

Content-based Recommendations for Radio Stations with Deep Learned Audio Fingerprints

Stefan Langer,¹ Liza Obermeier,² André Ebert,³ Markus Friedrich,⁴ Emma Munisamy,⁵ Claudia Linnhoff-Popien⁶

Abstract: The world of linear radio broadcasting is characterized by a wide variety of stations and played content. That is why finding stations playing the preferred content is a tough task for a potential listener, especially due to the overwhelming number of offered choices. Here, recommender systems usually step in but existing content-based approaches rely on metadata and thus are constrained by the available data quality. Other approaches leverage user behavior data and thus do not exploit any domain-specific knowledge and are furthermore disadvantageous regarding privacy concerns. Therefore, we propose a new pipeline for the generation of audio-based radio station fingerprints relying on audio stream crawling and a *Deep Autoencoder*. We show that the proposed fingerprints are especially useful for characterizing radio stations by their audio content and thus are an excellent representation for meaningful and reliable radio station recommendations. Furthermore, the proposed modules are part of the *HRADIO Communication Platform*, which enables hybrid radio features to radio stations. It is released with a flexible open source license and enables especially small- and medium-sized businesses, to provide customized and high quality radio services to potential listeners.

Keywords: Hybrid Radio; Multimedia Services; Recommender Systems; Unsupervised Learning; Deep Audio Fingerprints; Deep Learning

1 Introduction

Despite emerging competition from on-demand content services, linear radio broadcasting still remains one of the most popular entertainment and information media in Europe. Its advantage lies in its technical simplicity, its topicality, and its personal approach conveyed by professional moderators. However, with services like Spotify, Deezer, or Google Play, strong competitors have recently appeared. Those on-demand media services have the

¹ Ludwig-Maximilians-Universität München, Institut für Informatik, Oettingenstrasse 67, 80538 Munich, Germany, stefan.langer@ifi.lmu.de

² inovex GmbH, Data Management and Analytics, Lindberghstrasse 3, 80939 Munich, Germany, lobermeier@inovex.de

³ inovex GmbH, Data Management and Analytics, Lindberghstrasse 3, 80939 Munich, Germany, aebert@inovex.de

⁴ Ludwig-Maximilians-Universität München, Institut für Informatik, Oettingenstrasse 67, 80538 Munich, Germany, markus.friedrich@ifi.lmu.de

⁵ Ludwig-Maximilians-Universität München, Institut für Informatik, Oettingenstrasse 67, 80538 Munich, Germany, emma.munisamy@ifi.lmu.de

⁶ Ludwig-Maximilians-Universität München, Institut für Informatik, Oettingenstrasse 67, 80538 Munich, Germany, linnhoff@ifi.lmu.de

advantage of providing a personalized listening experience and a wide range of contents combined with precise recommendation systems.

The challenge for radio broadcasters is now to enrich their classic, linear radio programme with online-based, personalized technologies in order to improve the listening experience and to bridge the gap between linear and on-demand content providers. So-called hybrid radio technologies comprise of techniques for privacy-preserving user data collection that are capable of providing individual and secure feedback channels. These channels enable customers to actively take part in the composition of their radio programme. Another important enhancement compared to the established linear programme design is the on-the-fly substitution of content, e.g., replacing ads with songs from a pre-selected music playlist or the skipping of disliked content from one radio station with preferred content from another radio station. In combination with other hybrid radio technologies, such features help mainstream stations as well as small radio businesses to catch up with online streaming services and to even surpass them in specific disciplines, e.g., interactivity and feedback channel communication. To realize such rich features, a precise and comprehensive recommender engine is necessary. Therefore, on the one hand a large amount of well-structured and rich data is needed, on the other hand different concepts for the analysis of meta and audio data have to be selected, implemented and carefully evaluated.

This paper focuses on the latter and thus on the question of how meaningful recommendations for radio stations can be generated on the basis of rich metadata provided by the *HRADIO Communication Platform*, which was implemented by the authors within the scope of previous work [Fr19]. The findings and implemented concepts presented in this work are part of the open accessible HRADIO project⁷. In the following, two main approach categories are distinguished for analyzing and recommending radio stations and programmes: Recommendations based on *Collaborative Filtering* (CF) and *Content-based Filtering* (CBF) approaches. CF focuses on user opinions and behavior, which can lead to privacy issues as well as the *Cold Start Problem* (see Section 2). Thus, a CBF-based recommender system working with characteristics derived from available station data is preferred in context of this work. In [Fr19], it could be shown that station recommendation based on available metadata is possible, but only if the metadata reaches a certain quality level. This is provably often not the case. Other meaningful characteristics can be directly derived from the station's audio signal which significantly reduces metadata quality requirements. Therefore, we propose a *Deep Learning*-based audio crawling and fingerprint extraction pipeline for the characterization of radio stations and show visual results for numerous stations. Furthermore, we detail on how a recommender system based on the developed station fingerprints can be implemented.

Due to these unique technical circumstances as well as the fact that the *HRADIO Communication Platform* is released under a permissive open source license and available at

⁷ <https://www.hradio.eu/> - The HRADIO project and thus this work was funded by H2020, the EU Framework Programme for Research and Innovation. is driven by several radio stations, small development companies, technology experts, research institutes and the Ludwig-Maximilians-University Munich

no cost, new opportunities are offered especially for small- and medium-sized businesses within the radio landscape. It significantly increases the visibility of radio stations and makes them searchable via a search service that can be used by websites and mobile applications regardless of the size of the company. For this purpose, the daily programme and the audio stream of a station are analyzed. This is done automatically, without additional effort and without costs. This creates enormous added value for both radio stations and potential listeners. In particular, small or regional stations that were previously unknown to non-regional listeners can now be found on the basis of a listener's programme taste. In addition, previously untapped potentials are opened up with regard to findability, routing and provision of niche topics as well as the allocation of tailor-made advertising relevant only to certain listener groups.

To present the underlying concept, its implementation, and its evaluation of these ideas, this paper is structured as follows: Section 2 provides a brief overview of recommendation concepts and related work. Section 3 explains the basic principles of the proposed fingerprinting pipeline and its implementation. Its recommendation capabilities are evaluated in Section 4 while Section 5 summarizes this work.

2 Related Work

Next to recommendation approaches such as *You May Like* (YML), *Knowledge-based Filtering* (KF), and *Demographic Filtering* (DF), there are two mainly recognized concepts: *Content-based Filtering* (CBF) and *Collaborative Filtering* (CF) [Be07, Bu02, RRS11]. CF-, DF-, and KF-based systems show good performance if enough user data is available. But an open issue is the so-called *Cold Start Problem*, which occurs in the initial phase of the system where not enough user data is available to create meaningful recommendations [Ch15, La08, CAS18]. Another issue for CF and DF systems are so-called *Filter Bubbles*, describing the creation of closed-off, synthetic environments, in which always the same items are recommended, disregarding the existence of contrary or different items beyond the bubble [Pa11, GLH19]. In contrast to that, an initial issue of CBF-based systems is the need for high-quality metadata precisely describing the items to recommend [Fr19]. The concept presented in this paper is part of the *HRADIO Communication Platform*, which provides a vast amount of metadata and audio information within a hybrid radio context [Fr19] and is completely open source. For this reason, and to avoid issues like *Filter Bubbles* or the *Cold Start Problem*, it takes a CBF-based approach.

In order to compare radio services on basis of their audio features, existing approaches which utilize *Deep Learning* for music genre recognition (MGR) can be utilized [DBS11, Or18]. Gwardys et al. propose a concept using transfer learning in combination with a convolutional neural network for MGR [GG14]. Logan et al. and Siddiquee et al. present methods for measuring the similarity of music on basis of audio signals [LS01, Si16]. After clustering raw audio features, Logan et al. compare entities using the *Earth Movers Distance* (EMD) [RTG00]. Çataltepe et al. use adaptive features and user grouping to take note to the aspect

that different traits within music are of different importance for each user by including historical information about the users' listening behaviour [ÇA07]. Together with others, these works provide valuable input for the proposed concept.

3 Concept

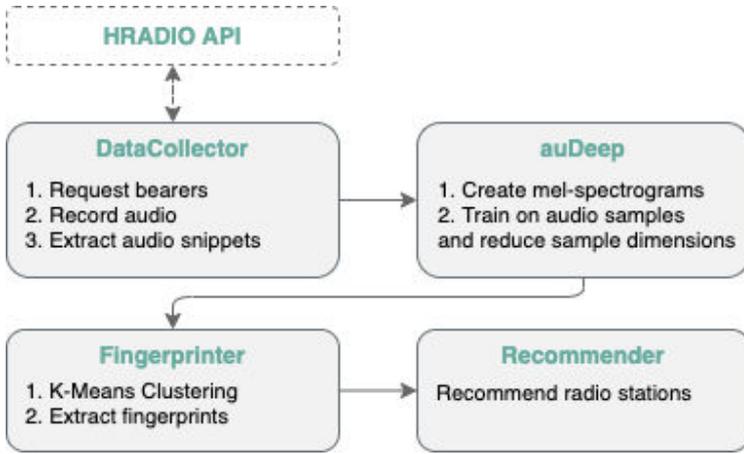


Fig. 1: The proposed pipeline consists of four modules: the *Data Collector*, the *auDeep* autoencoder, the *Fingerprinter*, and the *Recommender*. The *Data Collector* ensures the binding to the *HRADIO API*.

This chapter details the steps of the proposed analysis pipeline consisting of a *Data Collector*, the *auDeep* autoencoder, the *Fingerprinter*, and a *Recommender*, as depicted in Figure 1. In this context, the *Data Collector* unit (see Chapter 3.1) ensures the binding of the pipeline to the *HRadio API* (which is used for requesting a list of radio services from the *HRADIO Communication Platform* [Fr19]), records their audio streams and extracts audio bytes in predefined intervals. The raw audio snippets are transformed into mel-scaled spectrograms. These spectrograms express human perceptible image representations of audio signals and serve as training input for the *Deep Neural Autoencoder* provided by the *auDeep* toolkit. It is trained to reduce the dimensions of the audio representations [Fr17]. By applying the trained encoder component in our concept, the input data is compressed and the samples are reduced to vectors with 1024 dimensions. These vectors are the input of the *Fingerprinter*, which trains a *K-Means* clustering model [Pe11]. The distribution of samples in each cluster per radio station is regarded as the station's fingerprint. The last component shown in Figure 1 is the *Recommender*, which recommends radio stations similar to a particular input station on basis of the Euclidean distance of their fingerprints. A small distance implies a comparably high similarity (see Section 3.3 for more details).

3.1 Data Collector

The *DataCollector* unit (see Figure 1) consists of several components that collect and process radio data. One of them records radio stations by requesting a list of radio services and their HTTP *bearers*⁸ from the *HRADIO Communication Platform* [Fr19] via its REST API. Currently, a list of 461 valid and unique streams is received. The HTTP stream address of a sender is queried in order to receive the audio bytes of the corresponding audio stream using ICY ('I Can Yell') tag technology. The ICY tag protocol defines the transmission of textual content within audio streams and is commonly used due to its simple integration. Using *RabbitMQ*⁹ the received audio bytes are transferred to a separate component which is responsible for extracting audio snippets out of the radio streams. During 24 hours, each station is recorded for five seconds within intervals of two minutes. In order to reduce the amount of news sequences included in the recorded snippets, the five minutes before and after full hours are not recorded. This leads to a total number of 576 samples per radio station. Thereby, a full day of samples for 431 radio stations could be recorded, while 30 stations could not be recorded entirely due to server-side connection problems. In total, we collected 266,239 audio snippets, whereas 17,983 belong to incomplete radio station recordings.

3.2 Deep Fingerprints

The audio signals extracted in the previous step are transformed into a time-frequency representation, called mel spectrogram (see Figure 2), using the *auDeep* toolkit. Mel spectrograms are close to the human perception of audio signals with the mel scale being an assignment of actual to perceived frequencies. In the example in Figure 2, time is depicted on the x-axis and f frequencies are shown on the y-axis. In addition, the colour represents frequency amplitudes (or loudness): dark areas indicate low amplitudes, bright areas indicate high (or loud) amplitudes. For mel spectrogram generation we use the hyperparameters suggested by *auDeep* (a mel scale of magnitude 128, a window width of 0.08 with an overlap of 0.04, a fixed length of 5 seconds per input snippet, and clipping of values below $-60db$) [Fr17].

Subsequently, the *auDeep* autoencoder was trained on all 266,239 audio files, whereas the fingerprinting is only applied to complete sets of audio snippets (for 431 radio stations). The network is trained across 64 epochs with a batch size of 64 on 2 layers with each having 256 gated recurrent units (GRU), a learning rate of 0.001 and a dropout of 0.2. Training the network for 7 days resulted in a loss of 0.237. After the training is finished the *Fingerprinter* creates fingerprints of all complete stations, using all 266,239 vectors generated by the *auDeep* component. All these data points are divided into n clusters by applying the *K-Means* clustering algorithm. The parameter n is determined by using the

⁸ Connection information, how a radio station can be received by a device (broadcast or streaming)

⁹ An open source message broker (<https://www.rabbitmq.com/>)

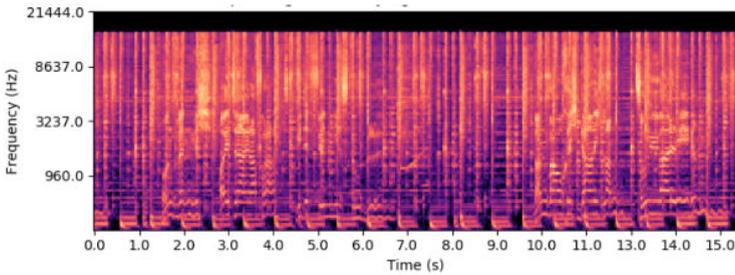


Fig. 2: An example of a mel spectrogram with 128 mel features. The x-axis represents the time scale and the y-axis the frequency scale.

silhouette coefficient [AT07], resulting in a range from 9 to 16 and showing a peak value at 11 as can be seen in Figure 3. Each data point is assigned to exactly one cluster and the fingerprint is derived from a histogram across all clusters for each station. This fingerprint vector now serves as input for the *Recommender*.

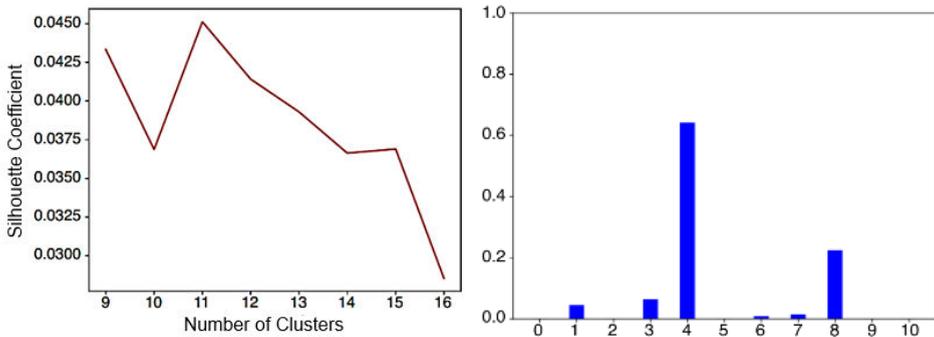


Fig. 3: Silhouette coefficients for different numbers of clusters (left) with an optimum cluster count of 11. On the right side a sample histogram of *BR KLASSIK* is depicted after the feature extraction process. The cluster index is depicted on the x-axis whereas the per-cluster frequency of occurrence is on the y-axis.

3.3 Recommender System

Based on learned per-station fingerprints (see Section 3.2) it is possible to define a similarity metric for radio stations which is essentially the Euclidean distance between fingerprint vectors, considering the space of all 11-dimensional fingerprint vectors as an Euclidean

vector space. This similarity metric can be used to establish a content-based comparison of radio stations (the fingerprint represents the content of a station): A small distance between two station fingerprints implies high similarity. Using this mechanism for recommendations, there are essentially two possibilities:

1. The k nearest radio stations are suggested to the user as similar.
2. Only radio stations within a certain Euclidean distance are listed.

The first possibility carries the risk that the returned station list may also contain distant radio stations. In contrast to that, for the second advance, the distances of provided radio stations are small in any case, while it carries the risk that the result set may be empty.

Because of the subjectivity of user taste and ratings, providing high-quality recommendations is not a trivial task and cannot only be based on station fingerprints. The *HRADIO Communication Platform* offers therefore a multitude of different recommendation strategies with the fingerprint-based recommender being one of it. The different modules are highly configurable and can be combined to a customized radio station recommendation experience. Examples for additional modules are trend-, metadata- and location-based recommenders.

4 Evaluation

In this section, the recommender system is evaluated quantitatively and its results are discussed by example (Section 4.1). Furthermore, we were interested in identifying station archetypes that represent certain categories (or genres) of stations. For that purpose, an *Archetypal Analysis* of the station fingerprint dataset was conducted with its results being presented in Section 4.2. Finally, fingerprint deltas between different times of day are analyzed and discussed while highlighting some significant examples in Section 4.2.1.

4.1 Recommender System

The evaluation of the recommender system is based on the station fingerprint dataset and corresponding genre labels that stem from existing metadata as visualized in Figure 5. In the following, we show how the distance measures map to these genre labels by choosing the most meaningful examples.

Table 1 shows the three closest radio stations (according to their fingerprint distance) with their distance to a requested service. The station *BR Klassik* is assigned to the genres *Classical Music* and *Special Music*. The closest 3 stations are *NDR Kultur* with the genres *Classical Music* and *Cultural* with an Euclidean distance of 59.58, *HR 2* with the genres *Classical Musik*, *Cultural*, and *Special Music* and an Euclidean distance of 60.81, and *Classic FM* with the genres *Classical Music* and *News* and an Euclidean distance of 60.93.

Requested station	1st closest station	2nd closest station	3rd closest station
BR Klassik <i>Classical Music</i> <i>Special Music</i>	NDR Kultur <i>Classical Music</i> <i>Cultural</i> distance: 59.58	HR2 <i>Classical Music</i> <i>Cultural</i> distance: 60.81	Classic FM <i>Classical Music</i> <i>News</i> distance: 60.93
Heart UK <i>Classic/Dance/Pop-rock</i> <i>Disco</i>	Capital XTRA Reloaded <i>Rap/HipHop/Raggae</i> distance: 41.09	3FM Isle of Man <i>Hit-Chart</i> distance: 86.34	FFH Rock <i>Rock</i> <i>Soft Rock</i> distance: 88.29
LBC UK <i>Local/Regional</i> <i>News</i>	Bayern 5 Plus <i>Information</i> distance: 117.97	WDR 3 <i>Classical Music</i> <i>Cultural</i> distance: 118.22	SWR 2 Archiv Radio <i>Documentary</i> distance: 118.25

Tab. 1: Three examples of recommendation requests and the three closest results including their Euclidean distances.

All 3 stations near to *BR Klassik* play similar content within the genres *Classical Music* and *Cultural*. The second radio station *Heart UK* is assigned to the genres *Classic/Dance/Pop-rock*, *Disco*, *Local/Regional*, *Dance/Dance-pop*, and *Showbiz*. The closest three stations are *Capital XTRA Reloaded* with the genres *Rap/Hip Hop/Reggae* with an Euclidean distance of 41.09, *3FM Isle of Man* with the genre *Hit-Chart* and an Euclidean distance of 86.34, and *FFH Rock* with the genres *Rock*, *Soft Rock*, *Grunge*, *Heavy Rock*, and *Rock & Roll* and an Euclidean distance of 88.29.

The only station close to *Heart UK* by genre is *3FM Isle of Man*. The other two stations can be considered not similar, being assigned to *Rap/HipHop/Raggae* versus *Rock*, *Soft Rock*, *Grunge*, *Heavy Rock*, *Rock & Roll*. The third radio station *LBC UK* is assigned the genres *Non-fiction*, *Local/Regional*, and *News*. The closest 3 non-variant stations are *Bayern 5 Plus* with the genre *Information* with an Euclidean distance of 117.97, *WDR 3* with the genres *Classical Musik* and *Cultural* with an Euclidean distance of 118.22, and *SWR 2 Archiv Radio* with the genre *Documentary* and an euclidean distance of 118.25.

All three stations near to *LBC UK* publish a lot of spoken content, as suggested by their genres *Local/Regional*, *News*, or *Documentary*.

As we could show and explain by example, the proposed station fingerprints in combination with the discussed similarity metric is a well-working approach for content-based station recommendations which outputs similar stations for a given station.

4.2 Archetypal Analysis

In order to distill station categories, we conducted an *Archetypal Analysis* on the fingerprints which is an unsupervised learning method used to extract representative individuals in a dataset. A data point is defined by its affiliation to k archetypes [CB94]. In soccer, for example, a player could be described by 10% defender, 50% midfielder, and 40% striker. We use this concept to find representative stations on basis of their fingerprints. In our case, the optimal amount of four archetypes was determined using the so-called *elbow criterion* [EL09]. Therefore, the residual sum of squares (RSS) for different numbers of archetypes is visualized in a scree plot (see Figure 4). The optimal number of archetypes is the one where the curve has its strongest bend.

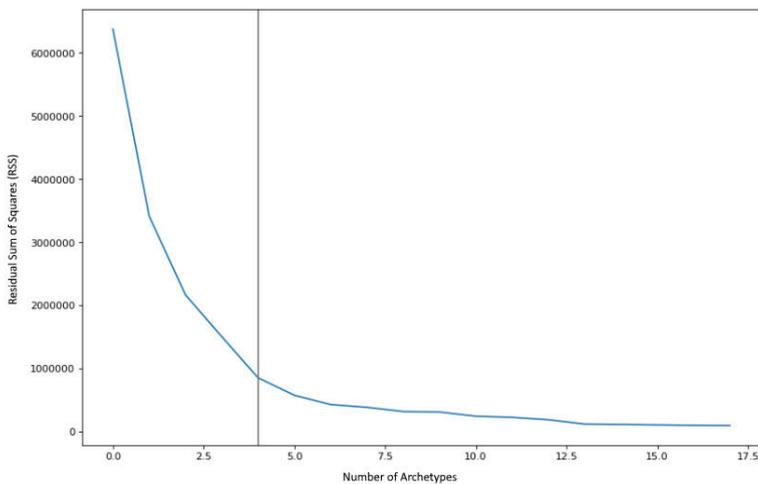


Fig. 4: Number of archetypes and corresponding RSS values. This so-called scree plot is used for the selection of the best number of archetypes.

Figure 5 shows a plot of all fingerprints, reduced to 2 dimensions by using a *Principal Component Analysis (PCA)* [Jo11], where stations are points coloured according to their genre. Archetypes are represented as black triangles. The first archetype is at [115.77, -164.82]. *Antenne P* defined as genre *Oldies* is the closest station to it with a distance of 19.86. Within a radius of an Euclidean distance of 150, 71 stations could be found in total.

SWRI RP is nearby next with a distance of 51.90, encompassing the genres *Pop Music* and *Regional*. The second archetype is at [292.96, 385.21]. *Noods Radio* is the closest station with a distance of 19.23 not listing any genres. Within a radius of an Euclidean distance of 150, 7 stations could be found in total. 95.3 *KGY Olympia* is the next nearest station with a distance of 41.56, classified as *AOR / Slow Rock / Soft Rock*.

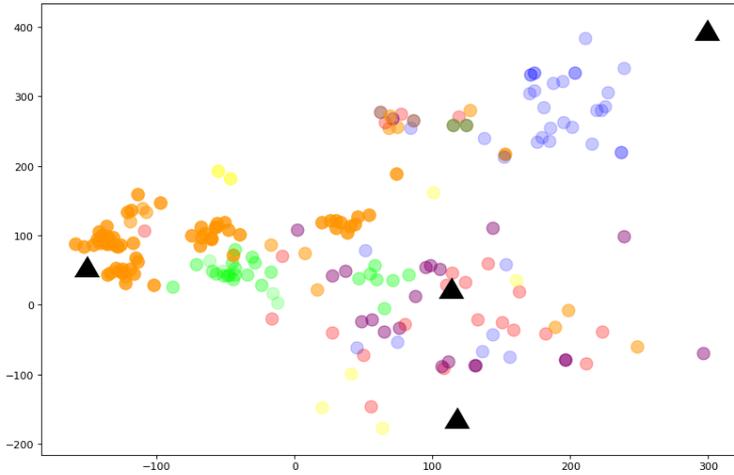


Fig. 5: Visualization of fingerprint vectors, using PCA. The dots mark the radio services, colored by genre. Black triangles represent archetypes.

The third archetype is located at $[-144.58, 50.05]$. *Heart Crawly* has a distance of 8.76 to it, encompassing the genres *Classic/Dance/Pop-rock*, *Disco*, *Local/Regional*, *Dance/Dance-pop*, and *Showbiz*. Within a radius of an Euclidean distance of 150, 175 stations could be found in total. The next other station providing genre metadata is *Heart Dorset* with a distance of 11.32, classified by the same genres.

The last archetype is at $[138.10, 60.20]$. *Bayern 2 Nord* is the closest station with a distance of 88.83 and genre *Cultural*. Within a radius of an Euclidean distance of 150, 48 stations could be found in total. The next other station offering genre meta data is *Classic FM* with a distance of 99.47 to the archetype, playing *Classical Music* and *News*.

In summary, the four found archetypes map well on genres and also target groups: The first archetype corresponds to stations with a target group that likes to listen to oldies / old pop music and demands for regional information. The second archetype has nearby stations with a very diverse programme that does not fit into specific music genres (*Noods Radio* for example plays almost everything from post punk to electro) and thus has a very broad target group. The third archetype represents stations with a focus on dance music whereas the fourth archetype is related to classical music and news. Interestingly, there is no specific archetype for contemporary pop music which could be due to the fact that this kind of music is more or less a mixture of existing genres.

4.2.1 Comparing Recommendations by Day Times

It is well known that radio stations change their programme over the day with, for example, fast wake-up music in the morning and slower, relaxing music in the evening. Our hypothesis is, that this effect also manifests in the fingerprints of a station extracted for different day times.

In order to test this hypothesis, we created three additional, normalized fingerprints per times of day, namely night, morning, and day. Night fingerprints consider audio samples from 09:00 pm to 05:00 am, morning fingerprints consider samples from 05:00 am to 09:00 am, and day fingerprints consider samples from between 09:00 am and 09:00 pm. See Figure 6 for a visualization of the different fingerprints using per-station trajectories. Comparing those three fingerprints to each other and to the whole-day fingerprints provides interesting insights into how radio stations change throughout the day.

E.g., large changes appear on *Gold 60s* which had comparably large distances between morning (square) and other times of the day (triangle, diamond) and whole day (circle) fingerprints, leading to the assumption that the style of content changes heavily between morning and rest of the day. This pattern is also visible in other stations' trajectories (e.g. *FFH Die 90er*, *FFH BrandNeu* and *106 Jack FM*). On the other hand, some stations show only minor (*FFH Workout*) or moderate (*Radio City Talk*) changes over the day.

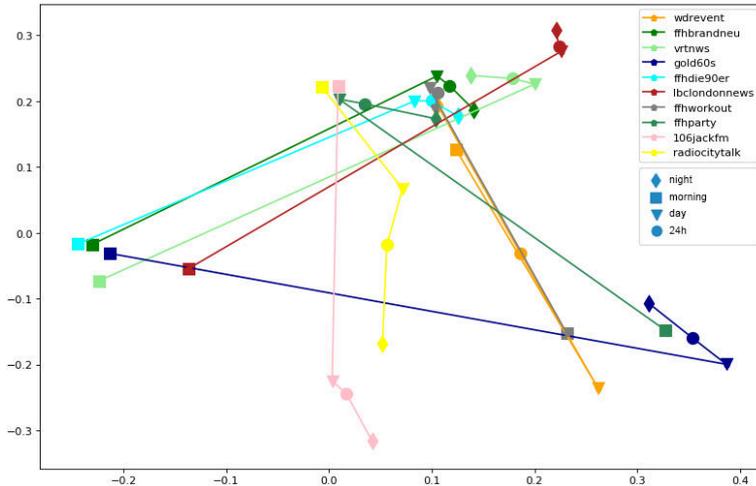


Fig. 6: Fingerprint trajectories visualizing deltas for stations for different times of day. Trajectory color corresponds to the station, circle shapes are whole-day fingerprints (24h), diamonds night, squares morning and triangles day time.

An additional hint for daytime-based changes visible in the extracted fingerprints is depicted in Figure 7. The day time and morning time archetypes (light blue, yellow) significantly differ from the night time archetypes (dark blue) which clearly shows that the assumed programme difference dependent on the time of day is visible in the fingerprint data-set as well.

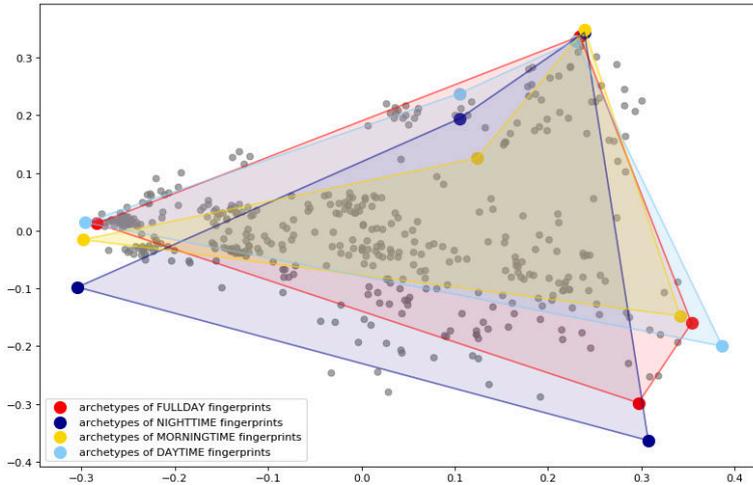


Fig. 7: Extracted archetypes for different times of day.

To summarize, we could prove our hypothesis, that time dependent radio content differences are visible in the fingerprint data by showing exemplary per-station differences and differences in the shapes and locations of the archetypes.

5 Conclusion

In this paper, we presented a holistic pipeline for deep radio station fingerprinting and recommendation. The first component utilized throughout the process is the *HRADIO Communication Platform*, which delivers metadata and bearer addresses of radio stations. The next component, the *DataCollector* recorded 461 radio stations over 24 hours, generating a total of 266,239 audio samples. Following this step, we trained a *Deep Neural Autoencoder* with *auDeep* in order to reduce each sample's dimensions. The *Fingerprinter* then clustered the samples and created a deep fingerprint for each service. The last component, the *Recommender*, is able to give recommendations depending on the calculated distance between the fingerprints.

We evaluated the *Recommender* by analyzing the recommendation results and noticed similar genres for close radio stations. Additionally, we conducted an *Archetypal Analysis* on the fingerprints, leading to four archetypes. By analyzing their closest radio stations, we noticed dissimilar target groups and genres between the archetypes and similar ones between the closest stations surrounding them.

Finally, we compared whole-day fingerprints to fingerprints extracted in the night, morning and day time. The current time of day seems to have a big influence on recommendations for some services. In contrast to that, others stayed somewhat constant throughout 24 hours. The data-set we created offers many more opportunities for further research and analysis. To further validate this, user studies should follow with which the quality of a recommendation can be quantified. In addition, the night-morning-day approach could be evaluated in far more depth.

We see three important practical use cases for the aforementioned techniques and findings:

- The proposed content-based recommender can complement existing recommender engines that mostly use collaborative filtering which is rather domain-unspecific. Within the HRADIO project, a station recommender system was developed, that uses a fingerprint-based recommendation module together with modules based on metadata, location and user behavior.¹⁰
- The described audio analysis pipeline can be used to automatically extract station and programme metadata (e.g. genre-like classifications) which is a big step forward considering the current situation regarding metadata availability in the radio domain.
- The extracted fingerprints and the proposed distance metric can also be helpful for radio stations that would like to check if their current play-out matches with the station's target programme and groups. This check can be conducted automatically and could provide even more detailed guidance like what to play in the next two hours to match the taste of a certain target group.

In summary, this approach provides valuable recommendations only on basis of audio signals and without the need of additional metadata. Moreover, the presented concepts are included and available within the open *HRADIO Communication Platform* and thereby enable also small businesses and radio stations to actively take part in a radio service landscape driven by smart services.

Bibliography

[AT07] Aranganayagi, S; Thangavel, K: Clustering categorical data using silhouette coefficient as a relocating measure. In: International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007). volume 2. IEEE, pp. 13–17, 2007.

¹⁰ Android App that uses the modularized station recommender: <https://play.google.com/store/apps/details?id=lmu.hradiu.hradioshowcase>

- [Be07] Bennett, James; Lanning, Stan et al.: The netflix prize. In: Proceedings of KDD cup and workshop. volume 2007. New York, NY, USA, p. 35, 2007.
- [Bu02] Burke, Robin: Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12(4):331–370, 2002.
- [ÇA07] Çataltepe, Zehra; Altinel, Berna: Music recommendation based on adaptive feature and user grouping. In: 2007 22nd international symposium on computer and information sciences. IEEE, pp. 1–6, 2007.
- [CAS18] Camacho, Lesly Alejandra Gonzalez; Alves-Souza, Solange Nice: Social network data to alleviate cold-start in recommender system: A systematic review. *Information Processing & Management*, 54(4):529–544, 2018.
- [CB94] Cutler, Adele; Breiman, Leo: Archetypal analysis. *Technometrics*, 36(4):338–347, 1994.
- [Ch15] Chang, Shiyu; Zhou, Jiayu; Chubak, Pirooz; Hu, Junling; Huang, Thomas: A space alignment method for cold-start tv show recommendations. In: Twenty-Fourth International Joint Conference on Artificial Intelligence. 2015.
- [DBS11] Dieleman, Sander; Brakel, Philémon; Schrauwen, Benjamin: Audio-based music classification with a pretrained convolutional network. In: 12th International Society for Music Information Retrieval Conference (ISMIR-2011). University of Miami, pp. 669–674, 2011.
- [EL09] Eugster, Manuel; Leisch, Friedrich: From spider-man to hero-archetypal analysis in R. 2009.
- [Fr17] Freitag, Michael; Amiriparian, Shahin; Pugachevskiy, Sergey; Cummins, Nicholas; Schuller, Björn: audeep: Unsupervised learning of representations from audio with deep recurrent neural networks. *The Journal of Machine Learning Research*, 18(1):6340–6344, 2017.
- [Fr19] Friedrich, Markus; Ebert, André; Hahn, Carsten; Schneider, Georg; Obermeier, Liza; Erk, Alexander; Jennes, Iris: A Distributed Metadata Platform for Hybrid Radio Services. In: International Conference on Innovations for Community Services. Springer, pp. 166–183, 2019.
- [GG14] Gwardys, Grzegorz; Grzywczak, Daniel: Deep image features in music information retrieval. *International Journal of Electronics and Telecommunications*, 60(4):321–326, 2014.
- [GLH19] Geschke, Daniel; Lorenz, Jan; Holtz, Peter: The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers. *British Journal of Social Psychology*, 58(1):129–149, 2019.
- [Jo11] Jolliffe, Ian: *Principal component analysis*. Springer, 2011.
- [La08] Lam, Xuan Nhat; Vu, Thuc; Le, Trong Duc; Duong, Anh Duc: Addressing cold-start problem in recommendation systems. In: Proceedings of the 2nd international conference on Ubiquitous information management and communication. ACM, pp. 208–211, 2008.
- [LS01] Logan, Beth; Salomon, Ariel: A Music Similarity Function Based on Signal Analysis. In: ICME. pp. 22–25, 2001.
- [Or18] Oramas, Sergio; Barbieri, Francesco; Nieto, Oriol; Serra, Xavier: Multimodal deep learning for music genre classification. *Transactions of the International Society for Music Information Retrieval*. 2018; 1 (1): 4-21., 2018.

- [Pa11] Pariser, Eli: *The filter bubble: What the Internet is hiding from you*. Penguin UK, 2011.
- [Pe11] Pedregosa, Fabian; Varoquaux, Gaël; Gramfort, Alexandre; Michel, Vincent; Thirion, Bertrand; Grisel, Olivier; Blondel, Mathieu; Prettenhofer, Peter; Weiss, Ron; Dubourg, Vincent et al.: *Scikit-learn: Machine learning in Python*. *Journal of machine learning research*, 12(Oct):2825–2830, 2011.
- [RRS11] Ricci, Francesco; Rokach, Lior; Shapira, Bracha: *Introduction to recommender systems handbook*. In: *Recommender systems handbook*, pp. 1–35. Springer, 2011.
- [RTG00] Rubner, Yossi; Tomasi, Carlo; Guibas, Leonidas J: *The earth mover’s distance as a metric for image retrieval*. *International journal of computer vision*, 40(2):99–121, 2000.
- [Si16] Siddiquee, Md Mahfuzur Rahman; Rahman, Md Saifur; Chowdhury, Shahnewaz Ul Islam; Rahman, Rashedur M: *Association rule mining and audio signal processing for music discovery and recommendation*. *International Journal of Software Innovation (IJSI)*, 4(2):71–87, 2016.