
Cross-country User Connections in an Online Social Network for Music

Christine Bauer

Johannes Kepler University Linz
Linz, Austria
christine.bauer@jku.at

Markus Schedl

Johannes Kepler University Linz
Linz, Austria
markus.schedl@jku.at

ABSTRACT

Social connections and cultural aspects play important roles in shaping an individual's preferences. For instance, people tend to select friends with similar music preferences. Furthermore, preferences and friending are influenced by cultural aspects. Recommender systems may benefit from these phenomena by using knowledge about the nature of social ties to better tailor recommendations to an individual. Focusing on the specificities of music preferences, we study user connections on Last.fm—an online social network for music. We identify those countries whose users are mainly connected within the same country, and those countries that are characterized by cross-country user connections. Strong cross-country connection pairs are typically characterized by similar cultural, historic, or linguistic backgrounds, or geographic proximity. The United States, the United Kingdom, and Russia are identified as countries having a large relative amount of user connections from other countries. Our results contribute to understanding the complexity of social ties and how they are reflected in connection behavior, and are a promising source for advancements of personalized systems.

KEYWORDS

online social networks; cross-country user connections; friends; social ties

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI'19 Extended Abstracts, May 4–9, 2019, Glasgow, Scotland UK

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5971-9/19/05.

<https://doi.org/10.1145/3290607.3312831>

INTRODUCTION

Social connections play an important role in shaping an individual’s taste and preferences [9]. For instance, people tend to select friends with similar life style and music taste, and also adopt their friends’ preferences [13]. Recommender systems may benefit from this phenomenon and use knowledge about social ties to better tailor recommendations to an individual [18].

In this work, we focus on the music domain (in particular, analyzing the OSN for music Last.fm (www.last.fm), with the goal to use findings in personalized (music) recommender systems. Besides social ties, many other factors influence an individual’s music taste and preferences including, for instance, demographics [10] or personality traits [5]. In user modeling and recommender systems research, cultural aspects—as drivers of music taste and preferences—have only recently started gaining considerable attention. With country-specific differences considered as a proxy for national culture, studies have found strong relationships between culture and music preferences (e.g., [15, 17]).

As social ties are always set in a particular social setting, it may be particularly beneficial for personalized systems to put both influencing factors—social ties and culture—into context. Yet, social ties and cultural aspects have been studied in separate research streams. The novel asset of our work is that we interweave these topics. Thereby, the objective of our ongoing research is to leverage the interweaving of studying social ties and cultural aspects to ultimately advance personalized systems. The overall research question is: Which countries are particularly important for social ties and are, thus, relevant to be considered for taste and preferences for personalized systems?

With the aim to exploit findings for personalized systems, this paper approaches the following specific research questions presented in Sidebar 1.

RQ1: Which are the countries whose users show a high proportion of within-country user connections? Which are the ones that show transnational behavior?

RQ2: What are the most important “attractor” countries, i.e., countries whose users are substantially more often the target of a user connection than other countries (in relative numbers)?

Sidebar 1: Research Questions (RQs)

RELATED WORK

Basically, our work on cross-country user connections in an OSN for music bridges several threads of research, including the role of shared music preferences in friending, social connectedness in OSN, and the nature of within-country and cross-country user connections on OSN. We briefly discuss the most important related literature in these areas.

The foundation of our work is research showing that shared music preferences create bonds between people [4] and that user connections play a role in shaping music preferences [9], for instance, as friendships and music taste tend to co-evolve [16]. Further, research in the field studies friendship strength prediction based on music preferences [3] or focuses on the social influence in music OSN [7].

Already early studies on social connectedness in OSN reported that users tend to manifest and intensify locally bounded social circles also in their online interactions [19], although OSN enable users to transcend geographical borders [6].

Reasons for focusing on the online social network for music: Last.fm (www.last.fm).

- User connections on Last.fm are not yet studied in detail; in contrast to Facebook (e.g., [2]) and Twitter (e.g., [11]).
- Focusing on Last.fm allows for exploiting a rigorous publicly available dataset (i.e., the LFM-1b dataset [14]), while with other OSN we would have to deal with unbalanced user samples with less coverage (because—from outside—we do not have access to those OSN’s complete sets of user connections).
- As Last.fm is a domain-specific OSN that accumulates the music taste of its users over the lifetime of their accounts, findings may be particularly beneficial for the advancement of music recommender systems.
- In contrast to Twitter, user connections on Last.fm are always bidirectional or commutative, i.e., if a user u_1 is connected to u_2 , user u_2 is also connected to u_1 .

Sidebar 2: Why we focus on the online social network for music Last.fm

¹ Note that rows do not sum up to 100% as only the top 20 countries are depicted.

Another relevant research thread studies proposes various methodological approaches to study cross-country user connections. For instance, Nash [12] examined the connection patterns of individual OSN users, whereas Barnett and Benefield [2] used an aggregation of individual cross-country connections per country to measure the relational strength between countries. Bailey et al. [1] analyzed connections across counties within the United States (US), but also connections between US counties and foreign countries.

METHODS AND MATERIALS

In order to assess the cross-country relationships between users, we investigate user connections on Last.fm. We take an approach similar to Bailey et al. [1] to calculate these relationships for pairs of countries. For each pair c_1 and c_2 , we compute the share of users in c_1 that are connected to users in c_2 . This yields a (per-row) normalized *country connection matrix* (cf. Table 1), which is asymmetric.

Using this country connection matrix (Table 1), we approach our research questions (Sidebar 1) as follows: To answer RQ1, for each country c , the share of user connections maintained with other users in c is compared to the share maintained with users from other countries on a per-country basis. Addressing RQ2, we define an *attractor measure* for a country c . This measure models the (relative) amount of users from countries other than c that are attracted to establish connections with users in c . More precisely, we define the attractor measure as the median of all shares of user connections from other countries maintained with users in c . In other words, we compute the median over all rows in the country connection matrix for the column representing country c , resulting in the attractor measure, which is given in the last row of Table 1. (Note that the country connection matrix is asymmetric and therefore allows to investigate which countries serve as an attractor for other countries.) We use the median instead of the mean to correct for outliers, e.g., if only one or two countries account for a vast share of user connections to country c , we do not consider c as a global attractor.

For various reasons (cf. Sidebar 2), we use the publicly available LFM-1b dataset [14]. Because we are interested in the cross-country user relationships, we exclude all users for whom the dataset does not report country information. Then, we use the Last.fm API endpoint *user.getFriends* (<https://www.last.fm/api/show/user.getFriends>) to obtain all connected users. We cross-match these friends with the users in the LFM-1b, which eventually yields a total of 55,191 users and 1,087,662 user-user connections. For our analysis, we consider the top 20 countries in terms of total number of users.

RESULTS AND DISCUSSION

Table 1 shows the cross-country user connections, in relative numbers, for Last.fm. For a given row (country c), the values denote the share of connections by users in c maintained with users in all countries (columns).¹ For example, Table 1 shows that 16.47% of user connections from Australia (first row) are maintained with users in the United States, but only 0.69% with users in Japan. Likewise,

the diagonal elements contain the percentage of within-country user connections, i.e., connections maintained between users within the same country.

As concerns RQ1, Table 1 reveals that the countries with the most *within-country connections* are Poland (76.87%) and Brazil (76.04%), followed—with a large margin—by Finland (65.27%), the Czech Republic (64.21%), and Russia (62.80%). The *most transnational* countries (i.e., having the largest proportions of cross-country connections) are Canada (29.01%) and France (38.60%).

With respect to *strong cross-country connections*, Belorussian and Ukrainian users are particularly well connected to Russians (17.12% and 19.22%, respectively). Russians themselves, in contrast, do not connect to that extent to Belorussians (1.93%) and Ukrainians (4.84%).² Canadian users connect a lot to users in the United States (26.46%). So do users in Australia (16.47%) and the United Kingdom (13.37%). In contrast, users from the United States show a substantial share of cross-country connections only to users from the United Kingdom (6.29%). We observe that the connected countries share similar cultural, historic, or linguistic backgrounds, or geographic proximity, which seems to be reflected in their connections. This is largely in line with findings in [2] analyzing connection behavior on Facebook and [8] comparing connection behavior on Facebook, Twitter, and offline.

Addressing RQ2, we illustrate the attractor measures in the last row of Table 1. Recall that this measure is defined as the median share of users in other countries who connect to users in the country under consideration. The countries with the *highest attractor* values are the United States (7.07%), the United Kingdom (4.22%), Russia (3.30%), and Germany (2.92%). The countries with the *lowest attractor* values are Belarus (0.32%), the Czech Republic (0.35%), and Norway (0.48%).

CONCLUSIONS AND OUTLOOK

In this work, we examined the nature of user connections in OSN across countries.

The contribution of our work is two-fold: First, we identified those countries whose users are mainly connected with other users from within the same country, and those countries that are characterized by a wide spectrum of cross-country user connections. Particularly strong cross-country connection pairs share that their countries are related culturally, historically, or linguistically. Second, we identified the United States, the United Kingdom, and Russia as the major “attractor” countries, being characterized by having a large relative amount of user connections from various other countries.

Beyond contributing to understanding the complex social ties, how they are reflected in connection behavior in OSN, and their dependency on cultural, historic, and linguistic characteristics of countries, our findings are a promising source for further developments in the modeling for personalized systems. In-depth investigations in this research field will help to advance particularly in cold start scenarios, where little is known about users’ social ties. Furthermore, user connections—and tie strength in particular—play an important role for decision-support systems or in online information diffusion. Knowing the users’ countries may be a good indicator for tie strength prediction.

²Note that these figures already account for different numbers of users in different countries; the values are therefore comparable.

Table 1: Cross-country user connections in relative numbers for Last.fm, for the top 20 countries. Rows denote countries in which connections originate; columns countries which they point to. Percentages are further visualized by different intensities of gray. The last row contains the attractor values. Country names are abbreviated according to ISO 3166-1 alpha-2.

	AU	BR	BY	CA	CZ	DE	ES	FI	FR	IT	JP	MX	NL	NO	PL	RU	SE	UA	UK	US
AU	44.92%	2.66%	0.35%	2.28%	0.32%	2.60%	0.72%	0.86%	0.89%	0.90%	0.69%	0.80%	1.24%	0.66%	1.75%	2.90%	1.01%	0.71%	7.01%	16.47%
BR	0.37%	76.04%	0.20%	0.59%	0.21%	1.53%	0.57%	0.53%	0.48%	0.80%	0.41%	0.81%	0.47%	0.20%	1.73%	1.78%	0.41%	0.49%	2.00%	4.08%
BY	0.55%	2.18%	50.99%	0.59%	0.49%	2.38%	0.52%	0.57%	0.64%	0.73%	0.40%	0.44%	0.52%	0.21%	2.62%	17.12%	0.40%	4.85%	2.14%	3.71%
CA	2.08%	3.85%	0.34%	29.01%	0.36%	3.47%	0.93%	0.88%	1.40%	1.11%	0.94%	0.99%	1.28%	0.55%	2.02%	3.58%	1.14%	1.01%	6.71%	26.46%
CZ	0.44%	2.10%	0.44%	0.55%	64.21%	2.36%	0.45%	0.66%	0.71%	0.93%	0.36%	0.48%	0.67%	0.31%	2.65%	3.48%	0.45%	1.17%	2.85%	3.84%
DE	0.68%	2.86%	0.40%	0.99%	0.44%	56.79%	0.98%	1.12%	1.08%	1.23%	0.59%	0.70%	1.28%	0.51%	2.51%	3.74%	0.94%	1.05%	3.82%	6.89%
ES	0.60%	3.38%	0.28%	0.85%	0.27%	3.12%	54.45%	0.88%	1.07%	1.74%	0.67%	2.05%	1.06%	0.33%	2.22%	2.71%	0.70%	0.63%	4.29%	6.25%
FI	0.57%	2.55%	0.25%	0.64%	0.32%	2.88%	0.71%	65.27%	0.65%	0.90%	0.77%	0.56%	0.83%	0.38%	2.14%	3.16%	1.14%	0.80%	3.00%	4.92%
FR	0.96%	3.72%	0.45%	1.66%	0.56%	4.50%	1.40%	1.05%	38.60%	1.79%	1.28%	1.10%	1.32%	0.54%	3.44%	5.75%	1.12%	1.40%	5.52%	9.12%
IT	0.62%	3.90%	0.32%	0.83%	0.46%	3.21%	1.43%	0.92%	1.13%	54.96%	0.69%	0.81%	1.11%	0.46%	2.62%	3.44%	0.77%	0.88%	4.44%	6.41%
JP	0.91%	3.88%	0.34%	1.35%	0.34%	2.96%	1.06%	1.51%	1.55%	1.32%	47.12%	1.05%	1.14%	0.66%	2.93%	3.98%	0.88%	1.13%	4.14%	9.08%
MX	0.81%	5.91%	0.29%	1.11%	0.35%	2.74%	2.53%	0.85%	1.04%	1.20%	0.81%	46.68%	0.82%	0.30%	2.57%	3.11%	0.68%	0.69%	3.09%	8.82%
NL	0.88%	2.38%	0.24%	0.99%	0.34%	3.48%	0.91%	0.88%	0.87%	1.15%	0.61%	0.57%	57.20%	0.61%	2.60%	2.60%	0.89%	0.69%	4.53%	7.26%
NO	1.08%	2.29%	0.22%	0.98%	0.36%	3.15%	0.65%	0.92%	0.81%	1.10%	0.82%	0.48%	1.39%	55.68%	2.77%	2.81%	2.06%	0.86%	4.76%	7.56%
PL	0.27%	1.94%	0.26%	0.35%	0.30%	1.51%	0.42%	0.50%	0.50%	0.60%	0.35%	0.39%	0.58%	0.27%	76.87%	1.98%	0.35%	0.70%	2.37%	2.62%
RU	0.51%	2.23%	1.93%	0.69%	0.44%	2.51%	0.57%	0.83%	0.93%	0.88%	0.53%	0.53%	0.64%	0.30%	2.21%	62.80%	0.50%	4.84%	2.67%	4.88%
SE	1.00%	2.92%	0.26%	1.23%	0.32%	3.57%	0.83%	1.69%	1.02%	1.11%	0.66%	0.65%	1.23%	1.25%	2.21%	2.79%	53.96%	0.76%	4.56%	8.25%
UA	0.49%	2.44%	2.17%	0.77%	0.59%	2.81%	0.53%	0.83%	0.90%	0.90%	0.60%	0.47%	0.67%	0.37%	3.12%	19.22%	0.53%	46.09%	2.61%	4.81%
UK	1.59%	3.24%	0.31%	1.66%	0.47%	3.32%	1.17%	1.02%	1.15%	1.47%	0.72%	0.68%	1.44%	0.66%	3.42%	3.44%	1.05%	0.85%	47.52%	13.37%
US	1.76%	3.11%	0.25%	3.09%	0.30%	2.82%	0.81%	0.79%	0.90%	1.00%	0.74%	0.92%	1.09%	0.50%	1.78%	2.97%	0.89%	0.74%	6.29%	59.77%
Attractor	0.75%	2.89%	0.32%	0.99%	0.35%	2.92%	0.82%	0.88%	0.92%	1.10%	0.68%	0.69%	1.10%	0.48%	2.58%	3.30%	0.88%	0.85%	4.22%	7.07%

Including more OSN and wider scales of user connections as well as more information thereof in future analyses will open up a broad spectrum of new insights for various areas, including research tasks such as the design of personalized systems or the analysis of social user behavior across cultures and time. Research questions for future work include the following: How do cross-country user connections vary in different domains (e.g., in a professional OSN such as LinkedIn versus Care2 for activists around the globe versus YouTube for video sharing)? How do OSN vary concerning strong and weak ties? Which cross-country differences exist with respect to strong and weak ties? Are cross-country ties equally strong as ties with users within a country?

ACKNOWLEDGMENTS

This research is supported by the Austrian Science Fund (FWF): V579.

REFERENCES

- [1] Michael Bailey, Ruiqing Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. 2017. *Measuring social connectedness*. Technical Report. National Bureau of Economic Research.
- [2] George A Barnett and Grace A Benefield. 2017. Predicting international Facebook ties through cultural homophily and other factors. *New Media & Society* 19, 2 (2017), 217–239.

- [3] Nancy K Baym and Andrew Ledbetter. 2009. Tunes that bind? Predicting friendship strength in a music-based social network. *Information, Communication & Society* 12, 3 (2009), 408–427.
- [4] Diana Boer, Ronald Fischer, Micha Strack, Michael H Bond, Eva Lo, and Jason Lam. 2011. How shared preferences in music create bonds between people: Values as the missing link. *Personality and Social Psychology Bulletin* 37, 9 (2011), 1159–1171.
- [5] Richard A. Brown. 2012. Music preferences and personality among Japanese university students. *International Journal of Psychology* 47, 4 (2012), 259–268. <https://doi.org/10.1080/00207594.2011.631544>
- [6] Sejung Marina Choi, Yoojung Kim, Yongjun Sung, and Dongyoung Sohn. 2011. Bridging or bonding? A cross-cultural study of social relationships in social networking sites. *Information, Communication & Society* 14, 1 (2011), 107–129.
- [7] Sanjeev Dewan, Yi-Jen Ho, and Jui Ramaprasad. 2013. Quantifying Social Influence in an Online Music Community. In *34th International Conference on Information Systems (ICIS 2013)*. AIS, Atlanta, GA, 3259–3267.
- [8] R.I.M. Dunbar, Valerio Arnaboldi, Marco Conti, and Andrea Passarella. 2015. The structure of online social networks mirrors those in the offline world. *Social Networks* 43 (2015), 39 – 47. <https://doi.org/10.1016/j.socnet.2015.04.005>
- [9] John Johnstone and Elihu Katz. 1957. Youth and Popular Music: A Study in the Sociology of Taste. *Amer. J. Sociology* 62, 6 (1957), 563–568. <http://www.jstor.org/stable/2773131>
- [10] Thomas Krismayer, Markus Schedl, Peter Knees, and Rick Rabiser. 2017. Prediction of User Demographics from Music Listening Habits. In *15th International Workshop on Content-based Multimedia Indexing (CBMI '17)*. ACM, New York, NY, Article 8, 7 pages. <https://doi.org/10.1145/3095713.3095722>
- [11] Seth A Myers, Aneesh Sharma, Pankaj Gupta, and Jimmy Lin. 2014. Information network or social network?: the structure of the twitter follow graph. In *Proceedings of the 23rd International Conference on World Wide Web (WWW '14 Companion)*. ACM, New York, NY, 493–498. <https://doi.org/10.1145/2567948.2576939>
- [12] Nathan Nash. 2010. International Facebook “friends”: Toward McLuhan’s global village. *The McMaster Journal of Communication* 5, 1 (2010), 1–12.
- [13] Peter J Rentfrow and Samuel D Gosling. 2006. Message in a ballad: The role of music preferences in interpersonal perception. *Psychological science* 17, 3 (2006), 236–242.
- [14] Markus Schedl. 2016. The LFM-1b Dataset for Music Retrieval and Recommendation. In *Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval (ICMR '16)*. ACM, New York, NY, 103–110. <https://doi.org/10.1145/2911996.2912004>
- [15] Markus Schedl and Christine Bauer. 2017. Introducing Global and Regional Mainstreamness for Improving Personalized Music Recommendation. In *Proceedings of the 15th International Conference on Advances in Mobile Computing & Multimedia (MoMM2017)*. ACM, New York, NY, 74–81. <https://doi.org/10.1145/3151848.3151849>
- [16] Christian Steglich, Tom A. B. Snijders, and Patrick West. 2006. Applying SIENA: an illustrative analysis of the coevolution of adolescents’ friendship networks, taste in music, and alcohol consumption. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences* 2, 1 (2006), 48–56. <https://doi.org/10.1027/1614-2241.2.1.48>
- [17] Gabriel Vigliensoni and Ichiro Fujinaga. 2016. Automatic music recommendation systems: do demographic, profiling, and contextual features improve their performance?. In *17th International Society for Music Information Retrieval Conference (ISMIR) (7–11 August)*. International Society for Music Information Retrieval (ISMIR), Canada, 94–100.
- [18] Xin Wang, Steven C.H. Hoi, Martin Ester, Jiajun Bu, and Chun Chen. 2017. Learning personalized preference of strong and weak ties for social recommendation. In *Proceedings of the 26th International Conference on World Wide Web (WWW '17)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 1601–1610. <https://doi.org/10.1145/3038912.3052556>
- [19] Shanyang Zhao and David Elesh. 2008. Copresence as ‘being with’: Social contact in online public domains. *Information, Communication & Society* 11, 4 (2008), 565–583.