

Increasing the information content of social network groups and clients using Social Mining

Valery Kaziev
*Candidate of Physics and
 Mathematics, Associate Professor
 Kabardino-Balkarian State
 University
 Nalchik, Russia
 kkvkvm@yandex.ru*

Bella Kazieva
*Candidate of Economic Sciences,
 Associate Professor
 Kabardino-Balkarian State
 University
 Nalchik, Russia*

Faiz Khizbullin
*Doctor of Technical Sciences,
 Professor of the Department of
 Management and Service in
 Technical Systems
 Ufa State Petroleum Technical
 University
 Ufa, Russia*

Oksana Takhumova
*Candidate of Economics Sciences,
 Associate Professor of the
 Department of Economic Analysis
 Kuban State Agrarian University
 Krasnodar, Russia*

Abstract–Connections, contacts, and commerce on social networks are expanding dynamically and have become attractive to many visitors who enjoy, for example, predicting events in a group, finding a person, finding out necessary information about him. Intelligent systems for activating the client, the client base, increasing information content from them began to be actively studied, taking into account the processes occurring in social networks. New forms of feedback, mechanisms (regulators) are activated; their new systemic and synergistic effects are being investigated.

The problems of the traditional and the tasks of modern (media network) sociology, social network technology, the methodology of research on society are investigated in the work on the basis of the principles of system dynamics. Hierarchical client structures in which the client, the cluster has its own weight, rank, are considered. The important task of identifying the rank and client, for example, the initiator (coordinator) of network processes, is investigated. A graph model of such hierarchical structures, taking into account hierarchical subordination, and the measures of connectivity necessary in assessing the evolutionary potential of social network groups, is proposed. The procedure for assessing the potential of a network (group) is given.

The results are the application of Social Mining in practice.

Keywords–social, network, customer base, informative, Social Mining

I. INTRODUCTION

The social circle in social networks expands dynamically and is attractive to many. For example, there is not only an exchange, viewing photo-video, but also targeted manipulation. Many enjoy finding links and predicting events in groups. Social networks, social networking sites allow you to find a person, perform geolocation, setting the necessary characteristics. For example, such a search may be useful to recruiters.

The systems of customer activation, increasing the information content of the client base, for example, a travel company [1] or a bank [2], have been actively used. The problems of ethnic relations, from the point of view of systemic relations and processes, in particular interethnic, are investigated [3]. The specific processes taking place in social networks, where new aspects of feedback, regulators, effects are possible, add their effect. The principles of system dynamics work [4], in particular:

- 1) the existence of internal regulatory system mechanisms activated by feedbacks;
- 2) the proportionality of influences (intrasystemic, with the environment) with the dynamic state of the system, etc.

Sociology (in the traditional sense) degenerates into an intellectual discourse, which is far from media education, social network processes [5], does not take into account modern technologies, methods of research of society, in particular, multi-agent [6].

II. METHODOLOGY

The empiricism of social, social, and media processes shows that social processes must also be investigated in the “human dimension” (that is, in terms of human values). Often used as a non-systemic process, the generator of random (more precisely, pseudorandom) numbers is a system number, that is, it has a goal, structure (algorithm), relations (data), resources and emergent property [7].

Social networks can lead to underestimation-overstatement of self-esteem. American researchers conducted an experiment by asking tests on the nature of lightning in two different groups of students. Students of the first group were given printouts, the other group was only allowed to use the Internet. Students of the second group decided that they possess all the knowledge about the nature of lightning (“I used

the Internet!”). The side of the important categories of “know real” and “know virtually” is erased by a web search for information. This leads to negative consequences.

Systems sociology, interacting with systems and networks of Social Mining, Big Data, Neuro Computing, Computational Social Science, e-Social Science, etc., builds the infrastructure base of the modern network community. Systemic and synergetic thinking, systemic and adaptive methods (technologies) are activated [8]. Recently, computer and media technologies, in particular, Data Mining and Social Mining, have been effectively used for mining. This is necessary when the number of users in social networks grows (for example, VKontakte dynamics are shown in Figure 1).

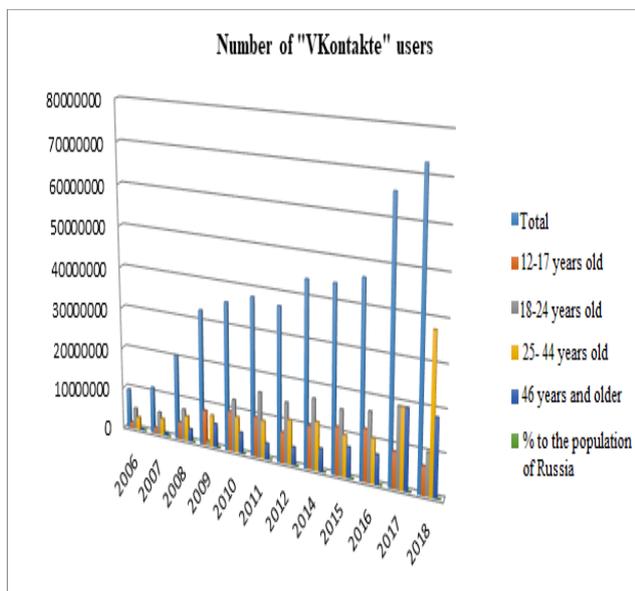


Fig.1. Dynamics of “VKontakte”-users by years and ages

The results obtained using Data Mining and Social Mining may cast doubt on some of the provisions of traditional sociology (the effect of the “sociological myth”).

A social network is understood as a virtual network user infrastructure, an environment of objects and processes, as in peer-to-peer networks connected by common relations. It is often identified with a cognitive graph structure: nodes are actors (groups, individuals, companies, etc.), arcs are their infological relationships (information flows - socio-economic, family, corporate, etc.).

Now philosophers and sociologists understand this category more broadly as a special philosophy of analysis, integration and updating of systemic network resources of a society, a group [9]. Social Mining is an adaptive application of Data Mining tools in search, identification, analysis of dependencies in social networks [10].

Social Mining, being an analogue of Data Mining (data analysis), searches for connections that are not

“on the surface”, dependencies in social networks. The whole history of the development of social networks indicates: the necessary data on the social life of a potential buyer is indicated in his profile for communication in a social network. Registration forms make up all the necessary information about a person for an initial acquaintance.

Functionality, applications used in Social Mining provide the accumulation, processing of the necessary information, get the possibility of probabilistic and situational modeling of characteristics, dependence on social networks, behavior factors, user relationships. Based on past knowledge and current situation, Social Mining will be able to analyze and produce the correct (i.e. correct, stable) result.

Social Mining can help identify leaders of social networks, PR events, objects that communicate more actively in groups that have many diverse connections, generate traffic, fall into this group (network segment), etc. Such tasks are relevant in marketing, advertising, security, gestalt psychology, customer retention, etc.

Data Mining is also used to analyze data on consumer behavior. Marketers send offers through call centers, portals. Using web and business intelligence is important. It is necessary for a comprehensive study and analysis of business evolution, both systemic and narrow-profile, using all available resources and criteria for returning invested funds with a profit.

Examples of solutions using Social Mining important tasks are given below.

Example 1. The formation and maintenance of the customer base of the bank. A well-formed client base that takes into account credit histories is the basis for sustainable bank lending in the context of the evolution of credit instruments and banking services. Especially, commercial banks that respond quickly to the behavior and capabilities of the client base, as well as a set of banking services.

According to researchers (for example, [2]), it is advisable to consider the client base only as a combination of classes to which certain algorithms and taxonomy criteria are applied, in particular, segmentation into real or potential clients [11], the knowledge of the “holding ability” of the first or “Attraction” of the second. Segmentation will require heterogeneous, varying depth and breadth information. However, only when using Social Mining is it possible to develop a client base that will affect the bank’s income, its predictability, if the necessary completeness and depth of processing of Social Mining information, a relevant interpretation of the results is ensured.

For example, identifications of indices like [12]:

- 1) NPS (Net Promoter Score) - measures of recommendation by the client to others;
- 2) CSAT (Customer Satisfaction Score) - measures of customer satisfaction;

3) CES (Customer Effort Score) - measures of subjectivity of the client's assessment of efforts to resolve their issue.

In addition to the main source of customer information - the Bureau of credit histories, as well as annoying calls, banks should use social networking opportunities, Social Mining (within the law). Only 20% of poorly structured data is analyzed in Russian banking practice [13].

Example 2. Insider identification. Let us consider in detail the problem of insider information leaks. Using Social Mining, the person with whom partners and counterparties most often communicate can be identified. This will allow the company's security service to reduce the risk of information leakage, provide targeted audits and monitoring of the most vulnerable points of the system. The most important tasks in this case are the identification of employees who are very actively communicating with employees of a competitor or partner.

Personalized access to data is achieved by searching, collecting information about the use of social networks, segmentation and interpretation, the so-called profiles of groups, segments. These data make it possible to predict the choice of people. The network also contains users who are highly active, or their behavior is significantly different from the "profiled" one.

The simplest example is spam (or viral marketing), often containing links to viruses, phishing, etc. Such "activists" must be found and neutralized on time. Organizations using Social Mining have the ability to prevent fraud.

III.RESULTS

Social networks, resources in them are structured according to a hierarchical, network or matrix type (we do not consider a linear structure as trivial). Hierarchical structures in which each user, each group has a certain weight, rank in the hierarchy, while the task of identifying this rank is important, are discussed below. For example, it is important to identify the initiator of network processes.

Not only social, but also financial, psychological, and other factors are important here, for example, stratification in social networks, as well as in life [14].

In some professional social networks, the factor that forms the hierarchical structure (levels, depth, width, tightness of connections or ranks) is competence. In other social networks, the opposite trend can be observed here: the higher the professional level, the greater the degree of rejection of opinion, the lower the status in social networks. The social energy of the individual is important, not its functionality. Marginal, adventurous persons with deviation behavior can gain social network capital.

In such situations, the task of recognizing the level, status of the user, his social goals has become urgent. It is necessary to involve mathematical, infological

methods, tools for analysis, verification, formalization of situations for this.

Hierarchical structures are usually described by the apparatus of graph theory. To simulate the social network structure, their connections, hierarchical subordination, with ranks, measures of connectivity, a mathematical (formal) graph of the form is defined: $G = \langle V, R \rangle$, где $V = \{v_1, v_2, \dots, v_n\}$ is the set of vertices corresponding to the elements of the network organization, $R = \{r_{ij} = (v_i, v_j): i, j \in \{1, 2, \dots, n\}\}$ is the set of arcs corresponding to the connections of elements, creative, group, commercial (SMM), image, security, etc.

Let the sequence of arcs in the graph G (channel of influence) from vertex i to vertex j :

$$(v_i, v_k) \rightarrow (v_k, v_l) \rightarrow \dots \rightarrow (v_p, v_q) \rightarrow (v_q, v_j).$$

If the vertices v_i and v_j are connected by a hierarchy relation ($v_i > v_j$), then there is a zero-length path from v_i and v_j in which v_j is a "slave" (subordinate) - directly (legitimately, consciously) or indirectly (unconsciously, "blindly") . Graph G has no loops; it does not contain loops (a path in which vertices do not repeat, and the initial vertex coincides with the final one).

In the social network structure G , we single out the interests leading to the grouping (taxonomy) $S = \langle L_1, L_2, \dots, L_m \rangle$, where the set $L_j, j = 1, 2, \dots, m$ corresponds to the interest $i \neq j, L_i \neq \emptyset$, if the S - group is unique. The subset $L_i \subset S$ is the set of members of social networks with the same rank (as in peer-to-peer networks), independently with the internal structure of the network or formally (i.e. model) - the rank is assigned.

Grouping S is ordered if

$$\forall p, q \in \{1, 2, \dots, m\}, \forall i, j: \exists (v_i, v_j) \in R, p < q.$$

This means that S arcs $r_{ij} = (v_i, v_j)$ in an ordered group are always directed from a subset with a lower number to a set with a higher number, for example, there are only vertices without ancestors in the first L_1 , and in the last L_m without descendants. For example, the initiator, moderator of the group belongs to a higher level, the social layer. Such ordering is possible, for example, using fuzzy mathematics [15].

Various measures of hierarchical rank can be specified. For example, you can determine using a measure of complexity (structural):

$$\mu = \max \mu_i(L_p),$$

$$\mu_i(L_p) = \sum_{j=1}^m (j - p) N_j^i,$$

where N_j^i - the number of vertices in the subset L_j reachable from the vertices i .

The hierarchical rank of a social network user is a multifaceted characteristic, each face defines one side

of the organizational social structure, for example, the number of followers or the distance (on the graph) to them.

When assessing the evolutionary potential of social network connections (groups), the significance and interconnection of the considered indicators is investigated and its integral indicator is determined. For example, given the complexity of the constituent elements and relationships, you can determine the indicator by the formula:

$$s = \sum_{i=1}^n s_i k_i,$$

Where i – number of network subsystems; s_i – complexity of a subsystem of i -th type; k_i – число элементов i -го типа системы the number of elements of the i -th system type.

The potential assessment procedure is implemented in stages:

- 1) basic (potential-forming) factors are determined;
- 2) social network indicators are identified by factors (communication coefficients, updates, etc.) and data collection;
- 3) assessment of significance, ranking of factors, selection of significant and their deviations;
- 4) algorithms for assessing potential (calculation of parameters, indicators, levels, scores, points).

IV. DISCUSSION

Within social networks, groups, communities (blogs, chats, forums, etc.), connections (thematic, regional, interests, etc.), goals, resources, hierarchies are formed.

Personalization of the buyer, visitor, his needs, interests, location allows targeted marketing. Modern systems of intellectual and neuro-support (for example, neuromarketing) track only those products and services that the client is interested in. Everyone will find the data, according to sociological and expert estimates on YouTube (63% of all users), VKontakte (61%), Odnoklassniki (42%), FaceBook (35%), Instagram (31%) and others try to personalize the client (data from Statista agency)

All the most popular, “trending” in the world in 2018 social networks gave the following representation by “volume” (in million people): Facebook (2061), YouTube (1500), WhatsApp (1300), WeChat (963), QQ (850).

In the "network life, client:

- 1) fills out a profile (data, photos, notes, friends, etc.);
- 2) joins groups using social networks on the Internet and subscriptions;
- 3) makes appointments;

- 4) read news, comment;
- 5) just puts “likes”.

Everything serves to collect information about a person. The following factors are taken into account: where did he get on the network, did he use Wi-Fi or mobile Internet, where did he take a photo, where did he get to the network. It seems to be insignificant connections, but they are analyzed deeply, at levels lower and lower. As a result, each user has an array of data that has everything you need. Therefore, social networks in our lives can tell everything about the user: actions, habits, intentions, needs, many of which the user would not tell in a confidential conversation with a person.

To prevent terrorist attacks, fraud, networks track people who can do harm. For example, the NSA uses programs that can monitor and identify those that fit certain patterns of behavior on social networks. This is done to predict a criminal conspiracy, an event. Neuro-mathematicians, computer scientists, psychologists, linguists did a good job here; suitable tools - cognitive schemes, data analysis, etc. are used.

Social Mining-based analytics help evaluate and process collected corporate information, increasing business activity, market stability; at the maximum adaptability of management, tracking of business processes, tracking, for example, using the Process Mining program.

It is important to monitor business processes in real time, analyzing their bottlenecks, effectively and intelligently monitoring. A successful business has a reasoned adaptive strategy and relevant tools.

It is necessary to take into account social network relations, the intellectual capital of user groups (clusters), visitor satisfaction, the socio-economic income of groups, the entire target audience, its level of culture. These indices confirm the evolution of network structures; businesses are better than economic indicators.

Our results will help to systematically and formally solve the problems of data enrichment using this toolkit and Social Mining.

V. CONCLUSION

To summarize, we note that the use of Social Mining technologies will become an additional engine for the customer-oriented approach of the company. This is a good opportunity to also remind you of a slightly forgotten brand. Intelligent social network analytics is a powerful management and security toolkit.

The proposed approach to the description of hierarchical structures and methods for setting complexity measures of hierarchical social structures is the basis for identifying the user's social network status. The criteria for the quality and quantity of customer information for the evolution of the customer base should be relevant to reality, relevant, ranked by

importance (objectivity, impact), efficiency, and effectiveness.

ACKNOWLEDGMENTS

The article is published with partial support from the Russian Federal Property Fund, project No. 17-02-00467-OGN “Improving organizational and information forms of managing innovative investment processes in the context of cluster diversification and transformation of problematic regional economies”.

REFERENCES

- [1] A. A. Limpinskaya, “Low-budget methods of forming a travel agency client base”, in *Resorts. Service. Tourism*, 2018, No 2 (39), pp. 14-22.
- [2] R. G. Imambaeva, “Ways of development of the bank's client base in modern conditions of global financial integration”, in *Global markets and financial engineering*, 2014, No 1 (1), pp. 63-72.
- [3] A. Y. Petukhov, A. O. Malkhanov, V. M. Sandalov, and Y. V. Petukhov, “Modeling Conflict in a Social System Using Diffusion Equations”, in *Simulation: Transactions of the Society for Modeling and Simulation International*, 2018, Vol. 94 (12), pp. 1053-1061.
- [4] A. A. Davydov, *Systems Sociology: Social Networks Mining*. Moscow: IS RAS, 2009, 228 p.
- [5] G. Abend, “Thick Concepts and Sociological Research”, in *Sociological Theory*, 2019, Vol. 37, Is. 3, pp. 209-233.
- [6] R. Sawyer, “Artificial Societies: Multi agent systems and the micro-macro link in sociological theory”, in *Sociological Methods and Research*, 2003, Vol. 31, No 3, pp. 325-363.
- [7] L. V. Glukhova, V. M. Kaziev, B. V. Kazieva, “System rules of financial control and management of innovative business processes of an enterprise”, in *Bulletin of the Volga University named after V.N. Tatishchev*, 2018, Vol. 2, No 1, pp. 118-126.
- [8] *Big Data in Computational Social Sciences and Humanities*. In book: *Big Data in Computational Social Science and Humanities*, 2018, pp. 1-25.
- [9] D. Kim, N. Allum, and A. Denman, *Who Is Doing Computational Social Science? Trends in Big Data Research*. Available at: <https://us.sagepub.com/sites/default/files/compsocsci.pdf>.
- [10] P. Attewell, and D. Monaghan, *Data Mining for the Social Sciences*. University of California Press; First edition May, 2015, pp. 264.
- [11] D. N. Vladislavlev, “Customer service management system as a competitive advantage”, in *Creative Economy*, 2011, No 11 (59), pp. 124-130.
- [12] E. De Haan, P. C. Verhoef, and T. Wiesel, “The Predictive Ability of Different Customer Feedback Metrics for Retention”, in *International Journal of Research in Marketing*, 2015, No 32 (2), pp. 195–206.
- [13] *The level of customer service in Russian banks: KPMG analysis*. Available at: <https://assets.kpmg.com/content/dam/kpmg/pdf/2015/02/Banking-survey2015-fin.pdf>.
- [14] I. V. Mostovaya, and G. A. Ugolnitsky, “Social pyramids: mathematical and sociological analysis”, in *Scientific Thought of the Caucasus*, 1998, No 3, pp. 37-43.
- [15] M. G. Matveev, and N. A. Aleinikova, “Mathematical modeling of a network planning problem using fuzzy mathematics”, in *Bulletin of the Voronezh State University (Series "System Analysis and Information Technologies")*, 2018, No 3, pp. 155-162.