Art and Performance Letters (2021)
ISSN 2523-5877 Vol. 2 Num. 5

DOI: 10.23977/artpl.2021.020503
Clausius Scientific Press, Canada

Evolution and revolutionary trend of artists and schools based on cluster analysis

Yu Fan, Yuxin Chen, Jinrong Yao, Zhanchen Li, Zhangliang Song

College of Port, Coastal and Offshore engineering, Hohai University, Nanjing 210098, Jiangsu

Keywords: Music influence, directional network, cluster analysis

Abstract: From the beginning, music is a part of human society and an important part of cultural heritage. As part of an effort to understand the role of music in the collective human experience, we attempted to develop a way to quantify the evolution of music.We have developed a comprehensive evaluation model to evaluate the influence of music.We mainly take "the number of followers of a genre" and "the number of influencers" as the quantitative evaluation index of music influence.We establish cosine similarity model to measure music similarity. After data normalization, we can conclude that the music similarity of artists from different schools is greater than that between different schools.

1. Background

In recent years, more and more music has entered people's daily life.People not only enjoy the artistic experience brought by all kinds of music, but also have a new interest in the evolution of music culture.The significance of music is one of the important issues in the study of music philosophy and aesthetics.[1] Charles Rosen said that the meaning of music only exists in music itself.To understand music, we must introduce intention within the framework of music language itself, not outside music.It is of great significance to study the evolution of music in the process of social development by measuring the influence of music and understanding the evolution and revolutionary trend of music artists and music schools.

When artists create a new piece of music, there are many factors that affect them, including their inherent creativity, current social or political events, the opportunity to use new instruments or tools, or other personal experiences.We mainly understand and measure the impact of previous music production on new music and music artists. The impact here is the music impact we study.

Specifically, some artists can list a dozen or more other artists who they think have influenced their own music.Others argue that influence can be measured by the similarity between song features, such as structure, rhythm or lyrics.

In the process of music evolution, music types will be constantly updated and changed.Many songs have similar sounds, and many artists have contributed to a major shift in the music genre.Sometimes these changes are due to the influence of one artist on another.Sometimes, it is a change in response to external events (such as major world events or technological advances). We consider music development mainly by considering the interaction between artists.

2. Research and analysis

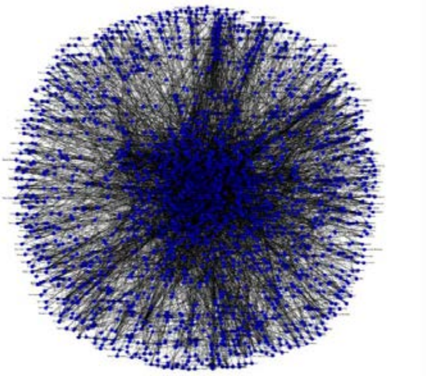
We need to use data to create a directed network of influencers and followers, which can reflect the influence of music. Obviously, the influence of artists' music will be reflected by the density of network connection lines.The embodiment of specific impact will find the relationship from the specific data to extract indicators, so as to quantitatively evaluate the impact.Select a music genre, start from the founder of the genre to create a sub network to quantify the music influence. Through the processing of the original data, we can get the indicators that can be used to judge the music influence of artists.For example, the most intuitive, the number of people affected by an artist.In the process of quantifying the influence of music, we can get the characteristics and laws of music.

The problem requires the development of music similarity model using known data.It is not difficult to see that there is a certain progressive relationship between the second question and the third question, so we decided to establish a similarity model between each song and use full_music_data set. The full representation of music features and types is directly expressed in data sets_music_As a parameter of music similarity, date is used to normalize the data to eliminate the influence of dimension on the data. In many methods of calculating similarity, we use cosine similarity to measure the similarity. Now let's look at the similarity relationship between artists within the genre and between schools. Let's look at the data, Artists of the same and different genres are replaced in the music similarity model for analysis and comparison.

3. Directional network and music influence evaluation model

3.1 Directional network

All the data given in the problem are mapped by mpai to followers through influencers.Large amount of data and high image density. As shown in Figure 1 below, we can see that there are many nodes and the mapping relationship is complex. It is difficult to see each mapping relationship in the whole network diagram.So we're going to make a subnet.



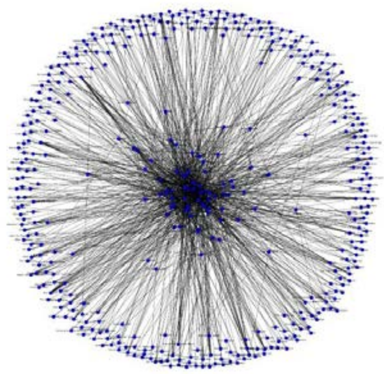


Figure 1: overall network diagram *Figure 2: directional subnet diagram in different years*

For the same type of music, we randomly selected the major schools of influencers as the blues data set to draw the directional network map of influencer followers. As shown in Figure 2 above, the screen is not intuitive enough.

In the data set of 1930, refine the selection again and make a directed network diagram, as shown in Figure 3.At this time, it can well reflect the directional network relationship of the same type of music genre in the same year.The remaining data can be selected according to the same method to obtain the directed network graph

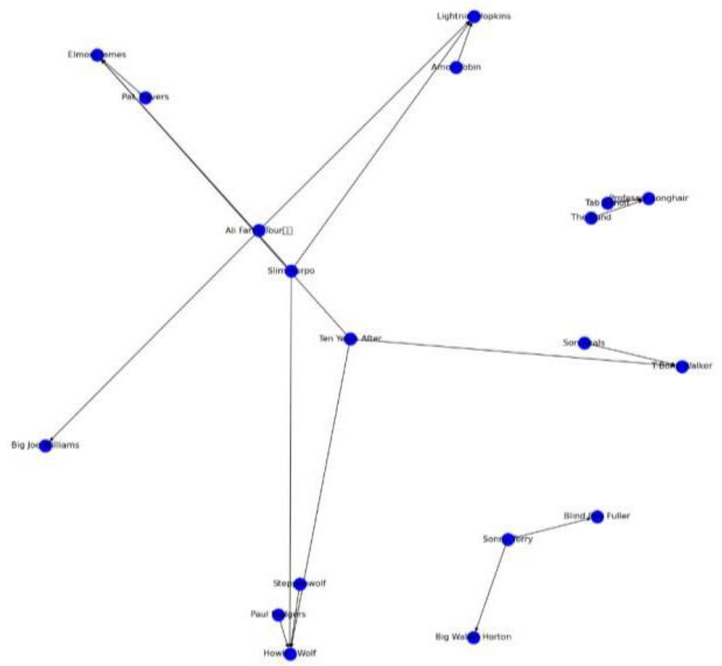


Figure 3: directional sub network of the same music type and year

3.2 Music influence evaluation model

We determine the relevant indicators of "music influence" in the sub network, establish the evaluation index system, and quantitatively describe the relationship between subnets. We use the combination evaluation method to calculate the weight of each index.

The process of developing parameters and quantifying music impact is shown in Figure 4

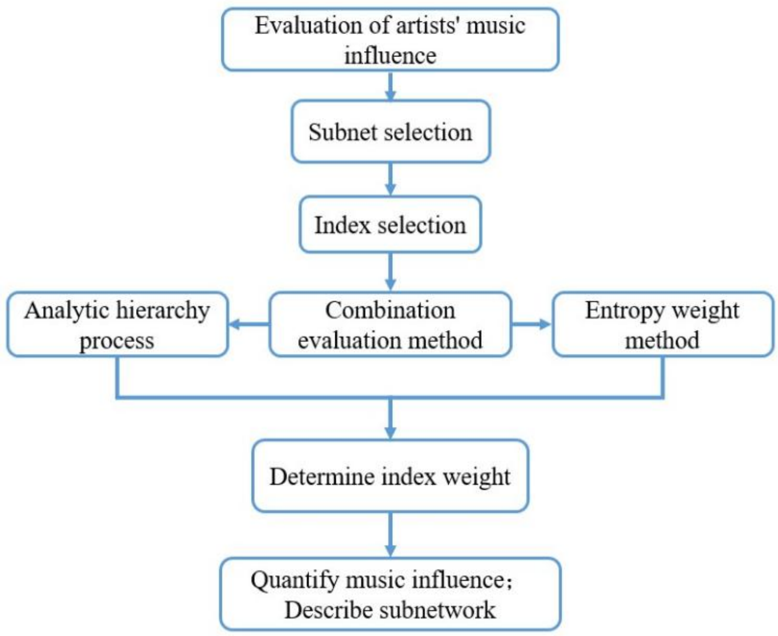


Figure 4: flow chart for quantifying music impact

4. Music similarity analysis model

According to the literature review, based on the research of Angelo Cesar Mendes da Silva and Wang Hailong, it is found that "relative date" and "year" have little influence on music similarity, so these two indexes are deleted and not included in the similarity evaluation index. After discussion and research, we use cosine similarity algorithm to calculate music similarity.

We group the data in the artist data dataset by genre.Because there are so many artists in each field, it is difficult to compare the music similarity between artists in the same field and artists in different fields. Therefore, sampling method is used to compare music similarity. In the previous 19 groups of musicians, 3 musicians in the same group were randomly selected by MATLAB software, the average value of the results is taken as the music similarity of musicians in the same field.Then select three musicians in the rest group, and calculate the similarity between musicians in the selected group and musicians in different fields.

Taking the first group as an example, the calculation results are as follows:

Table 1: Music similarity of musicians of the same school

type	The first pair	The second pair	The third pair	mean value
(1,1)	0.925	0.960	0.934	0.940

According to the results of similarity calculation, the similarity of musicians in the first genre is 0.94, which is very close to 1, indicating that the music styles of musicians of the same genre are very similar.The music similarity between the first genre and other genres is less than 0.94, which indicates that the similarity of music is low and the music styles of different genres are different.The similarity between the first category and the eighth and ninth categories is very close to 0.94, and the similarity is very high.The similarity of the 16 genres is only 0.737, and the difference of music styles is the biggest.

Through the analysis, it is concluded that the similarity of music in genre is greater than that between genres, but there may be similarities between genres of some genres, close to or even greater than that of music in genres.Through the analysis, it is found that there are 15 indicators for the description of musicians. The influence degree within and between schools is compared, because the more artists of a school are influenced by artists of a certain school, the greater the influence between the two schools.Therefore, we calculate the number of influencers in a genre and construct an influence matrix.

We use the artist data set to calculate genre similarity. First, we remove count from the dataset.Because unlike other indicators, count does not show the characteristics of a genre.In order to quantify the model, integer and Boolean data are unified into floating-point data in [0,1].In order to extract and process musicians from the same genre, the string data of the genre is mapped to integers.At the same time, because the musicians of "unknown" musicians have no genres in fact, it is impossible to calculate the similarity between "unknown" musicians and other genres, so it is necessary to exclude them.We use the impact data set to calculate genre impact. At the same time, we need to delete data for musicians whose music type is "unknown."

The calculation formula of similarity is still the same as that of the music similarity model in question 2. The calculation of influence degree is based on the statistics of the number of schools influencing each other to construct the interaction matrix

$$M = \begin{bmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,18} & m_{1,19} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,18} & m_{2,19} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ m_{18,1} & m_{18,2} & \cdots & m_{18,18} & m_{18,19} \\ m_{19,1} & m_{19,2} & \cdots & m_{19,18} & m_{19,19} \end{bmatrix}$$

$$S = \begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,18} & s_{1,19} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,18} & s_{2,19} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s_{18,1} & s_{18,2} & \cdots & s_{18,18} & s_{18,19} \\ s_{19,1} & s_{19,2} & \cdots & s_{19,18} & s_{19,19} \end{bmatrix}$$

We study the similarities between schools and musicians to reflect their similarities.The similarities are as follows:

generl	Pop/R	Electro	Avant-	Blues	Jazz	Count	Comec	R&B;	Regga	Classic	Latin	Vocal	Folk	Easy Li
Pop/R	1	0.964	0.541	0.873	0.325	0.972	0.825	0.805	0.917	0.433	0.661	0.582	0.859	0.856
Electro	0.964	1	0.632	0.885	0.31	0.944	0.799	0.818	0.947	0.425	0.701	0.596	0.826	0.865
Avant-	0.541	0.632	1	0.732	0.584	0.603	0.39	0.424	0.569	0.448	0.74	0.676	0.514	0.767
Blues	0.873	0.885	0.732	1	0.331	0.853	0.656	0.76	0.904	0.398	0.883	0.664	0.797	0.88
Jazz	0.325	0.31	0.584	0.331	1	0.429	0.417	0.318	0.17	0.853	0.316	0.412	0.305	0.626
Count	0.972	0.944	0.603	0.853	0.429	1	0.845	0.737	0.863	0.54	0.708	0.672	0.845	0.886
Comec	0.825	0.799	0.39	0.656	0.417	0.845	1	0.793	0.75	0.556	0.471	0.634	0.873	0.736
R&B;	0.805	0.818	0.424	0.76	0.318	0.737	0.793	1	0.813	0.402	0.561	0.447	0.685	0.691
Regga	0.917	0.947	0.569	0.904	0.17	0.863	0.75	0.813	1	0.282	0.731	0.552	0.802	0.844
Classic	0.433	0.425	0.448	0.398	0.853	0.54	0.556	0.402	0.282	1	0.378	0.443	0.422	0.601
Latin	0.661	0.701	0.74	0.883	0.316	0.708	0.471	0.561	0.731	0.378	1	0.741	0.597	0.795
Vocal	0.582	0.596	0.676	0.664	0.412	0.672	0.634	0.447	0.552	0.443	0.741	1	0.739	0.687
Folk	0.859	0.826	0.514	0.797	0.305	0.845	0.873	0.685	0.802	0.422	0.597	0.739	1	0.774
Easy Li	0.856	0.865	0.767	0.88	0.626	0.886	0.736	0.691	0.844	0.601	0.795	0.687	0.774	1

Figure 5: Genre similarity thermography

The influence of each genre is as follows:

generl	Pop/Roc	Electro	Avant-G	Blues	Jazz	Country	Comedy/	R&B:	Reggae	Classic	Latin	Vocal	Folk	Easy Li
Pop/Roc	0.808	0.011	0.002	0.025	0.015	0.023	0.001	0.060	0.006	0.002	0.004	0.014	0.020	0.001
Electro	0.450	0.332	0.015	0.001	0.041	0.002	0.001	0.079	0.030	0.014	0.006	0.002	0.003	0.000
Avant-G	0.297	0.000	0.189	0.000	0.054	0.000	0.000	0.054	0.000	0.297	0.000	0.000	0.000	0.027
Blues	0.089	0.000	0.000	0.701	0.062	0.019	0.000	0.094	0.000	0.000	0.000	0.028	0.006	0.000
Jazz	0.045	0.000	0.000	0.008	0.832	0.005	0.000	0.033	0.000	0.001	0.017	0.040	0.000	0.004
Country	0.160	0.000	0.000	0.002	0.007	0.770	0.001	0.012	0.000	0.001	0.000	0.011	0.033	0.000
Comedy/	0.200	0.000	0.000	0.000	0.033	0.007	0.673	0.027	0.000	0.000	0.000	0.027	0.033	0.000
R&B:	0.066	0.003	0.000	0.027	0.041	0.002	0.001	0.788	0.002	0.000	0.002	0.054	0.002	0.000
Reggae	0.032	0.000	0.000	0.004	0.015	0.001	0.001	0.140	0.785	0.000	0.000	0.007	0.000	0.000
Classic	0.309	0.012	0.062	0.000	0.062	0.000	0.000	0.025	0.000	0.210	0.012	0.185	0.012	0.000
Latin	0.186	0.001	0.000	0.001	0.058	0.003	0.000	0.044	0.014	0.001	0.570	0.063	0.001	0.004
Vocal	0.090	0.000	0.000	0.006	0.126	0.008	0.004	0.045	0.001	0.012	0.010	0.642	0.004	0.005
Folk	0.269	0.000	0.000	0.039	0.018	0.163	0.008	0.010	0.000	0.000	0.000	0.026	0.439	0.000
Easy Li	0.167	0.000	0.000	0.000	0.222	0.000	0.000	0.000	0.000	0.019	0.093	0.037	0.000	0.333

Figure 6: Temperature spectrum of influence degree

From the analysis of thermal images, it is not difficult to see that the most influential school is the same school, which interacts with each other in learning.This is consistent with our habitual understanding.When we explore and learn in a certain field, we should first refer to the previous experience in this field.

In previous studies, we found that genre popularity was very high, and genre size was the largest.We think it has something to do with the characteristics of "pop / rock", which are easy to understand and popular with the audience.Children's songs are mainly heard by children, but the development of a school mainly depends on adult musicians.Natural children has little influence on other schools.

In the data set of "data by artist", we select the musician who has the most followers of each genre as the representative of the genre, and uses the data represented by the genre to conduct systematic cluster analysis

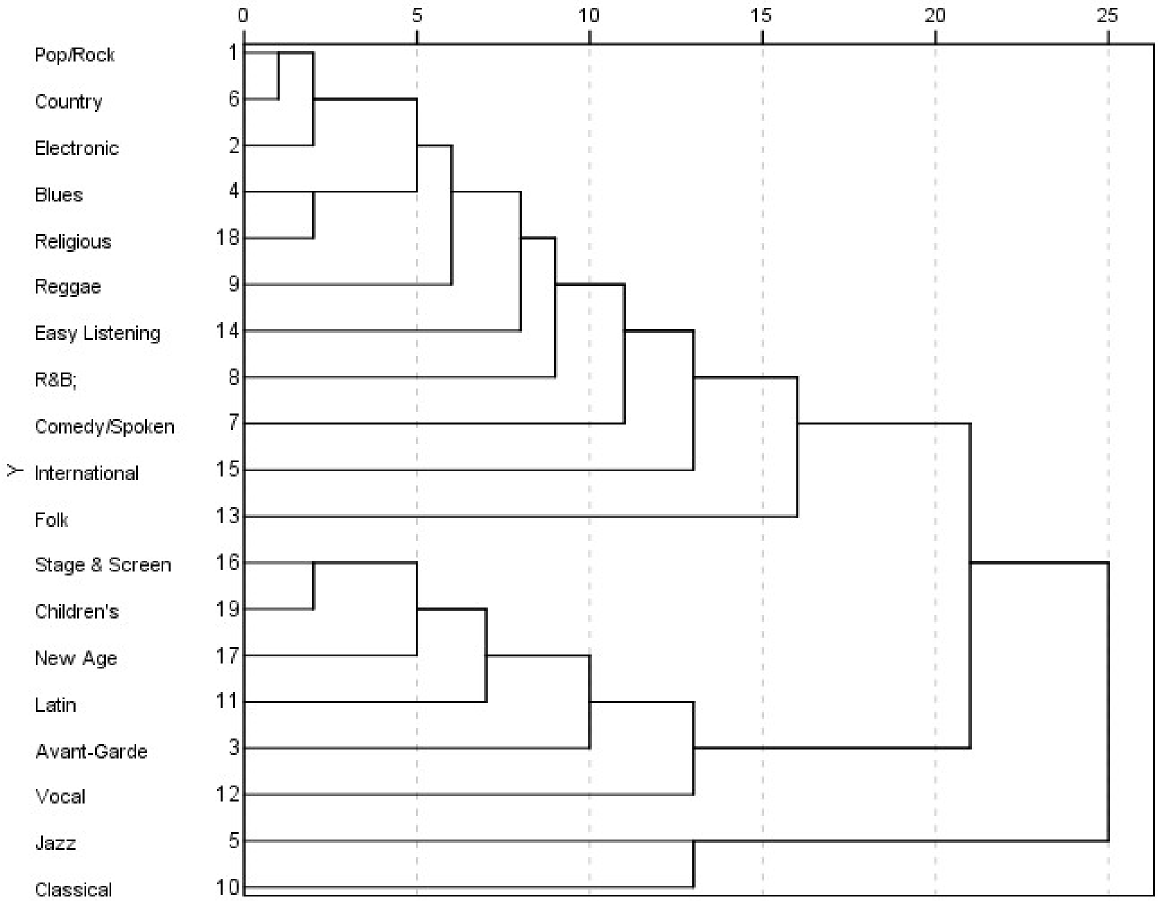


Figure 7: System clustering results

We can see that the two indicators "duration"_MS' and 'popularity' play the most important role. We use them as graphs of cluster analysis results.

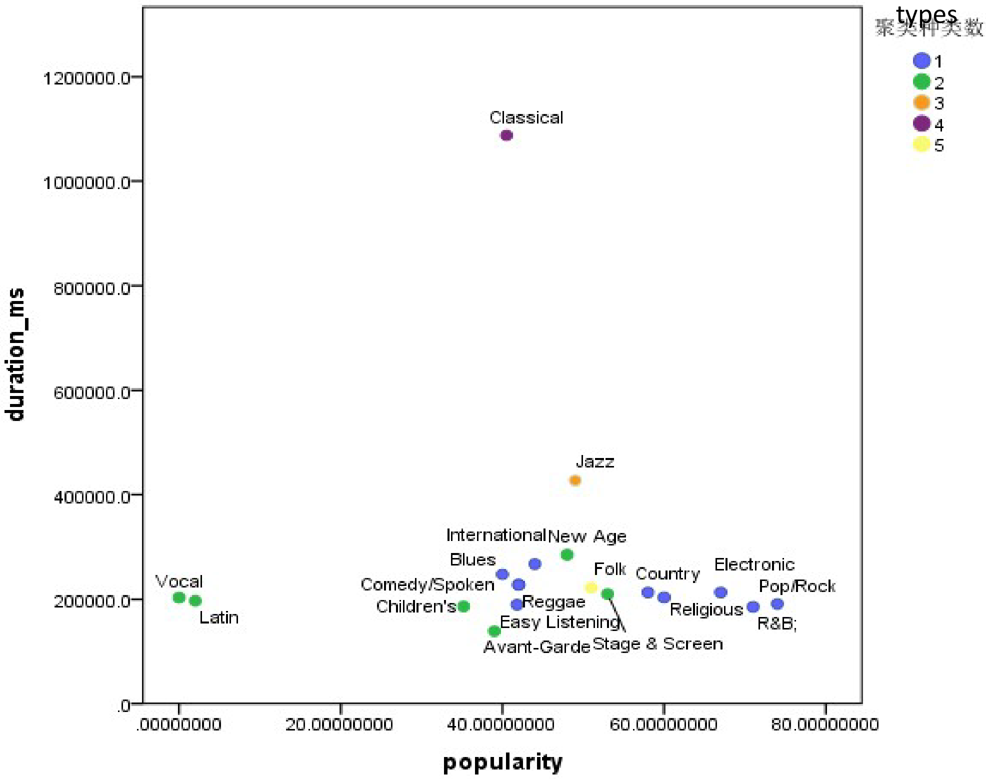


Figure 8: Graph clustering results

5. Conclusion

We establish followers in the influence network through problem analysis and orientation. We use directional network to analyze the similarity and influence between schools. We find that due to the interaction between musicians, one music genre can begin to be similar to other schools in some music indicators. Even because of the change of the music index of the genre, the popularity of the genre has increased,for having heard it many times.At the same time, it is precisely because of the mutual influence of musicians that a hundred schools of thought and new schools have emerged in the field of music.The combination of the influence of music and the development of the times will lead to the emergence of new music forms, such as rock music, which is influenced by blues and country music, and combines more and more American teenagers who are eager to resist the background of the times.

With more and more data, it is necessary to develop models in data mining, simulate the original human learning process, complete data mining with models, and apply statistical and machine learning methods.Data mining can help us find more valuable information in massive data.At the same time, huge data can be divided into subnets one by one, and valuable information can be extracted by the interconnection between networks.The abundance of data puts forward new requirements for the selection of indicators in the study of music influence and genre influence, and more representative evaluation indicators should be formulated.

References

[1] Jia Bo. Reflections on Charles Rosen's "musical significance" thesis [J]. Qilu art garden, 2020 (06): 14-20.
[2] Mendes A, Mauricio Archanjo Nunes Coelho, Neto R F. A Music Classification Model based on Metric Learning Applied to MP3 Audio Files [J]. Expert Systems with Applications, 2019.

[3] Wang Hailong, Liulin, Lin Min, Pei Dongmei. Recommendation algorithm for music Personalization Based on MIR and K-means label clustering [J/OL]. Journal of Jilin. University (Engineering Edition): 1-7 [2021-02-07].
[4] Buda, Andrzej; Jarynowski, Andrzej (December 2010).Life time of correlations and its applications. Wydawnictwo Niezalezne. pp. 5-21.ISBN9788391527290.