offer another possibility to further support and extend our results.

In addition, it is not clear whether these suspected bots are successful at convincing human accounts that a track is more popular than it is, and thereby causing them to become more active on the tracks themselves. To examine this causal relationship, we plan to carry out a time series analysis that can establish the order of events, that is, whether comments by bots lead to subsequent comments by non-bot accounts. Additionally, we will use time-series data to determine if the selected period of one month from upload is optimal for a sample of track activity.

Finally, the collected data does not contain deleted accounts and their activities, because the SoundCloud API does not provide an archive. Therefore, bots already deleted by SoundCloud are excluded from our analysis. The true extent of bot activity is probably even higher than what we are able to observe.

6 Conclusion

Summarising, we have found a simple method to differentiate between two groups of SoundCloud accounts which exhibit characteristic behaviours. Further analysis, for example regarding the active contribution to the platform in the form of track uploads or curation of playlists, supported our interpretation of these groups exhibiting bot-like or human-like behaviour respectively. This strong evidence for the widespread existence and visibility of automated accounts on SoundCloud is by itself relevant and strongly suggests that further efforts in detecting and investigating these automated accounts are needed. The manipulation of activity records arguably impacts platform dynamics, thereby affecting the success of tracks. This not only has implications regarding the user experience of SoundCloud, but also cultural, monetary, and economical consequences for the platform, its users, and artists.

References

Abokhodair, N., D. Yoo, and D. W. McDonald (2015). “Dissecting a Social Botnet.” In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing – CSCW ’15*, pp. 839–851. ISBN: 9781450329224. DOI: 10.1145/2675133.2675208.

Aggarwal, R. and H. Singh (2013). “Differential influence of blogs across different stages of decision making: The case of venture capitalists.” *MIS Quarterly* 37 (4), 1093–1112.

Alderman, J. (2008). *Sonic boom: Napster, MP3, and the new pioneers of music*. London: Fourth Estate.

Bai, Y., Z. Yao, and Y.-F. Dou (2015). “Effect of social commerce factors on user purchase behavior: An empirical investigation from renren.com.” *International Journal of Information Management* 35 (5), 538–550. DOI: 10.1016/j.ijinfomgt.2015.04.011.

Bastian, M., S. Heymann, M. Jacomy, et al. (2009). “Gephi: an open source software for exploring and manipulating networks.” In: *Proceedings of the Third International Conference on Weblogs and Social Media (ICWSM-2009)*. Vol. 8, pp. 361–362.

Bessi, A. and E. Ferrara (2016). “Social bots distort the 2016 U.S. Presidential election online discussion.” *First Monday* 21 (11). ISSN: 13960466. DOI: 10.5210/fm.v21i11.7090.

Brachten, F., S. Stieglitz, L. Hofeditz, K. Kloppenborg, and A. Reimann (2017). “Strategies and Influence of Social Bots in a 2017 German state election.” In: *Proceeedings of the Australasian Conference on Information Systems*, Paper 207.

Bruns, A., B. Moon, F. Münch, P. Wikström, S. Stieglitz, F. Brachten, and B. Ross (2018). “Detecting Twitter Bots That Share SoundCloud Tracks.” In: *Proceedings of the International Conference on Social Media + Society*.

Chen, H., P. De, and Y. Hu (2015). “IT-enabled broadcasting in social media: An empirical study of artists’ activities and music sales.” *Information Systems Research* 26 (3), 513–531. DOI: 10.1287/isre.2015.0582.

Chu, Z., S. Gianvecchio, H. Wang, and S. Jajodia (2012). “Detecting automation of twitter accounts: Are you a human, bot, or cyborg?” *IEEE Transactions on Dependable and Secure Computing* 9 (6), 811–824. DOI: 10.1109/TDSC.2012.75.

Cook, J. (2017). *SoundCloud raised $70 million in debt funding*. URL: http://uk.businessinsider.com/soundcloud-raises-70-million-debt-funding-companies-house-documents-2017-3 (visited on 10/13/2017).

Edwards, C., A. J. Beattie, A. Edwards, and P. R. Spence (2016). “Differences in perceptions of communication quality between a Twitterbot and human agent for information seeking and learning.” *Computers in Human Behavior* 65 (Supplement C), 666–671. ISSN: 0747-5632. DOI: https://doi.org/10.1016/j.chb.2016.07.003.

Ferrara, E., O. Varol, C. Davis, F. Menczer, and A. Flammini (2016). “The rise of social bots.” *Communications of the ACM* 59 (7), 96–104. DOI: 10.1145/2818717.

Ferrara, E. (2017). “Disinformation and social bot operations in the run up to the 2017 French presidential election.” *First Monday* 22 (8). ISSN: 13960466. DOI: 10.5210/fm.v22i8.8005.

Ghosh, R., T. Surachawala, and K. Lerman (2011). “Entropy-based Classification of ‘Retweeting’ Activity on Twitter.” In: *Proceedings of the KDD Workshop on Social Network Analysis (SNA-KDD)*.

Gu, B., J. Park, and P. Konana (2012). “The impact of external word-of-mouth sources on retailer sales of high-involvement products.” *Information Systems Research* 23 (1), 182–196. DOI: 10.1287/isre.1100.0343.

Hegelich, S. and D. Janetzko (2016). “Are social bots on Twitter political actors? Empirical evidence from a Ukrainian social botnet.” In: *Proceedings of the 10th International Conference on Web and Social Media, ICWSM 2016*, pp. 579–582. ISBN: 9781577357582.

Jordanous, A., D. Allington, and B. Dueck (2015). “Measuring cultural value using social network analysis: A case study on valuing electronic musicians.” In: *Proceedings of the Sixth International Conference on Computational Creativity*, pp. 110–117.

Knopper, S. (2009). *Appetite for self-destruction: The spectacular crash of the record industry in the digital age*. New York: Free Press.

Mahroum, S. (2016). "SoundCloud in Berlin, Sharing the Vibe." In: *Black Swan Start-ups: Understanding the Rise of Successful Technology Business in Unlikely Places*. London: Palgrave Macmillan UK, pp. 59–72. ISBN: 978-1-137-57727-6. DOI: 10.1057/978-1-137-57727-6_4.

Mjos, O. J. (2013). *Music, social media and global mobility: MySpace, Facebook, YouTube*. New York: Routledge Chapman & Hall.

Mønsted, B., P. Sapieżyński, E. Ferrara, and S. Lehmann (2017). "Evidence of complex contagion of information in social media: An experiment using Twitter bots." *PLOS ONE* 12 (9), 1–12. DOI: 10.1371/journal.pone.0184148.

Munger, K. (2017). "Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment." *Political Behavior* 39 (3), 629–649. ISSN: 1573-6687. DOI: 10.1007/s11109-016-9373-5.

Ratkiewicz, J., M. D. Conover, M. Meiss, B. Gonçalves, A. Flammini, and F. Menczer (2011). "Detecting and Tracking Political Abuse in Social Media." In: *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, pp. 297–304. eprint: 1011.3768.

Stieglitz, S., M. Mirbabaie, B. Ross, and C. Neuberger (2018). "Social media analytics – Challenges in topic discovery, data collection, and data preparation." *International Journal of Information Management* 39, 156–168. DOI: 10.1016/j.ijinfomgt.2017.12.002.

Stieglitz, S., F. Brachten, B. Ross, and A.-K. Jung (2017). "Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts." In: *Proceedings of the Australasian Conference on Information Systems*, Paper 206.

Varol, O., E. Ferrara, C. A. Davis, F. Menczer, and A. Flammini (2017). "Online Human-Bot Interactions: Detection, Estimation, and Characterization." In: *Proceedings of the Eleventh International Conference on Web and Social Media*, pp. 280–289.

Walker, R. (2015). *Can SoundCloud Be the Facebook of Music?* URL: https://www.bloomberg.com/news/features/2015-07-10/can-soundcloud-be-the-facebook-of-music- (visited on 10/13/2017).

Wikström, P. (2012). "A typology of music distribution models." *International Journal of Music Business Research* 1 (1), 7–20.

— (2013). *The music industry: Music in the cloud*. Cambridge: Polity Press.

Zhou, W. and W. Duan (2016). "The sales impact of word-of-mouth distribution across retail and third-party websites." In: *Proceedings of the Thirty-seventh International Conference on Information Systems (ICIS)*.

Acknowledgements

This work was supported in part by the UA-DAAD project "Emergent Music Engagement Practices via SoundCloud" and in part by the Deutsche Forschungsgemeinschaft (DFG) under grant No. GRK 2167, Research Training Group "User-Centred Social Media".