

ON THE IMPORTANCE OF CONSIDERING COUNTRY-SPECIFIC ASPECTS ON THE ONLINE-MARKET: AN EXAMPLE OF MUSIC RECOMMENDATION CONSIDERING COUNTRY-SPECIFIC MAINSTREAM

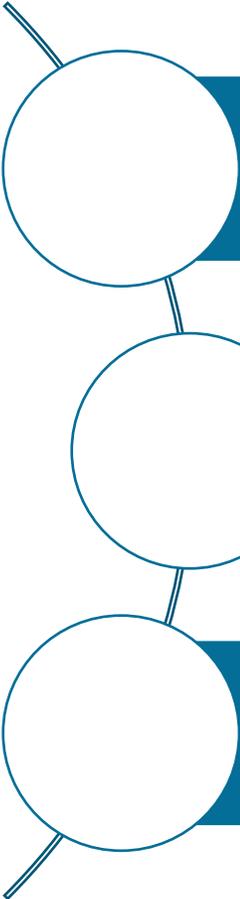
Christine Bauer & Markus Schedl
Johannes Kepler University Linz, Austria
christine.bauer@jku.at | markus.schedl@jku.at
www.cp.jku.at

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



Department of
Computational
Perception

BACKGROUND



huge amount of consumable online content → information overload

support users in searching, sorting, and filtering the massive amount of content
→ recommender systems important

example:
music recordings on YouTube, Spotify, or iTunes → music recommender systems

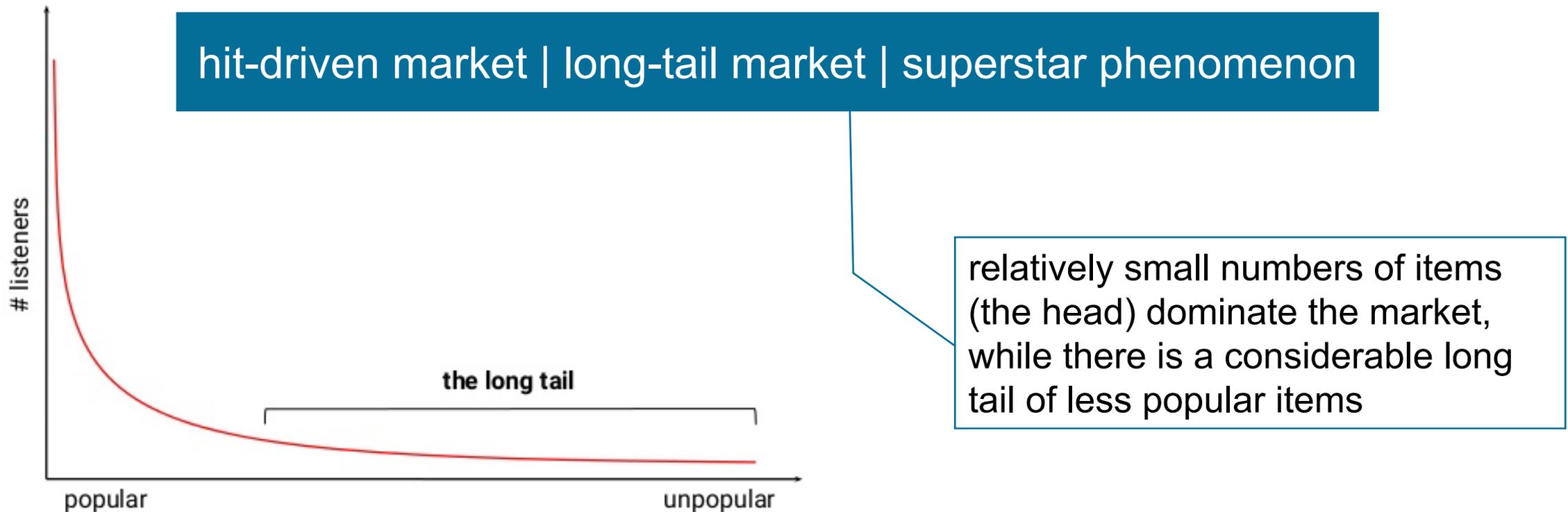
The success of a music recommender system depends
on its ability to propose
the right music,
to the right user,
at the right moment.

Laplante, A., “Improving Music Recommender Systems: What Can We Learn from Research on Music Tags?”, 15th International Society for Music Information Retrieval Conference (ISMIR), 2014, pp. 451-456.

This task is extremely complex!

POPULARITY-BASED APPROACH MAY HELP.

- popularity-based approach assumes that a random user is more likely to like a very popular music item than one of the far less popular items



LET'S TRY!

Are the most popular artists of the LFM-1b dataset good recommendations for us?

last.fm

- > 1b listening events (LE)
- > 120k users
- 47 countries
- LEs covering Jan 2005 – Aug 2014
- > 585k artists

Markus Schedl. 2016. The LFM-1b Dataset for Music Retrieval and Recommendation, Proceedings of the ACM International Conference on Multimedia Retrieval (ICMR), New York, USA, April 2016.

MOST POPULAR ARTISTS OF THE LFM-1B DATASET W.R.T. LISTENER FREQUENCY (LF)

Artist	LF
Radiohead	24,829
Nirvana	24,249
Coldplay	23,714
Daft Punk	23,661
Red Hot Chili Peppers	22,609
Muse	22,429
Queen	21,778
The Beatles	21,738
Pink Floyd	21,129
David Bowie	20,602

Global (53,258 users)

(subset with country information)

LF... listener frequency
(or playcount): the number
of unique listeners of the item

SOME PEOPLE FOLLOW POPULARITY TRENDS, SOME DO NOT...

“music mainstreaminess of a user”

describes a listener in terms of the degree to which he or she prefers music items that are currently popular or rather ignores such trends

Markus Schedl and David Hauger. 2015. Tailoring Music Recommendations to Users by Considering Diversity, Mainstreaminess, and Novelty. In Proceedings of the 38th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2015).

MUSIC MAINSTREAMINESS COMBINED WITH COLLABORATIVE FILTERING IMPROVES RESULTS...

... w.r.t. recommendation accuracy and rating prediction error than pure collaborative filtering approaches alone

Gabriel Vigliensoni and Ichiro Fujinaga. 2016. Automatic music recommendation systems: do demographic, profiling, and contextual features improve their performance?. In Proceedings of the 17th International Society for Music Information Retrieval Conference (August 7-11, 2016) (ISMIR 2016). pp 94–100.

Markus Schedl and David Hauger. 2015. Tailoring Music Recommendations to Users by Considering Diversity, Mainstreaminess, and Novelty. In Proceedings of the 38th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2015).

DOES GLOBAL MAINSTREAMINESS OCCLUDE THE COUNTRY-SPECIFIC ONE?

Top artists for selected countries,
according to listener frequency (LF)

Global (53,258 users)

Artist	LF
Radiohead	24,829
Nirvana	24,249
Coldplay	23,714
Daft Punk	23,661
Red Hot Chili Peppers	22,609
Muse	22,429
Queen	21,778
The Beatles	21,738
Pink Floyd	21,129
David Bowie	20,602

Finland (1,407 users)

Artist	LF
Metallica	703
Nightwish	695
Muse	693
Daft Punk	675
Queen	671
System of a Down	663
Coldplay	634
Nirvana	614
Pendulum	613
Iron Maiden	609

Italy (972 users)

Artist	LF
Radiohead	556
Pink Floyd	539
The Beatles	505
David Bowie	500
Muse	500
Nirvana	497
Coldplay	475
The Cure	466
Depeche Mode	459
Daft Punk	457

Turkey (479 users)

Artist	LF
Pink Floyd	292
Radiohead	289
Metallica	268
Coldplay	261
Nirvana	251
Massive Attack	249
The Beatles	240
Red Hot Chili Peppers	240
Queen	238
Led Zeppelin	236

YES, GLOBAL MAINSTREAMINESS OCCLUDES THE COUNTRY-SPECIFIC ONE

Global (53,258 users)

Artist	LF
Radiohead	24,829
Nirvana	24,249
Coldplay	23,714
Daft Punk	23,661
Red Hot Chili Peppers	22,609
Muse	22,429
Queen	21,778
The Beatles	21,738
Pink Floyd	21,129
David Bowie	20,602

Top artists for selected countries,
according to artist frequency–inverse listener frequency (AF-ILF)

Finland (1,407 users)

Artist	AF-ILF
St. Hood	70.526
The Sun Sawed in 1/2	67.490
tiko-μ	66.546
Worth the Pain	66.058
Cutdown	65.247
Katariina Hänninen	64.955
Game Music Finland	64.835
Daisuke Ishiwatari	63.565
Altis	63.235
Redrum-187	62.428

Italy (972 users)

Artist	AF-ILF
CaneSecco	68.451
DSA Commando	66.049
Veronica Marchi	65.864
Train To Roots	65.459
Alessandro Raina	64.228
Machete Empire	63.915
Danti	62.958
Dargen D'Amico	62.453
宝塚歌劇団・宙組	62.228
Aquefrigide	61.663

Turkey (479 users)

Artist	AF-ILF
Cüneyt Ergün	64.473
Floyd Red Crow Westerman	61.955
Fırat Tanış	58.666
Acil Servis	58.439
Taste (Rory Gallagher)	58.366
Mezarkabul	57.799
Rachmaninoff Sergey	57.733
Mabel Matiz	57.619
Grup Yorum	56.855
Yüzyüzeyken Konuşuruz	56.748

A RANK-BASED APPROACH ... CONSIDERING DIFFERENCES BETWEEN COUNTRIES

- Rank-based ($R_{u,g/c}$): rank-order correlation according to Kendall's τ between global/country's and user's preference profiles



$$R_{u,g} = \tau (\text{ranks} (PP_u) , \text{ranks} (PP_g))$$



$$R_{u,c} = \tau (\text{ranks} (PP_u) , \text{ranks} (PP_c))$$

where $\text{ranks}(PP_u)$ denotes a function that converts the real-valued preference profile of user u to ranks, $\text{ranks}(PP_c)$ accordingly on the country-level (country of user u), and $\text{ranks}(PP_g)$ on the global level, i.e. considering all users

higher values indicate closer to the mainstream, whereas lower ones indicate farther away from the mainstream

METHODS: EVALUATION APPROACH FOR MUSIC RECOMMENDATION TAILORED TO COUNTRY-SPECIFIC USER MAINSTREAMINESS

dataset

- subset of LFM-1b: 53,258 users from 47 countries
- users with country information; only countries with min. 100 users

evaluation method

- rating prediction on playcounts scaled to [0, 1000]

algorithm

- model-based collaborative filtering (SVD)

analysis

- different definitions and levels of mainstreaminess

definitions

- rank-based approach; global vs. country-specific

levels

- user tertiles w.r.t. mainstreaminess (lower, mid, upper 1/3)

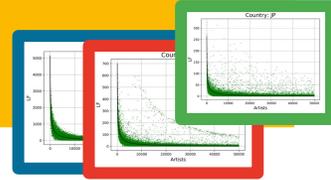
performance measures

- root mean square error (RMSE)

TAKE AWAY PREVIEW...

country-specific differences of users' listening behavior concerning music mainstreaminess

- national boundaries on the global online market do exist



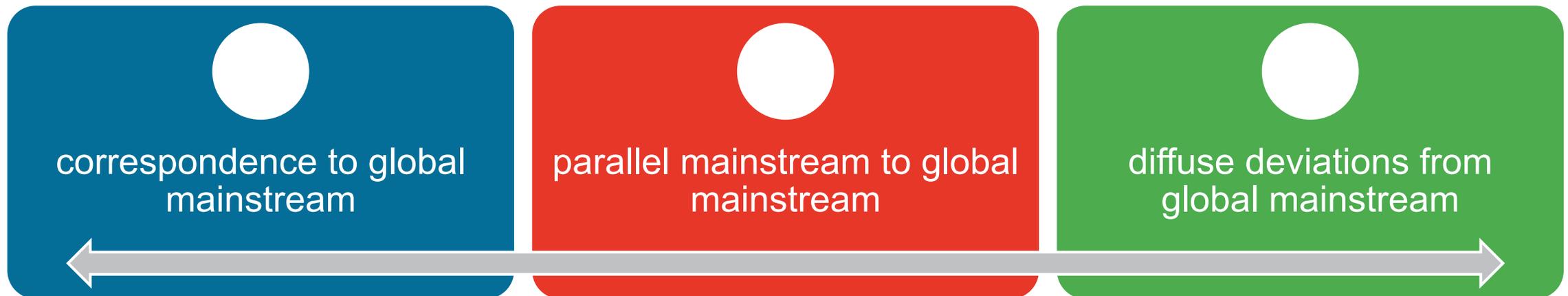
tailoring music recommendations to a user's country may improve recommendation accuracy

- overall improvement when using the country scope (considering *all* users of a country)
- improvement for the *low* mainstreaminess user set is remarkable
- combination of user mainstreaminess and “country filtering” works:
 - particularly well for countries far away from the global mainstream
 - outperforms the global mainstreaminess measure for *mid* and *low* for countries close to the global mainstream

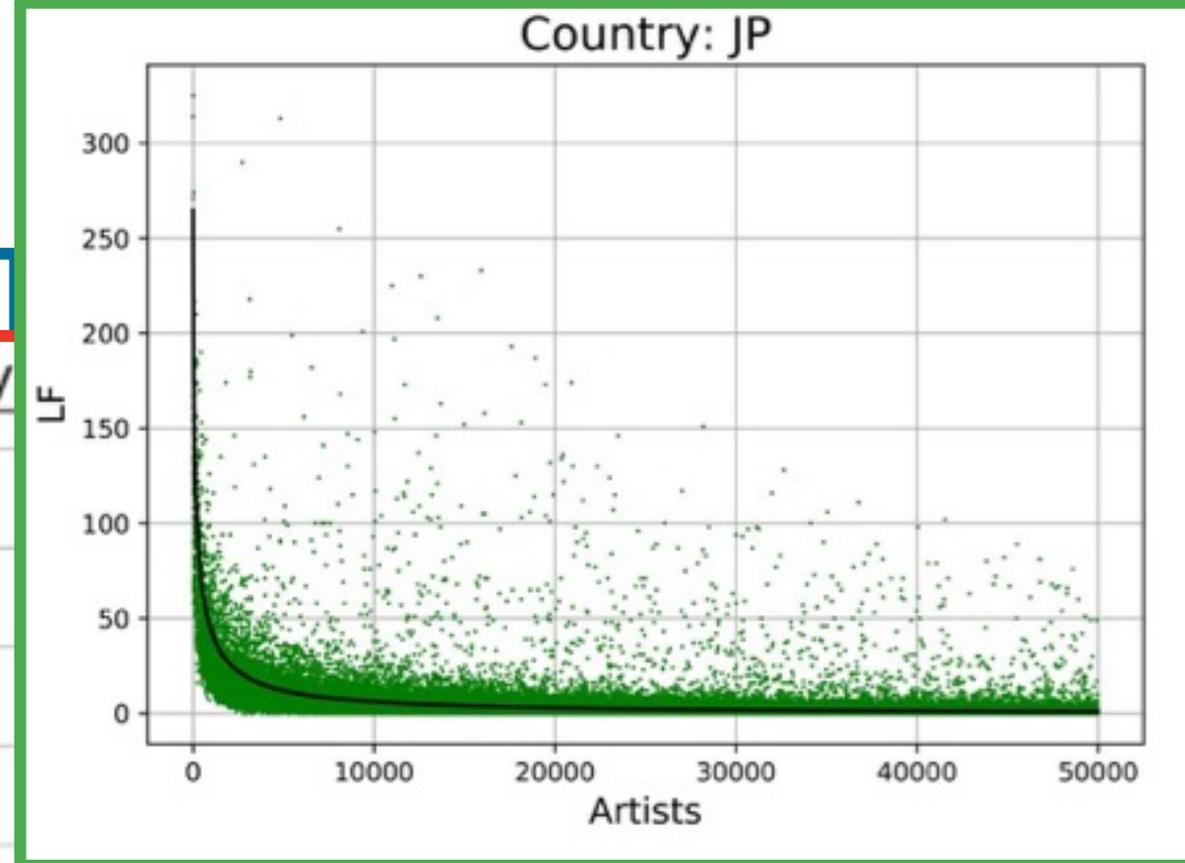
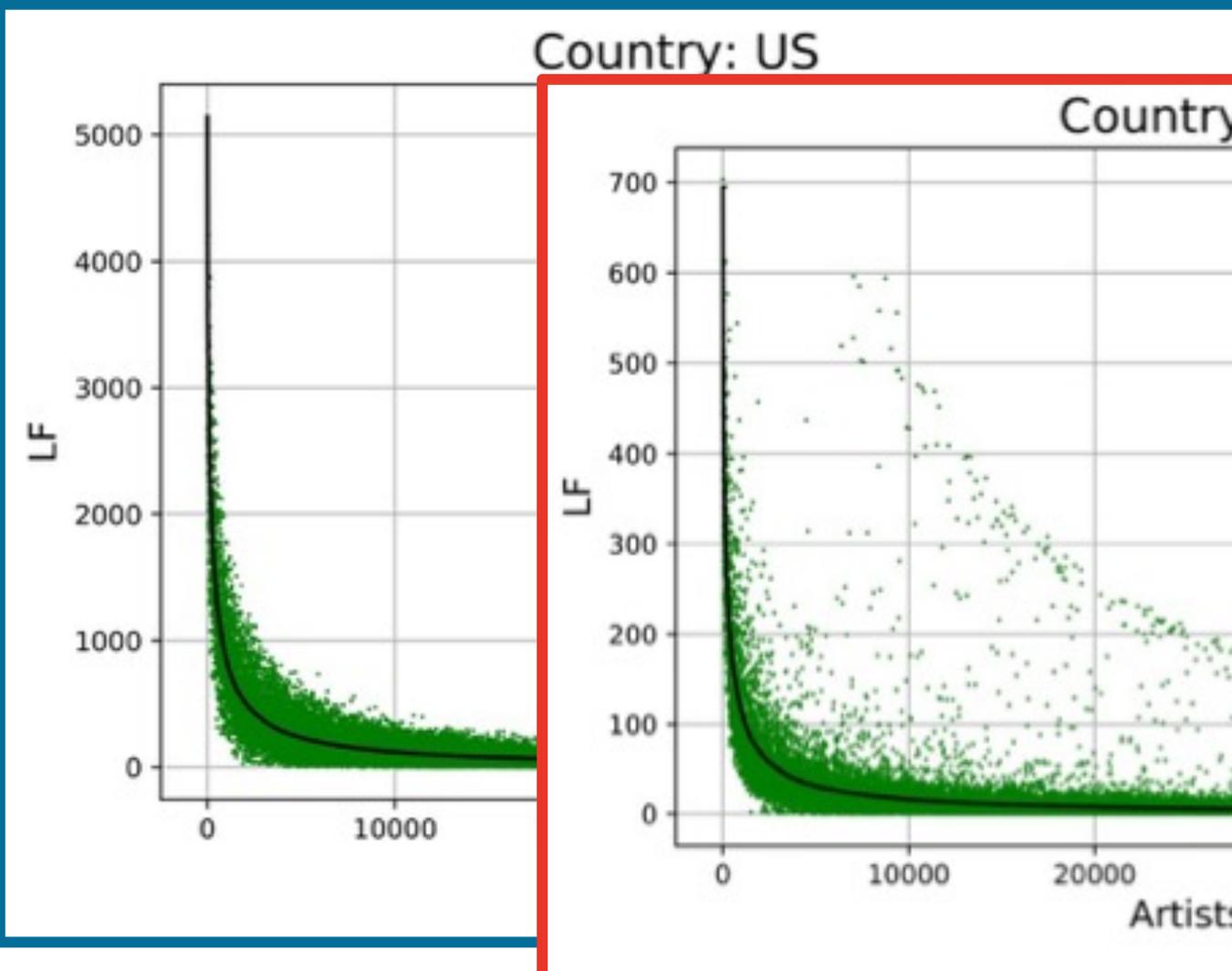
adopted measures do not perform equally well for all kinds of country mainstreaminess profiles → important to take into account which country is addressed

RESULT #1

There are country-specific differences of users' listening behavior concerning music mainstreamness.



RESULT #1, DETAILED



deviations from
mainstream

RESULT #2A

There is an
*overall improvement when using the country scope,
considering all user sets*

(i.e., no differentiation between high, mid, or low mainstreaminess)

measure	user set	w.RMSE
$R_{u,g}$	all	15.906
	high	3.680
	mid	7.443
	low	19.183
$R_{u,c}$	all	14.349
	high	3.687
	mid	4.270
	low	3.692



RESULT #2B

There is a
very slight fall back for the high mainstreaminess user set
 using the country scope.

measure	user set	w.RMSE
$R_{u,g}$	all	15.906
	high	3.680
	mid	7.443
	low	19.183
$R_{u,c}$	all	14.349
	high	3.687
	mid	4.270
	low	3.692

RESULT #2C

There is a *considerable improvement for the mid and low mainstreaminess user set using the country scope.*

measure	user set	w.RMSE
$R_{u,g}$	all	15.906
	high	3.680
	mid	7.443
	low	19.183
$R_{u,c}$	all	14.349
	high	3.687
	mid	4.270
	low	3.692



RESULT #2D

The *improvement for the low mainstreaminess user set is remarkable using the country scope.*

measure	user set	w.RMSE
$R_{u,g}$	all	15.906
	high	3.680
	mid	7.443
	low	19.183
$R_{u,c}$	all	14.349
	high	3.687
	mid	4.270
	low	3.692



RESULT #3A

Considering entire country user set (all), *global mainstreamness measure performs poorly for Finland (far from global mainstream) and very well for the United States and Sweden (oriented at the global mainstream).*

measure	country	user set	RMSE		measure	country	user set	RMSE
$R_{u,g}$	US (global mainstream)	all	5.327		$R_{u,c}$	US (global mainstream)	all	28.995
		high	5.396				high	5.360
		mid	24.845				mid	5.411
		low	28.544				low	5.434
	FI (parallel mainstream)	all	27.084			FI (parallel mainstream)	all	3.976
		high	3.909				high	4.058
		mid	4.135				mid	25.723
		low	4.077				low	4.085
	SE (diffuse deviations)	all	6.209			SE (diffuse deviations)	all	6.199
		high	6.278				high	6.225
		mid	6.318				mid	6.473
		low	6.436				low	6.331

correspondence to global mainstream

parallel mainstream to global mainstream

diffuse deviations from global mainstream

RESULT #3C

The *global mainstreaminess measure* performs particularly well for the high mainstream user set and for no differentiation (all); especially for the United States (*oriented at the global mainstream*).

measure	country	user set	RMSE		measure	country	user set	RMSE
$R_{u,g}$	US (global mainstream)	all	5.327		$R_{u,c}$	US (global mainstream)	all	28.995
		high	5.396				high	5.360
		mid	24.845				mid	5.411
		low	28.544	low			5.434	
	FI (parallel mainstream)	all	27.084			FI (parallel mainstream)	all	3.976
		high	3.909				high	4.058
		mid	4.135				mid	25.723
		low	4.077	low			4.085	
	SE (diffuse deviations)	all	6.209			SE (diffuse deviations)	all	6.199
		high	6.278				high	6.225
		mid	6.318				mid	6.473
		low	6.436	low			6.331	

correspondence to global mainstream

parallel mainstream to global mainstream

diffuse deviations from global mainstream

RESULT #4A

For the United States (*oriented at the global mainstream*), the *country-specific mainstreamness measure* performs poorly the *all* user set (no differentiation); and well for the specific user sets.

measure	country	user set	RMSE	measure	country	user set	RMSE
$R_{u,g}$	US (global mainstream)	all	5.327	$R_{u,c}$	US (global mainstream)	all	28.995
		high	5.396			high	5.360
		mid	24.845			mid	5.411
		low	28.544			low	5.434
	FI (parallel mainstream)	all	27.084		FI (parallel mainstream)	all	3.976
		high	3.909			high	4.058
		mid	4.135			mid	25.723
		low	4.077			low	4.085
	SE (diffuse deviations)	all	6.209		SE (diffuse deviations)	all	6.199
		high	6.278			high	6.225
		mid	6.318			mid	6.473
		low	6.436			low	6.331



correspondence to global mainstream

parallel mainstream to global mainstream

diffuse deviations from global mainstream

RESULT #4B

For Finland (*parallel mainstream*), the *country-specific mainstreamness measure* performs well for all sets except the *mid* user set.

measure	country	user set	RMSE
$R_{u,g}$	US (global mainstream)	all	5.327
		high	5.396
		mid	24.845
		low	28.544
	FI (parallel mainstream)	all	27.084
		high	3.909
		mid	4.135
		low	4.077
	SE (diffuse deviations)	all	6.209
		high	6.278
		mid	6.318
		low	6.436

measure	country	user set	RMSE
$R_{u,c}$	US (global mainstream)	all	28.995
		high	5.360
		mid	5.411
		low	5.434
	FI (parallel mainstream)	all	3.976
		high	4.058
		mid	25.723
		low	4.085
	SE (diffuse deviations)	all	6.199
		high	6.225
		mid	6.473
		low	6.331



correspondence to global mainstream

parallel mainstream to global mainstream

diffuse deviations from global mainstream

RESULT #4B

For Finland (*parallel mainstream*), the *country-specific mainstreamness measure* even outperforms the *global mainstreamness measure* in the *all* user set.

measure	country	user set	RMSE
	US (global mainstream)	all	5.327
		high	5.396
		mid	24.845
		low	28.544
$R_{u,g}$	FI (parallel mainstream)	all	27.084
		high	3.909
		mid	4.135
		low	4.077
	SE (diffuse deviations)	all	6.209
		high	6.278
		mid	6.318
		low	6.436

measure	country	user set	RMSE
	US (global mainstream)	all	28.995
		high	5.360
		mid	5.411
		low	5.434
$R_{u,c}$	FI (parallel mainstream)	all	3.976
		high	4.058
		mid	25.723
		low	4.085
	SE (diffuse deviations)	all	6.199
		high	6.225
		mid	6.473
		low	6.331



correspondence to global mainstream

parallel mainstream to global mainstream

diffuse deviations from global mainstream

RESULT #4C

For Sweden (*diffuse deviations*), comparable results for all segments.

measure	country	user set	RMSE
$R_{u,g}$	US (global mainstream)	all	5.327
		high	5.396
		mid	24.845
		low	28.544
	FI (parallel mainstream)	all	27.084
		high	3.909
		mid	4.135
		low	4.077
	SE (diffuse deviations)	all	6.209
		high	6.278
		mid	6.318
		low	6.436

measure	country	user set	RMSE
$R_{u,c}$	US (global mainstream)	all	28.995
		high	5.360
		mid	5.411
		low	5.434
	FI (parallel mainstream)	all	3.976
		high	4.058
		mid	25.723
		low	4.085
	SE (diffuse deviations)	all	6.199
		high	6.225
		mid	6.473
		low	6.331

correspondence to global mainstream

parallel mainstream to global mainstream

diffuse deviations from global mainstream

RESULT #5A

The combination of considering a user's mainstreamness and "country filtering" for music recommendation works particularly well for countries far away from the global mainstream (e.g., Finland).

measure	country	user set	RMSE	measure	country	user set	RMSE
$R_{u,g}$	US (global mainstream)	all	5.327	$R_{u,c}$	US (global mainstream)	all	28.995
		high	5.396			high	5.360
		mid	24.845			mid	5.411
		low	28.544			low	5.434
	FI (parallel mainstream)	all	27.084		FI (parallel mainstream)	all	3.976
		high	3.909			high	4.058
		mid	4.135			mid	25.723
		low	4.077			low	4.085
	SE (diffuse deviations)	all	6.209		SE (diffuse deviations)	all	6.199
		high	6.278			high	6.225
		mid	6.318			mid	6.473
		low	6.436			low	6.331



RESULT #5BI

The combination of considering a user's mainstreamness and "country filtering" for music recommendation performs poorly for countries close to the global mainstream (e.g., United States).

measure	country	user set	RMSE	measure	country	user set	RMSE
$R_{u,g}$	US (global mainstream)	all	5.327	$R_{u,c}$	US (global mainstream)	all	28.995
		high	5.396			high	5.360
		mid	24.845			mid	5.411
		low	28.544			low	5.434
	FI (parallel mainstream)	all	27.084		FI (parallel mainstream)	all	3.976
		high	3.909			high	4.058
		mid	4.135			mid	25.723
		low	4.077			low	4.085
	SE (diffuse deviations)	all	6.209		SE (diffuse deviations)	all	6.199
		high	6.278			high	6.225
		mid	6.318			mid	6.473
		low	6.436			low	6.331



correspondence to global mainstream

parallel mainstream to global mainstream

diffuse deviations from global mainstream

RESULT #5BII

The combination of considering a user's mainstreamness and "country filtering" for music recommendation outperforms the global mainstreamness measure for mid and low for countries close to the global mainstream (e.g., United States).

measure	country	user set	RMSE
	US (global mainstream)	all	5.327
		high	5.396
		mid	24.845
		low	28.544
$R_{u,g}$	FI (parallel mainstream)	all	27.084
		high	3.909
		mid	4.135
		low	4.077
	SE (diffuse deviations)	all	6.209
		high	6.278
		mid	6.318
		low	6.436

measure	country	user set	RMSE
	US (global mainstream)	all	28.995
		high	5.360
		mid	5.411
		low	5.434
$R_{u,c}$	FI (parallel mainstream)	all	3.976
		high	4.058
		mid	25.723
		low	4.085
	SE (diffuse deviations)	all	6.199
		high	6.225
		mid	6.473
		low	6.331



correspondence to global mainstream

parallel mainstream to global mainstream

diffuse deviations from global mainstream

FUTURE AVENUES OF RESEARCH

avenues for future research

- advancements in mainstreamness measurement
 - to further improve recommendation performance
- focus on algorithmic advancements that may be described as “recommender of recommenders”:
 - depending on the identified user country and the respective country profile, different measurements and/or algorithms would be adopted for further steps in the recommendation process

next steps

- delve into detail for a larger scale of countries (current limitation: just 3 rather dissimilar countries)
- analyze in which countries what kind of mainstreamness functions perform particularly well or poorly

further perspectives

- expand the perspective on cultural aspects
 - from user country to cultural regions, languages, urban vs. country-side,...

IMPLICATIONS

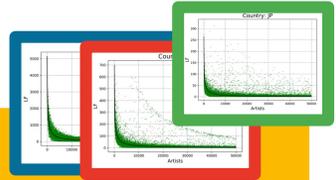
practical implications

- presented approach can be readily adopted in real-world MRS
- including music streaming services (e.g., Spotify, Pandora), but also multimedia platforms hosting music videos (e.g., YouTube)

theoretical implications

- national boundaries on the global online market do exist
- particularly interesting as the music recording industry is considered a “globally oriented market” compared to rather “locally oriented markets” (e.g., food products)

TAKE AWAY RECAPTURING



country-specific differences of users' listening behavior concerning music mainstreamness

- national boundaries on the global online market do exist



combination of considering a user's mainstreamness and "country filtering"

- rank-based approach considering differences between countries

$$R_{u,g} = \tau(\text{ranks}(PP_u), \text{ranks}(PP_g))$$

$$R_{u,c} = \tau(\text{ranks}(PP_u), \text{ranks}(PP_c))$$

tailoring music recommendations to a user's country may improve recommendation accuracy

- overall improvement when using the country scope (considering *all* users of a country)
- improvement for the *low* mainstreamness user set is remarkable
- combination of user mainstreamness and "country filtering" works
 - particularly well for countries far away from the global mainstream
 - outperforms the global mainstreamness measure for *mid* and *low* for countries close to the global mainstream

adopted measures do not perform equally well for all kinds of country mainstreamness profiles → important to take into account which country is addressed

JKU



**JOHANNES KEPLER
UNIVERSITY LINZ**

Christine Bauer & Markus Schedl
Johannes Kepler University Linz, Austria
christine.bauer@jku.at | markus.schedl@jku.at
www.cp.jku.at

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