The Use of Digital Footprints to Create Psychographic Portraits for Increased Efficiency in Advertising Messages

Abstract:
This article describes the process of studying users’ behaviour in the digital space and creating psychographic portraits using an automated system inside the VKontakte (VK) social network and collecting users’ information. Within the framework of this study, the authors attempted to prove the possibility of implementing these data for small and medium-sized businesses’ goals and to offer the use of psychographic segmentation to create more personalized brand communication with consumers. The study is based on the analysis of data from an application of the VK social network (TIPI Psychological Test) that was developed by the authors and implemented as an online survey of 261 respondents from Nizhny Novgorod to identify their psychographic profile, as well as a database (MySQL) analysis of the activities of these people in the VK social network (their group subscriptions). Based on a comparison of the results of using these two methods, the authors conclude that it is possible to develop a system that uses users’ digital traces to create psychographic portraits and set up more effective targeted advertising for a specific consumer segment.

Keywords: digital portrait, digital behaviour, advertising, efficiency, technology, marketing, network

Introduction
The main direction of the current trends in consumer segmentation in Russia (Fukolova, 2018) and abroad (Adams & Mazin, 2018) is the automation of the processes of selection and creation of the audience base. To do this, the most often used tools are search and “tracking” systems for the online audience and analysis of audience
behaviour. The main mechanisms of these systems are the comparison of the user’s demographic profile with their actions and the subsequent offer of targeted advertising based on such data. At the same time, the use of big data technologies is most often perceived as an expensive and complex system available only to the largest corporations (Y. Huang et al., 2019). Analytics in small and medium-sized businesses is limited by users’ socio-demographic parameters and collecting data on their actions in the network. The motivational and psychological aspects of these actions are not considered because psychographic audience research has long been inaccessible to small companies and is now quite expensive.

With the development of social networks and an increase in user activity, it became possible to use their own pages on networks as data sources for building psychographic portraits of the consumer. Modern analytical methods and technical capabilities allow such research to be carried out by companies of any level, but poor coverage of these resources in domestic literature and the lack of a ready-made commercial solution prevent psychographic user segmentation from becoming a significant part of modern marketing in Russia. A similar approach has been used on the foreign market since 2013, following the publication of Kosinski’s (2013) study.

The use of psychographic segmentation creates opportunities to develop more personalized brand communication with consumers (Gajanova et al., 2019). This will allow businesses to appeal to their audience more accurately and increase audience engagement in response to ads (Seung & Quach, 2021). Within the framework of this study, the authors attempted to offer an additional layer of targeted appeal to the audience segmented by demographic characteristics by creating psychographic portraits of users. The possibility of implementing such an approach in the presence of a developed methodology and modern computer technology has been proven by researchers worldwide. Considering this fact, the current state of the Russian market and the growing interest in studying the characteristics of user behaviour in social networks, the aim of this study was to develop a similar project in our country – to segment consumers by psychographic characteristics based on their behaviour in the digital space using an automated system operating within the VKontakte social network and to prove the possibility of using these data for small and medium-sized businesses.

The object of this study is the residents of Nizhny Novgorod who using the VKontakte social network. The subject of the research, in this case, is the psychological characteristics of these users. According to the aim of this work, the general hypothesis of the study is that it is possible to create an automated system capable of recreating psychographic portraits of users based on their data obtained in the VKontakte social network.
The following provisions were put forward by the authors as private hypotheses:

− It is possible to create a psychographic portrait of a consumer using their digital “footprint”.
− Presumably, the system can be built inside the VKontakte social network without using external analytics systems.

With the aim, the authors formulated the following research tasks:

− to study the theoretical aspects of using psychographic segmentation and the use of psychological targeting in Internet marketing;
− to revise the methodology of psychographic segmentation in the digital space and adapt it to the Russian market;
− to create an application for collecting data for psychographic segmentation using an online survey of 261 residents of Nizhny Novgorod that helps to identify their Big Five psychographic profile using the TIPI questionnaire, and to choose an automated data collection system (MySQL) that analyses users’ group subscriptions within the VKontakte social network; and
− to link the results of the analysis of the TIPI Psychological Test and database to create a psychographic portrait of users based on their digital trail.

**Literature and Research Review**

In the early 2000s, marketers already had knowledge about the need to introduce psychographic segmentation in the online environment, but the lack of tools for collecting such information and segmentation methodology prevented the use of these scientific works for commercial purposes. The first study on the possibilities of psychographic segmentation of consumers, considering modern technical means, was the work “Psychographic Profiling of the Online Shopper”, released in 2003 (Vijayasarathy, 2003). The aim of the work was to study the relationship between users’ psychographic profiles and their behaviour in online stores. A significant breakthrough in these relationships was not revealed in the work; however, in-depth analysis proved that there is no significant correlation between product categories and classical socio-demographic segmentation.

For almost 10 years, the online shopping industry and online user behaviour were not studied by marketers because it was difficult to collect the necessary information. With the advent of social networks and the era of daily Internet use by most of the world’s population, it became possible to collect the necessary data for marketers of commercial organizations to use in their work (S. Huang et al., 2012). Subsequent works were released in 2011: “Modeling the Psychographic Behavior of Users Using...”
Ontologies in Web Marketing Services” (Garcia et al., 2011) and “Towards a psychographic user model from mobile phone usage” (Armenta et al., 2011), the purpose of which was to use psychographics to predict consumer behaviour. The second work, presented in the form of the initial analysis of users according to the Big Five personality traits, proved the possibility of automatically collecting information about consumers and segmenting them according to psychographic characteristics. At the same time, the use of research materials for commercial purposes was strictly limited due to the technical difficulties and cost of such data.

Along with this, a strategy for predicting the user profile itself was developed, described in two works: “Demographic and Psychographic Estimation of Twitter Users Using Social Structures” (Ito et al., 2014) and “A psychographic framework for online user identification” (Adeyemi et al., 2014). The techniques presented in these works are based on foreign psychodiagnostics tools; therefore, application in Russia is possible only after their adaptation. Such studies have not been conducted before because most of the tools, as well as social networks, were intended for Western users.

In Russian practice, most studies related to psychographics do not imply its use in the marketing system. The use of psychographic segmentation in Russian marketing is most often aimed at offline campaigns and use in classical communication channels. Regarding the use of psychographic segmentation in the framework of Internet campaigns or daily communication with consumers, a study was released in Russian: “Digital Footprints as a Tool for Profiling Consumers in the Framework of SMART Marketing” (Nevostruyev, 2017). The work is a description of the mechanics of Kozinsky’s (2013) study and designed for the Western market, without examining the possibilities of its application in Russia.

In recent years, there has been a significant shift towards wider use of the Internet, with a diminishing concern for absolute anonymity (Friedman, 2012). The initial source of the new wave of psychographic segmentation is the dissertation “Measurement and prediction of individual and group differences in the digital environment” (Kosinski, 2013). The hypothesis of the work was the creation of psychographic portraits based on behaviour on a social network (Facebook). This became possible after the creation of a model of correlation between the results of user testing and users’ interactions on the social network. Thus, each object in the social network (groups, likes) was assigned its own coefficients. After that, using the new user’s page and collecting the base of their “subscriptions”, it was possible to calculate their psychographic portrait with minimal difference from the passed test. The study hypothesis was confirmed, and the study’s main result was the ability to predict the complete psychographic portrait of a user by analysing their page on a social network. The Facebook toolkit in 2013 also enabled
analysis of the websites visited by the user, which made the forecast more accurate, but then, the sites were replaced by groups on the network. Nevertheless, despite the huge sample and the accuracy of the study, only U.S. residents took part in it, which makes it impossible to use the survey data in Russia.

Such studies have not been conducted in Russian, even though the model proposed in 2013 uses a psychodiagnostics technique, an analogue of which exists in Russian practice. The main reason is the lack of a sufficient number of Russian-speaking users on social networks or the lack of domestic counterparts. Currently, advertising using the Russian social network VKontakte is one of the most popular marketing tools for companies, regardless of their size. Increasing the effectiveness of such advertising is one of the main tasks of specialists, exclusively using socio-demographic segmentation and limiting themselves to creativity in creating advertisements. However, the practice of foreign colleagues proves that with the correct segmentation of users according to psychographic characteristics, a significant increase in the effectiveness of this tool is possible (Bal et al., 2019), and the use of a social network as a database is a commercially justified method (Hundal, 2020).

Currently, the key study in this area is “Psychological targeting as an effective approach to digital mass persuasion” (Kosinski et al., 2017). Building on recent advances in assessing psychological characteristics from digital footprints, this article demonstrates the effectiveness of psychological mass persuasion, that is, adapting persuasive messages to the psychological characteristics of large groups of people to influence their behaviour. On the one hand, this form of psychological mass persuasion can be used to help people make better decisions and lead healthier and happier lives. On the other hand, it can be used to covertly exploit weak points in their character and convince them to make a decision against their own interests. Today, people are faced with compelling communication in a wide variety of contexts: Governments, companies, and political parties use compelling messages to encourage people to eat healthier, buy a particular product, or vote for a particular candidate. Laboratory research shows that such persuasive calls are more effective at influencing behaviour when they are adapted to people’s unique psychological characteristics (Gajanova, 2018). However, investigating large-scale psychological beliefs in the real world has been difficult due to the questionnaire nature of psychological assessments. This study shows that people’s psychological characteristics can be accurately predicted from their digital footprints, such as their Facebook likes or tweets. Convincing calls that matched the level of extroversion or openness to people’s experiences resulted in 40% more clicks and 50% more purchases than their inappropriate or non-personalized counterparts. These results prove that the use of psychological targeting can influence the behaviour of
large groups of people by adapting compelling appeals to the psychological needs of the target audience.

Considering the peculiarities of the research carried out in this area, one of the tasks of this work was the adaptation of a foreign approach to the psychographic segmentation of consumers and the use of this approach in the Russian social network VKontakte for the subsequent adjustment of more accurate targeted advertising.

**Data and Method**

The aim of the work was to develop a user segmentation system capable of providing a sufficient level of participant involvement. To achieve this goal, the following tasks were required:

1. Selection of psychographic testing methods
2. Development of a testing system considering users’ specifics
3. Initial testing

The main prospect of the research was the complete automation of the process because a ready-made system with modern computing power can show the needed result in a few seconds. An automated survey system built into the social network VKontakte would allow for obtaining the necessary sample due to high user involvement in psychological tests and the dissemination of their results.

For such foreign studies, the five-factor personality model is most often used, consisting of five relatively independent traits (Milfont & Sibley, 2012):

- Extraversion
- Benevolence (honesty, respect)
- Conscientiousness (impulsivity/reliability)
- Neuroticism (emotional stability of the user)
- Openness to experience (culture and intelligence)

The only adaptation of IPIP-NEO-120, an in-depth personality test, existing in Russia is under completion and is available exclusively for online testing; commercial use is not available.

Therefore, during the study, we assumed the possibility of self-translation and the use of the English-language questionnaire TIPI (Gosling et al., 2003).

The questions of the questionnaire were translated into Russian by the author of the study. Then, the formulations were adapted considering the traditions of formulating questions in psychological methods in general and tests using the Big Five model in
particular. Additionally, a transcript of the indicator terms was created for the Russian-speaking research participants.

Furthermore, as a subsequent model for developing a web application for users, it was necessary to create a tabular version of the questionnaire as an algorithm.

All personal data (e-mail, page link on a social network) must be hashed to preserve users’ privacy. The survey consists of 10 statements using scaling for the answer (1, 7); responses are then transferred to a point system according to the authors’ method.

When calculating the test results, the primary marks (1.7) in odd questions are replaced by the corresponding points of 7.1. The transferable points are subsequently summed up with the results of the even questions. The final indicator of each of the factors is the division of the sum by 2, according to the research methodology.

Indicators are calculated using a formula where the number is the question number. If the letter R is added to the number, then the number of points is the opposite. That is, if the answer is 7, then 1 point is awarded; if the answer is 6, then 2 points are awarded. Where there is no letter R, the number of points is equal to the answer itself. After adding up, the result is divided by 2.

- Extraversion: \((1 + 6R) / 2\)
- Benevolence: \((2R + 7) / 2\)
- Conscientiousness: \((3 + 8R) / 2\)
- Neuroticism: \((4R + 9) / 2\)
- Openness to experience: \((5 + 10R) / 2\)

Additionally, to increase user engagement in passing testing, it was decided not to use the standard testing option but to use colour coding and elements of a user-friendly interface.

The version of the system created in the form of a table had to be translated into a code that could be used as a VK application. For this, JavaScript was chosen, which allows one to embed applications hosted on a dedicated hosting platform into a social network. This is necessary because the creation and maintenance of a database is required for the full operation of the system.

**Results**

During the study, an application was developed to obtain users’ Big Five psychographic characteristics according to the TIPI questionnaire. At the first launch of the application, the user was required to confirm access to the information posted on their page, according to the social network’s policy.

The start screen of the app talked about TIPI and informed the user about the polling methodology.
When the test was complete, the user was shown a window with the result, as well as a description of the indicators obtained.

One of the key advantages of the constructed system was the complete privacy of the user from the study’s author; that is, the information from the person who completed the questionnaire could be perceived exclusively by the database in the form of cipher text. This made it possible to completely anonymize and not to violate the rules for creating and using VKontakte applications.

According to the rules for using VKontakte, any uploading of data to third-party servers is prohibited and can be punished by blocking, so it was necessary to find an approach to implementing the original algorithm without violating the rules for using the system. At the same time, the main set of tools of a modern Social Media specialist currently consists exclusively of applications and services that violate the rules for using the social network, but the VKontakte administration does not prohibit their functioning because they stimulate the network’s internal advertising market.

Thus, the application itself was modified so that additional functions appeared when opened on behalf of the developer. The buttons “Check the group” and “Check the user” became the final step of the application implementation, confirming the possibility of creating a system for predictive creation of psychographic user portraits on the VKontakte network. “Check the group” was in an input field where the administrator could provide a link to the desired group and see its estimated “metrics” recreated from previous test passes.

The reverse process – the creation of a psychographic user portrait by analysing their behaviour on a social network in Russian – was not performed due to the low prevalence of networks in Russia but because of the technical features of the platform. However, the main difference from the modern approach to collecting such data was access to identical information from the user’s friends, which made it possible to receive data from hundreds of potential consumers due to one passed test. In 2018, this system was revised due to the new privacy policy of social networks, so group subscriptions became the main source of data, as they remained the only points of contact with various brands on social networks.

The hypothesis of the reverse construction of the user’s portrait consists of choosing a random user (for research material) or a group of users (for planning an advertising campaign) with subsequent analytics of their “subscriptions” and calculating the arithmetic mean of the psychographic portrait.

Creating a Psychographic User Portrait Based on Their Digital Footprint
The main goal of the project was to prove the possibility and put into practice the creation of a psychographic user portrait based on their digital trail – in this case by the communities of which the user is a member on the VKontakte network.

The technique of reverse creation of a psychographic user portrait consisted of loading all user groups and calculating the arithmetic mean in accordance with the available data, as well as subsequent questioning using a classical questionnaire to compare the forecast accuracy.

The main step in the development of the system was the implementation of the “Check the user” button. The purpose of the study was to create psychographic portraits on a larger scale, and such functionality is ready, but for clarity, point measurements are used in the given text. As a measuring group, four VKontakte network users were selected, and their test results were voluntarily submitted for subsequent analysis and testing of the predictive system. As shown in the comparisons below, the key metric for system accuracy is the number of occurrences. The number of entries is the number of groups for which information is available in the project base. This confirms the hypothesis of the need to build a constantly updated and changeable system because the measurement accuracy depends on the current state of the group indicators, which could change depending on the content produced.

Thus, the number of occurrences becomes the only necessary factor to improve the system’s efficiency. Because the main audience of users who have used the application live in Nizhny Novgorod, there is a fairly large number of random users whose number of entries will be equal to or more than 50, which already allows using such data as the basis for subsequent decisions. Currently, the accuracy of the system is determined on the basis of 261 people who took part in the initial testing, creating a database of 15,741 lines. This number allows one to take each active Nizhny Novgorod user of the social network VKontakte and drive them through the system, having received their predictive result with a standard deviation within the norm.

At the same time, in the course of analysing the effectiveness of determining reverse typing, shortcomings similar to the results of the previous study were identified. The main disadvantage of the system is its dependence on the sample of respondents. The lack of a mobile version and technical flaws prevented the application from spreading by word of mouth. To subsequently get rid of this shortcoming, it is necessary to refine the technical and design part of the application with the subsequent addition to the VKontakte catalogue. This will allow the test to be offered to users interested in the topic and significantly increase the sample and therefore the accuracy.
Three random communities have been selected that can show the weight of each user who passed the test. On the Figure 1 you can observe the indicators which were obtained with 18 users from the same community who passed the test.

The standard deviation level ranged from 1.13 to 1.98, which, on a scale of 1–7, is a sufficiently high level of deviation and requires adjustment. In this case, for the groups in which the number of occurrences reaches 50 and higher, the standard deviation decreases in several parameters. It should be noted that the factor with the maximum level of standard deviation is neuroticism, regardless of the number of occurrences (Figure 2).

During the research and study of the model, additional indicators were identified that can be built into the system in the future:
- Education
- Languages
- Number of friends
- Number of subscribers
- Number of photos
- Number of posts on the wall

The dependence of the standard deviation on the number of possible entries into the group was also calculated with the construction of a trend line (Figure 3).

To build the chart, a group with 50 entries was used, and the trend line was drawn for 10 possible periods. The graph shows how strongly the size of the database is related to the standard deviation within each specific group. Thus, the main goal of subsequent studies becomes clear: the maximum improvement of the technical side of the system to increase the sample and the accuracy of the work.

**Conclusion**

The purpose of this study was to create a ready-made system for working with psychographic characteristics of users of the social network VKontakte. The feasibility of developing and launching such a system has been confirmed, but during the research, a significant number of new areas of work and potential system improvements have been discovered. The use of the program for commercial purposes is currently possible, considering minor adjustments of the functionality from academic to commercial.

On the other hand, now, and with the available database, the system does not fully meet the expectations set at the beginning of the study. The accuracy of the resulting psychographic user portrait depends on the sample size, and the current version of the system has a significant advantage over the previous algorithm. A closed survey, on
which the user spent 10–15 minutes, as well as data downloaded in table format and requiring a lengthy further processing, prevented the concept from being effectively used for commercial purposes. Now, the application “Psychological test TIPI” allows one, if one has access, to see such data in real time, and the database is updated with each new user who takes part in the test. A possible continuation of the research could be an application that combines the functionality of not only psychographic tools but also sociological surveys. Thus, it is possible to collect more data about each user, which will enable reliance on more significant factors when calculating and building a psychographic portrait.

Thus, the main result of the study was the confirmation of the hypothesis about the possible construction of psychographic user portraits based on their digital traces and the need to improve and supplement the finished system to increase its efficiency.

REFERENCES


ИСПОЛЬЗОВАНИЕ ЦИФРОВЫХ СЛЕДОВ ДЛЯ
СОЗДАНИЯ ПСИХОГРАФИЧЕСКИХ ПОРТРЕТОВ В
ЦЕЛЯХ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ
РЕКЛАМНЫХ СООБЩЕНИЙ

Сергеев П. П.
аспирант программы «Менеджмент» Национального исследовательского университета «Высшая школа экономики»
(Нижний Новгород, Россия)
pp sergeev@hse.ru

Самылина Д. А.
ассистент в Национальном исследовательском университете «Высшая школа экономики»
(Нижний Новгород, Россия)
dsamyлина@hse.ru

Аннотация:
В статье описывается процесс изучения поведения пользователей в цифровом пространстве и создания их психографических портретов с помощью автоматизированной системы, работающей внутри социальной сети ВКонтакте (ВКонтакте,) а также благодаря сбору информации об этих пользователях через стандартную компьютерную программу. В рамках данного исследования авторы попытались доказать возможность применения собранных данных для решения задач малого и среднего бизнеса, а также предложить использование психографической сегментации для создания более персонализированной коммуникации бренда со своими потребителями. Исследование основано на анализе данных, полученных через приложение в социальной сети ВКонтакте (TIPI Psychological Test), которое было разработано и реализовано авторами в виде онлайн-опроса 261 респондента из Нижнего Новгорода для выявления их психографического профиля. Также авторы опирались на результаты анализа базы данных (MySQL) активности этих людей в социальной сети ВКонтакте (их подписки на группы и сообщества в сети).
На основе сравнения результатов использования этих двух методов, авторы приходят к выводу, что возможно разработка и внедрение системы, использующей цифровые следы пользователей для создания их психографических портретов и настройки более эффективной таргетированной рекламы на заданную целевую аудиторию.

Ключевые слова: цифровой портрет, цифровое поведение, реклама, эффективность, технологии, маркетинг, сеть