Advances in Economics, Business and Management Research, volume 139
International Conference on Economics, Management and Technologies 2020 (ICEMT 2020)

Clustering 'Goods' on the Two-Side Market - on the VK Social Network

Andrei Plotnikov[1,*] Hiroko Kawamorita[2]

[1]*Perm National Research Polytechnic University, Perm, 614000, Russia*
[2]*Ondokuz Mayis University, Atakum-Samsun, 55200, Turkey*
[*]*Corresponding author. Email: plotnikov-av@mail.ru*

ABSTRACT
The article considers the concept of "two-side market", highlighting its main features, criteria and characteristics. A model of a multilateral platform is identified, which is a business model of these types of markets. The largest Russian social network VKontakte (VK) is considered from the point of view of a two-side market. The main difficulty in the analysis of two-side markets is the determination of external effects and assessment of their significance for each of the user groups. Researchers are already working in this direction, offering various tools, mainly based on relative indicators and surveys. The paper presents an empirical analysis of digital marketing services presented on the social network VKontakte product storefront. A total of 19,765 'products' from 3,127 thematic communities were collected. Search engine optimization, social media marketing, contextual advertising (Yandex Ads & Google Ads) were selected as subjects. These groups are characterized as commercial communities that provide services in the areas of search engine optimization, social media marketing, contextual advertising in the b2b, b2c sectors. The research methods are cluster analysis, Word2Vec. The results of the study determine that the model does not understand the concepts of contextual and targeting. The difference is formed only since these words when teaching the model had a different environment, respectively, they were associated to different vectors.
Keywords: *two-side market, social network, digital economy, Internet marketing, services, social media marketing*

1. INTRODUCTION

Two-side markets are a relatively new subject of research by scientists. Despite the seemingly, at first glance, simplicity, these markets differ from the classical ones in that the seller works immediately for two groups of buyers who consume different goods. In this case, the seller is far from a classic reseller or intermediary, he is a so-called platform through which two groups of consumers are interconnected. The main feature and reason for studying two-side markets is their inconsistency with classical economic theory. An example of such a failure to comply with economic laws in two-side markets is the establishment of a zero price for one of the consumer groups, which will lead to ultimate profit maximization due to increased demand from another consumer group. Back in the middle of the last century, markets were identified in which both groups of agents did not intersect in commodity-money relations. It was the relations established between them that mattered, for example, between the employee and the company. A characteristic feature of such a market is a two-side exchange, that is, both parties receive benefits from each other. Moreover, often, the price of a 'product' can be zero. [1]

It should be noted that a two-side market is a concept that defines the system of circulation of goods between consumer groups. A specific set of devices and software that creates a platform for integration and trade is called a multilateral platform [2]. A two-side market may appear on one of the platforms, and then migrate to other platforms. And vice versa, the market can exist on any of the platforms, and after the appearance of a new one, more interesting in terms of cost reduction, it may switch to it.

1.1. Related Work

Due to the relative novelty of the concept and the lack of in-depth research, the term "two-side markets" is not fully formed today due to the dynamism of the market. An up-to-date definition of "two-side markets" was proposed by Tirole J. and Rochet J. (2003) [3]. Currently, the generally accepted types of two-side markets are a "one on one" market and a market with the participation of the platform. Two-side markets belong to the concept of the digital economy (initially the concept was considered by Tapscott D. (1996) [4] and after 5 years by Mesenbourg T. (2001) [5]), and they are independent of geographical location. A fresh look at ideas about the development of two-side markets is presented in the works by Wang Z. (2019) [6], Jabbour C. (2019) [7], Bajo-Buenestado R. (2019) [8], Dizdar D. (2019) [9], Jang D. (2018) [10] and Frishammar J. (2018) [11].

Multilateral platform strategies can be varied. But the most important problem is the decrease in demand of one group in the event of a decrease in demand from other groups, especially when this happens regardless of the cost of services or goods offered. The chicken-and-egg problem is the main one for any platform, it is important to develop one's own set of tools that would attract the necessary volume to the user groups of the platform, thus triggering network effects between them.

The two most popular and effective strategies are the strategy of price differentiation and investing in users. We have already examined the first on the example of free television channels for the viewer, which are more than paid for by advertisers. The investment strategy is to invest in one of the parties of the platform in order to get the most return from another group of users or from another group of goods and services [12].

Social networks are mentioned as classical multilateral platforms. Neti S. (2011) [13] in his work considers a social network as an opportunity available to a manufacturer for cooperation with a potential consumer. Mangold W. G., Faulds D. J. (2009) [14] determine that the development of social networks has allowed one person to communicate with hundreds or even thousands of other people about products, services and companies that provide them. Chung C., Austria K. (2010) [15] in their work consider that social networks enable marketers to increase trust and build consumer loyalty.

Social Media Marketing

In this literature research, over 5000 metadata from the largest abstract and citation database of peer-reviewed literature specifically Elsevier were retrieved. Those were all examined in topic related literature between the year 2012 and 2020. The overlay visualizations are presented by using the VOS viewer software as shown in Figure 1 and 2 to support co-occurrence network analysis by authors and co-authors. Also, the important terms were extracted from the relevant literature. Finally, "online marketing", "digital economy", and "social media marketing" were the keywords matching the title of the article and which are most relevant to this research.

Each year has a different number of publications, however the number of research related to the above 3 keywords increased after 2015 as this is a rapid growing but still a relatively new and very broad field of scientific research. The overlay visualization shown in Figure 1 was created by the titles and "media", "future", "characteristics", "challenges", and "problems" were the most studied concepts in the aspect of their relationships with social network. Whereas the Figure 2 was created by the titles and abstracts, which displays the increased number of studies in "social network" such as Facebook on the effect of social media interactions between consumers in various sectors including medical care, finance, education, and technology. Unlikely to the traditional network, two-sided networks create dependency in consumers which require providers to scale both sides of the market which is often challenging. Successful examples of such technology companies are known worldwide including Amazon, Apple, Google, Facebook, and Uber.

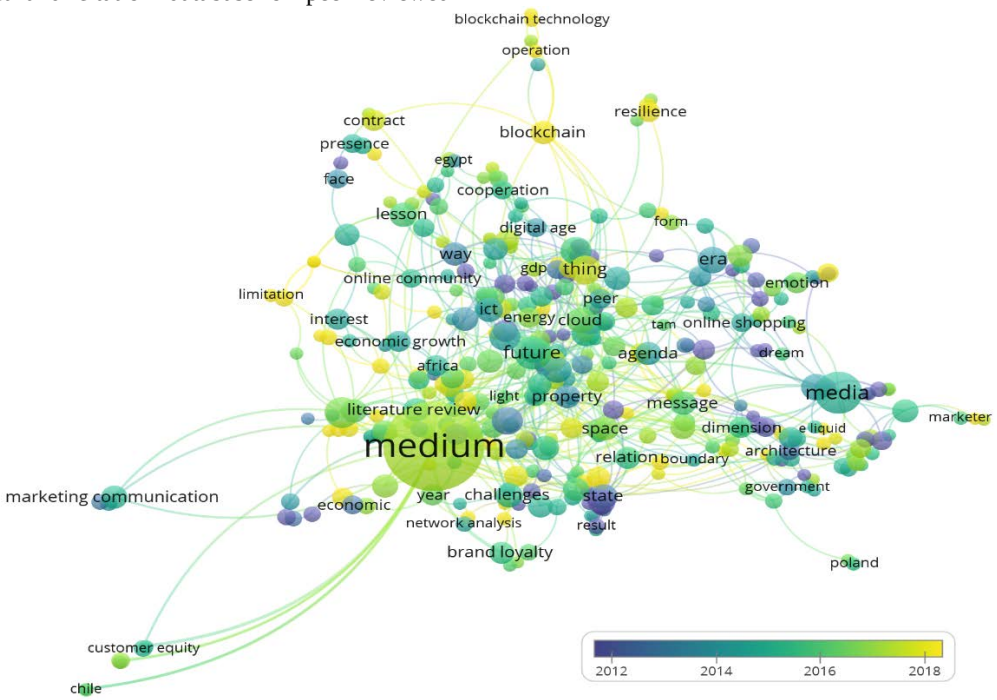

Figure 1 Overlay visualization (By Titles)

The largest and most advanced among Russian-language social networks is VKontakte. Using it as an example, we can study network effects and the movement of services between consumer groups. The possibilities of social

network VKontakte are considered by A.A. Morozova [16], the classification of promotion methods is considered in the work of R.A. Ganeev [17], the clustering of communities is considered in the work by S.I. Suslov [18], the analysis of users in terms of their professional interests is presented in the work by A.M. Obzhorin [19].

Let's start with the official advertising on the social network. The parties in this case will be a group of users and a group of advertisers. Users receive from a social network mostly free services (a means of communication, maintaining business pages, dating, etc.), while advertisers, offering their services to a group of users, pay the lion's share of social network revenues. In this case, we can talk about the platform using the strategy of income differentiation, when, attracting an increasing number of users, investing a lot of money in the development of the interface and the site functionality, the social network does not charge for this directly. All the expenses are covered by advertising contracts. When using the platform, users don't think about how much their actions affect certain advertising contracts. At the same time, a group of users also actively uses offers from the second group, acquiring stickers, paying bonuses in games or purchasing certain third-party goods.

Another option for using a social network from the point of view of a multilateral platform is crowdfunding. This process is a public funding of various projects. For such a process, a platform is necessary, and social network VKontakte is used for these purposes. In this case, the parties are the users and the authors of a business idea or a project. The interaction is carried out by way of obtaining means by the authors of the idea for the development of their business or organization. In exchange for this, the donors (a group of users) can count on a reward, which is negotiated in advance. If this is a kind of business idea, then as a result the group of users can receive some dividends. If the idea is creative, for example, recording a music album or publishing a book, then a bonus may be receiving these materials before their official release or signed directly by the performers.

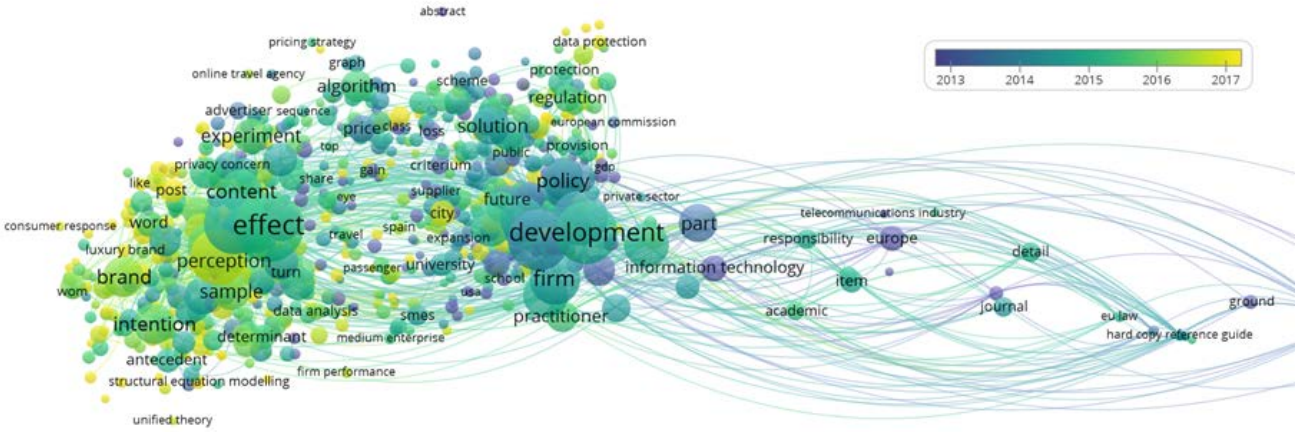

Figure 2 Overlay visualization (By Titles and Abstracts)

Communities in social networks is a very multifaceted phenomenon of a multilateral platform. User groups are the basis, they are united in communities based on common interests. In turn, the group of community administrators depends on how much the information they offer is interesting to users. If the community offers relevant data, creative ideas, can attract a new audience, this opens up new opportunities for it. For example, attracting advertisers to the community site. If the community is large and has a permanent audience, its authors can collaborate with third-party advertisers to earn money from the community. There is an interaction of external network effects between all groups united in a community on a single platform. Users consume data and information they are interested in through a single platform, while simultaneously receiving a portion of advertising integration. Next, we turn to the empirical part of the study.

2. RESEARCH METHODS

The research objective expresses the intention to classify 'goods' as services of social network VK by different clustering methods.

In September 2015, VK social network made a separate functional for selling goods and services - an online store within communities. Accordingly, most commercial communities used this VK e-commerce feature. As a result, we collected 19,765 'products' from 3,127 thematic communities. We have selected the following subjects: search engine optimization, social media marketing, contextual advertising (Yandex Ads & Google Ads). These groups are characterized as commercial communities that provide services in the areas of search engine optimization, social media marketing, contextual advertising in the b2b, b2c sector.

Data on product categories (services) of VK social network communities were collected during August and September 2019 using SaaS Target Hunter.
The software libraries used are: pandas, re, numpy, matplotlib.pyplot.
In conducting the study, we performed the following actions:
1. Loading a data file;
2. Determining the procedure for tokenizing the text field and checking its operation;
3. Tokenization of the "Name" field and verification of the result;
4. Class definition for counting the "aggregated" vector of words included in the text field;
5. Creating a Word2Vec model;
6. Creating a dictionary of words and their corresponding vectors;
7. Accumulating aggregated vectors corresponding to group names in the name_means array;
8. Clustering the aggregated vectors according to the principle of maximum similarity;
9. Adding a column with the obtained cluster number to the main dataset and displaying a graphical representation of the distribution by cluster;
10. Insuring the clustering adequacy: setting the keywords that appear in the "Title" field to see which clusters these groups are in.

3. RESULTS AND DISCUSSION

Distribution of services received by 10 clusters (Fig. 3).

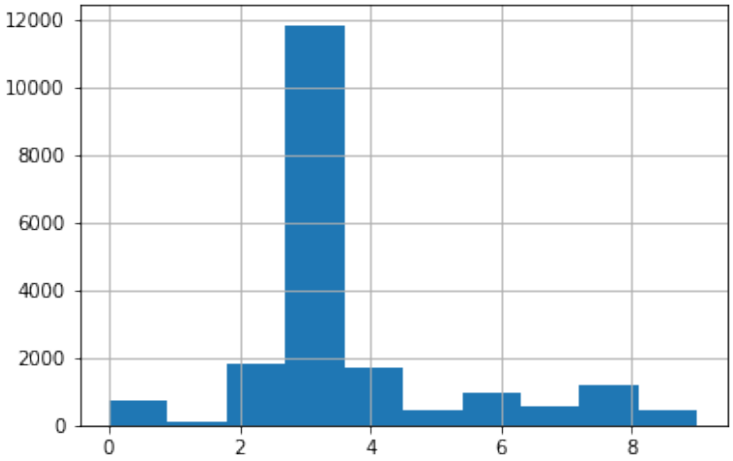

Figure 3 Distribution of services by clusters

Next, we verify the adequacy of clustering. We set the keywords that appear in the "Title" field to see at which clusters these groups are at Fig. 4.

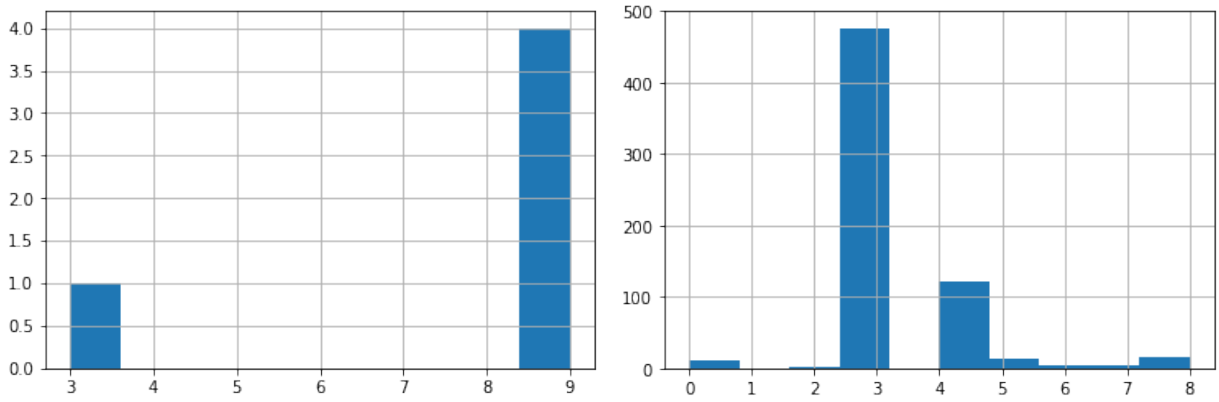

Figure 4 -The clusters presented according to: Google (left) and SMM (right)

We define the average cost of services for clusters. To begin with, we replace the obviously implausible price values

(<10 rubles and> 100 000 rubles) with median price values. We can see how the cost is distributed over the entire sample in Fig. 5.

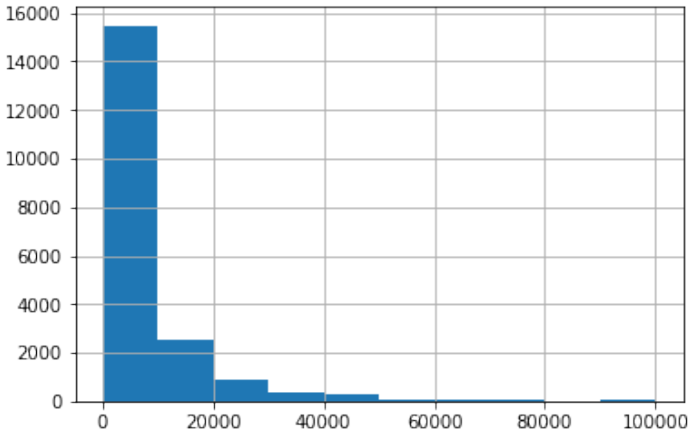

Figure 5 Average cost of services by cluster

We define a function for counting and issuing the most common words in the cluster name and calculate the average cost for the cluster by designating the name of the cluster with the most common word in it. We display the result in graphical form (Fig. 6).

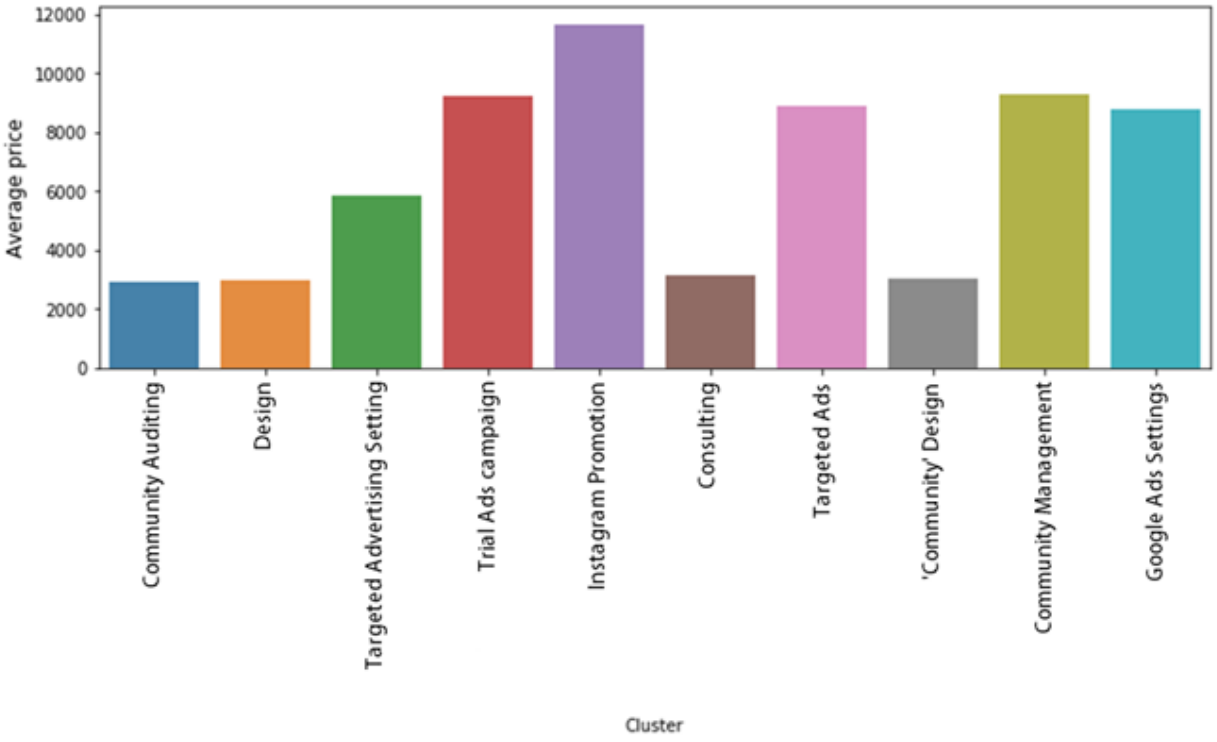

Figure 6 Average cost among frequently mentioned cluster services

Let us display the dependence of the price of a service on a cluster in the form of a “plotbox” graph. In this graph, the median (50 percentile) is shown by the yellow line, 25 and 75 percentiles are the boundaries of the box. The whiskers represent the entire scatter of points except the outliers, that is, the minimum and maximum values that fall into the interval (Q1 - 1.5IQR, Q3 + 1.5IQR), where IQR = Q3 - Q1 is the interquantile range. Dots on the graph indicate outliers - those values that do not fit into the range of values specified by the mustache of the graph. We display the result in graphical form in Fig. 7.

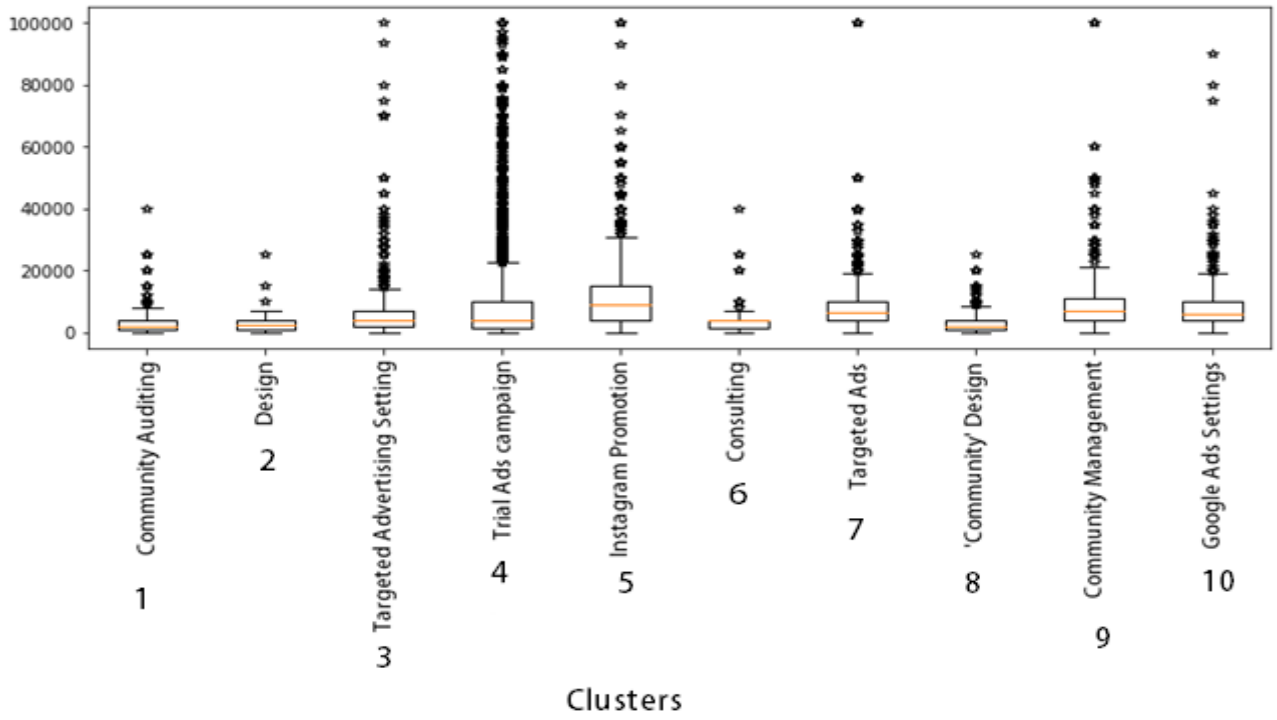

Figure 7 Distribution (plotbox) of prices for services

Now let's see how the price of cluster services changes over time. To do this, we group all the sample data by periods (quarter) and display them on the graph (Fig. 8).

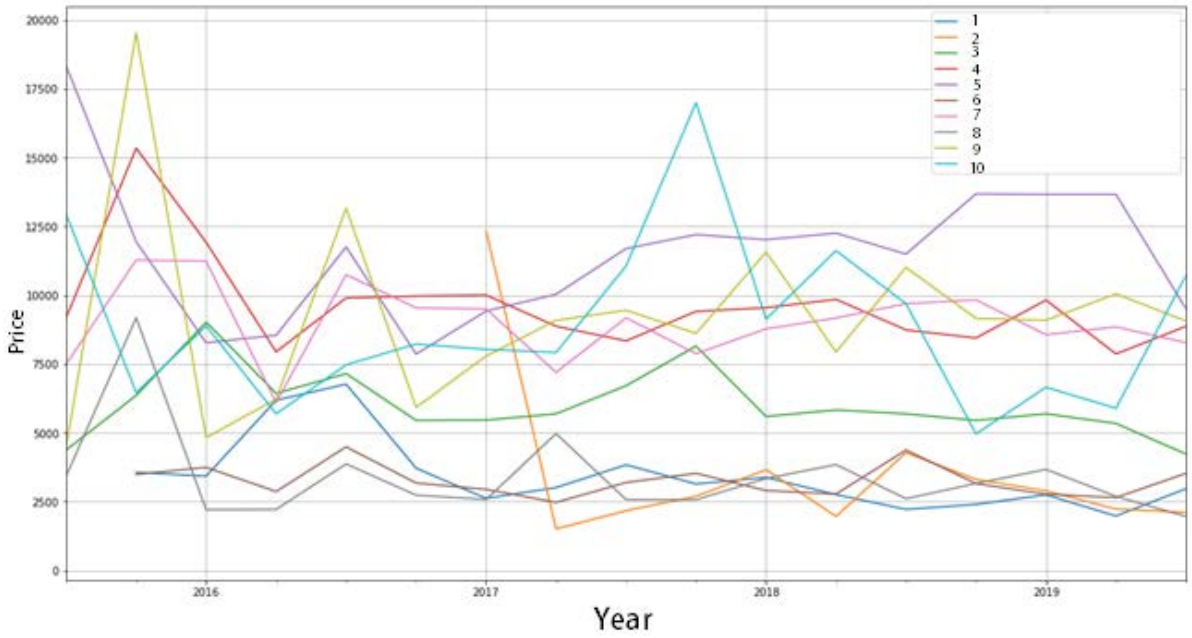

Figure 8 Change in prices of services by cluster over time

Let us consider in more detail cluster 2, containing package service offers. Let's see the TOP10 of the most frequently used words in the names of groups that fall into the cluster 'Targeted Advertising Setting': 'Targeted Advertising Setting', occurs 109 times; 'Community Design', 108; 'The design of the group', 80; 'Setting Yandex Direct', 56; 'Setting up advertising', 43; 'Setting up contextual advertising', 35; 'Setting Yandex.Direct', 34; 'Community Design', 25; 'Design of social networks', 21; 'Design of the VK group', 17

4. CONCLUSION

Despite the fact that 'Design of the VK group', 17 is present in the cluster - this does not affect the cluster, as targeted advertising can also be mentioned in the context of social networks. The model does not understand the concepts of contextual and targeted. The difference is formed only since these words when teaching the model had a different environment, respectively, they were associated to different vectors. From the point of view of common sense, perhaps these clusters just need to be combined, and some others, on the contrary, should be disunited. If a different number of clusters were chosen, they could be in the same one. In fact, this is the eternal compromise between model accuracy and selectivity.

ACKNOWLEDGMENT

The study was supported by the grant from the President of the Russian Federation for state support for research by young Russian scientists – Candidates of Sciences (project MK-698.2019.6).

REFERENCES

[1] Balandina M.S., Baskakova I.V. (2016). Bilateral markets: definition of a concept, key characteristics and assessment tools. Journal of new economy, 2 (64): 12-20.

[2] Kovalenko A.I. (2016). Research issues of multilateral platforms. Modern competition, 3 (57): 64-90.

[3] Rochet J. C., Tirole J. (2003). Platform competition in two-sided markets. Journal Of The European Economic Association, Vol. 1. No. 4: 990-1029.

[4] Tapscott D. (1996). The digital economy: Promise and peril in the age of networked intelligence. New York: McGraw-Hill, Vol. 1.

[5] Mesenbourg T. L. (2001). Measuring the digital economy. US Bureau of the Census.

[6] Li B. G., McAndrews J., Wang Z. (2019). Two-Sided Market, R&D and Payments System Evolution. International Monetary Fund. 48 p.

[7] Jabbour C. et al. (2019). Spatial data infrastructure management: A two-sided market approach for strategic reflections. International Journal of Information Management. Vol. 45: 69-82.

[8] Bajo-Buenestado R., Kinateder M. (2019). Armstrong meets Rochet – Tirole: On the equivalence of different pricing structures in two-sided markets. Economics Letters. Vol. 177: 43-46.

[9] Dizdar D., Moldovanu B., Szech N. (2019). The feedback effect in two-sided markets with two-side investments. Journal of Economic Theory. Vol. 182: 106-142.

[10] Jang D. C., Kim B., Lee S. Y. (2018). A two-sided market platform analysis for the electric vehicle adoption: Firm strategies and policy design. Transportation Research Part D: Transport and Environment. Vol. 62: 646-658.

[11] Frishammar J. et al. (2018). Digital strategies for two-sided markets: A case study of shopping malls. Decision Support Systems. Vol. 108: 34-44.

[12] Muravsky D.V., Yablonsky S.A. (2015). The principles of the multilateral platform. Innovations, 3 (197): 45-49.

[13] Neti, S. (2011). Social Media and its Role in Marketing. International Journal of Enterprise Computing and Business Systems, 1 (2): 1-15.

[14] Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. Business horizons, 52 (4): 357-365.

[15] Chung, C., & Austria, K. (2010). Social Media Gratification and Attitude toward Social Media Marketing Messages: A Study of the Effect of Social Media Marketing Messages on Online Shopping Value. Proceedings of the Northeast Business & Economics Association: 581-586.

[16] Morozova A.A. (2019). The multimedia capabilities of the social network VKontakte. Sign: the problem field of media education, 1 (31): 200-208.

[17] Ganeev R.A. (2016). Classification of methods of promotion "VKontakte". Economics, 5 (14): 53-57.

[18] Suslov S.I. (2016). Clusters of St. Petersburg political online communities in VKontakte. Bulletin of St. Petersburg State University. Series 12. Sociology, No. 4: 69-87.

[19] Obzhorin A.M. (2016). Professional identity of users of the VKontakte social network. Meteor City, 5: 36-41.