Adaptive method of color selection in application to social media images

Fedor Bushmelev[a], Maksim Abramov[a,b] and Tatiana Tulupyeva[a,b,c]

[a] *Saint-Petersburg State University, 7/9 Universitetskaya Emb.; Saint-Petersburg, 199034, Russia*
[b] *Saint-Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, 14-th Linia VO, 39; Saint-Petersburg, 199178, Russian Federation*
[c] *North-West Institute of Management of RANEPA, Sredny prospect VO, 57/43; Saint-Petersburg, 199178, Russian Federation*

Abstract

The paper presents an approach that allows one to obtain a numerical understanding of the emotional state of a social media user based on images published by him. The proposed algorithm is to apply an adapted color selections method to images. The work can be useful for determining and refining the assessments of the security of the user of the information system against social engineering attacks and their further research.

Keywords 1

Social engineering attacks, information security, image processing, psychological user's characteristics, social media, fuzzy sets of color selections

1. Introduction

The rapid pace of development of information technologies, as well as high human involvement in them, make people think about information security more often. Recently, there has been a significant increase in the number of attacks on information systems [1, 2]. Statistics provided in the annual report of the telecommunications company Verizon, as well as a number of other companies in the field of information security [3, 4, 5], show that a significant part of attacks on information systems is completed by social engineering. Thus, it is confirmed that the user is one of the most important and vulnerable parts of the information system, which directly affects the level of its security. Methods thats are used in social engineering are usually manipulative, psychological and analytical techniques for the hidden motivation of the user [6]. During preparation of an attack, the attacker relies on the psychological, social, culture-anthropological characteristics and characteristics of the user obtained from various sources [7, 8]. To assess the security level and countermeasures of social engineers, it is necessary to have an idea of the specified user characteristics. In this work, the main source of information about them is social networks themselves, meaning publications on the pages, containing large amount of information about user [9]. You can get the degree of expression of the user's personal functions by analyzing the posts, texts, graphic material, friends list, subscriptions to various communities, personal information provided in the questionnaire. It is worth noting that there are already developments in building a user's metaprofile on several pages of social networks [10], as well as in forming an idea of the user's psychological characteristics based on information extracted from social networks [11] ... When constructing assessments of the security of information systems against social engineering attacks, the profile of its vulnerabilities is used, which is based, among other things, on the psychological characteristics of the user [12]. Thus, on the basis of information about the user, extracted from various sources, we can talk about the severity of his personal functions, as a result, vulnerabilities and assessment of protection against social engineering attacks.

Russian Advances in Fuzzy Systems and Soft Computing: selected contributions to the 8-th International Conference on Fuzzy Systems, Soft Computing and Intelligent Technologies (FSSCIT-2020), June 29 – July 1, 2020, Smolensk, Russia
EMAIL: fedebushe@gmail.com; mva@dscs.pro; tvt@dscs.pro
ORCID: 0000-0003-2465-6859; 0000-0002-5476-3025; 0000-0003-3630-7971

Several social media are currently the most popular among others [13]. These are “Youtube”, “Facebook”, “Instagram”, “Twitter”, “Reddit”, in particular, in Russia and the CIS countries are also popular “VK.com”, “OK.ru”. There is an active growth in the audience of named services, at the same time, the number of graphic content published by users is growing. That is why it seems necessary to adapt and use the latter to build the user's psychological functions, namely, to assess his emotional state. Note that currently there are no methods, approaches and tools for determining the psycho-emotional state of the user. Such developments of aggregates would allow to assess the severity of their vulnerabilities. This article uses methods for processing images published on the network, using numerical, ordinal indicators in the current psychological state of the user.

2. Relevant works

This work is part of a general research devoted to the problem of analyzing the security of users of information systems from social engineering attacks, carried out on the basis of the SPIIRAS laboratory of theoretical and interdisciplinary problems of computer science. The researchers of the laboratory proposed a model of a socio-engineering attack and developed a prototype of a software package that simulates socio-engineering attacks on users of an information system [6]. As part of the research, one of the common tasks is to build a profile of user vulnerabilities. To solve this problem, work was carried out to aggregate data about the user from social networks [11, 14, 6, 12, 10, 15, 16]. They solve the problem of identifying the severity of the psychological characteristics of the user based on the analysis of the content published by him on his page in the social network. However, even the estimate obtained on the basis of the above works may be inaccurate due to the incompleteness of the data used. It is assumed that there is the use of graphic information from social networks as a source of information about the psycho-emotional state of the user. Testing this hypothesis will contribute to the development of the "personnel" model, the components of which are the psychological characteristics and the profile of the user's vulnerabilities, and, as a consequence, will help improve the security assessment.

As a theoretical basis for identifying the characteristics of the user's emotional state, the method of color selections will be used [17] - a psychological test, a modification of the Luscher color test, related to projective methods. The method uses eight colors and each color is assigned a number: *dark blue* (1); *dark green* (2); *light red* (3); *yellow* (4); *purple* (5); *brown* (6); *black* (7); *indifferent gray* (0). The testing procedure consists in ordering the colors of the subjects according to the degree of their subjective pleasantness. Color (combination of colors) at a certain position shows aspects of the internal psycho-emotional state of the tested person.

Also note that initially this method cannot be applied to images, especially those published on social networks. Within the framework of this work, a groundwork is being formed to test the hypothesis about the possibility of obtaining an assessment of the user's emotional state, in particular, the applicability of the Luscher test to quantitative indicators isolated from a digital image using the adapted method of color selections.

3. Formulation of the problem

This research is devoted to solving the problem of automated application of the adapted color selection method to images posted by a user on social networks. In other words, it is necessary to take an image as input, and as a result of the implementation of the algorithm, obtain the numerical values of the emotional state assessment. To obtain this assessment, it is necessary to solve several subtasks, consisting in the development of an approach to color segmentation, taking into account the peculiarities of the application of the color selection method; as well as the construction of an algorithm for applying the adapted color selection method to images.

4. Applying an adapted color selections method

In this section, we will consider obtaining the numerical characteristics of the user's emotional state using image processing, which will be performed based on the method of color selections. In view of the fact that the latter is not directly applicable to the graphic content of social networks and is intended to work in real conditions, we will propose an approach for its adaptation to digital images, for the possibility of further work on the formation and verification of a hypothesis about psychological characteristics.

The algorithm of the proposed approach is shown in Figure 1. Let's consider each stage in more detail, starting with the 2nd «*Conversion to HSV model*».

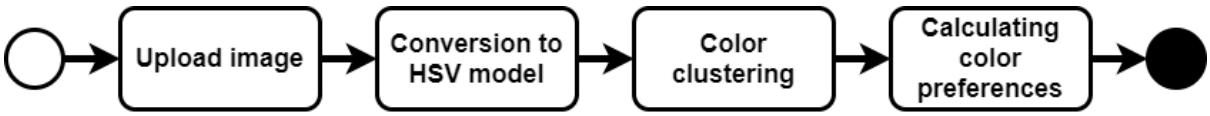

Figure 1: An algorithm for obtaining emotional characteristics from an image using the color selection method

4.1. Second level heading

In order to work with images automatically, you must choose the most appropriate color model. Additive (eg RGB) and subtractive (eg CMYK) models have a limited color space – all the possible color values given by the model - and are different on different devices. The L*a*b* model and those associated with it are devoid of this drawback, but are obtained by methods using cube roots and are highly nonlinear, which complicates the computation process. The HSV model was obtained by nonlinear RGB conversion, is componentally closer to human perception of colors and can be a worthy replacement for L*a*b*, which is confirmed in various studies related to computer vision, comparison and applicability of color models, graphics software [18, 19]. Thus, the HSV model was chosen to work with images.

4.2. Adapting the color selections method

In order to talk about the user's color preferences, based on the images published by him, it is necessary to understand the quantitative characteristics of colors. The color selections method works with 8 primary colors. They can be divided into monochrome (*gray*, *black*) and chromatic colors (the other 6 colors). In this case, the selected color model is applicable to this division: the *saturation* and *value* components will be responsible for the first group of colors, and *hue* for the second, respectively. Note that using standard metrics directly, such as Euclidean distance, is inappropriate because gives an incorrect result.

4.2.1. Chromatic colors

We will define them using the *hue* component, which is a color gradient circle that varies from 0 to 360°. Each chromatic color in the color selections method can be associated with an area of acceptable *hue* component of tones, about which it can be said unambiguously that they belong to one or another color.

For shades in the intervals between the boundaries of color areas, the belonging must be determined separately. Different approaches and metrics for matching hue to color can be applied, for example, in one image a group of pixels of one hue surrounded by contrasting hues may visually differ from the same group of pixels in another image surrounded by similar hues, affecting the perception and the final result. Within the framework of this work, we will restrict ourselves to the augmented Euclidean distance, taking into account the values of the *saturation* and *value* components, which vary from 0 to 100, in view of their greater influence on this area. Consider the spacing

between the boundaries of two colors C_1 and C_2. We will say that the hue $C_1 < C_2$ by the value of the *hue* component. Let's set the distance as follows:

$$\begin{cases} d_s^2(c,p) = \sqrt{\frac{(100-s_c-s_p)^2+d(c,p)}{2}} \\ d_v^1(c,p) = \sqrt{\frac{(100-v_c-v_p)^2+d(c,p)}{2}} \\ d(c,p) = \sqrt{(h_c-h_p)^2+(s_c-s_p)^2+(v_c-v_p)^2} \end{cases} \quad (1)$$

where, $d_s^2(c,p)$and $d_v^1(c,p)$ – distances to C_1 and C_2; c and p – one of two colors and a pixel; h_c, h_p, s_c, s_p, v_c, v_p – pairs of hue, saturation, and values for color and pixel, respectively.

Then, using equations (1) for each of the colors, we calculate the distance to a pixel belonging to this interval and select the minimum distance. In other words, if the shade of a pixel is dark enough, then it will be primarily referred to the color C_1, and if it is faded, then to the color C_2.

4.2.2. Monochrome colors

Figure 2 shows that gray and black colors are present in each field of hue shades, therefore, to determine them, we will use the *saturation* and *value* components. Gray is a weakly saturated color and black is a low brightness. Thus, you can set upper bounds b_s and b_v for saturation and brightness values or gray and black, respectively.

4.2.3. Color tone segmentation

Due to the fact that the values of the model components are continuous, a situation may arise when it is impossible to unambiguously determine which of the shades of tone a pixel belongs to, chromatic or monochrome. In this regard, it is proposed to introduce tone segmentation, as shown in Fig. 2.

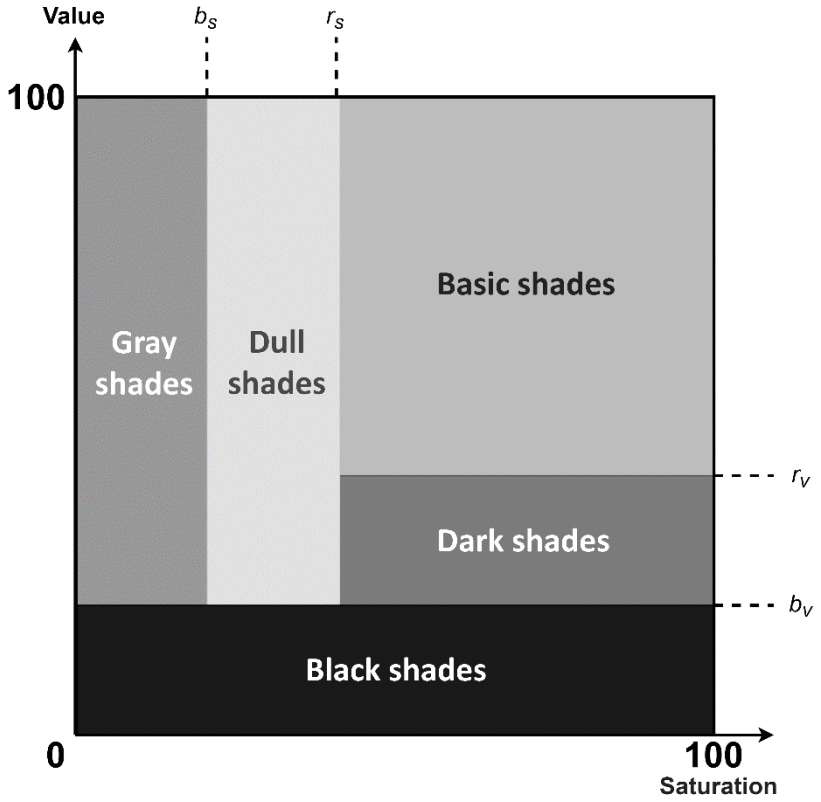

Figure 2: Segmentation of color shades by *saturation* and *value*

Where r_s and r_v denote the lower bounds for saturation and brightness, when we can tell what type of hue a pixel is. These boundaries, as well as b_s and b_v, can be either fixed or obtained, for example, using statistical methods. In this work, we will focus on the fact that the upper boundaries of monochrome shades are fixed at 15 units, and the lower boundaries of chromatic shades as the 25th percentile. This choice is justified by the fact that, depending on the image, the ratio of illuminated and dark pixels can be drastically different, affecting perception.

To work with pixels, the values of which lie in the areas of dim and dark shades, we will use the Euclidean distance formulas with some addition. For the area with dull shades, equations (2) are used.

$$\begin{cases} d_g(c,p) = \sqrt{a_g \left(\frac{s_c - s_p}{b_s}\right)^2 + (1 - a_g) \cdot d(c,p)^2} \\ a_g = \max\left(1 - \frac{\min(s_c, s_p)}{r_s}, 0\right) \end{cases} \quad (2)$$

For the area with dark shades, we apply equations (3).

$$\begin{cases} d_b(c,p) = \sqrt{a_b \left(\frac{v_c - v_p}{b_v}\right)^2 + (1 - a_b) \cdot d(c,p)^2} \\ a_b = \max\left(1 - \frac{\min(v_c, v_p)}{r_v}, 0\right) \end{cases} \quad (3)$$

Thus, the numerical characteristics of the colors of the image can be obtained. These results will be transmitted to specialists in the subject area to determine the emotional state of the user, his psychological characteristics and subsequent research.

5. Conclusion

The article proposes an approach to identify the internal psycho-emotional state of the user based on images published on social networks and media. The approach consists in the automated application of an adapted color selection method to designated images on social networks and obtaining the corresponding numerical characteristics. The proposed methodology lays the foundation for further work on conducting appropriate experiments, clarifying existing and developing new approaches, models and algorithms, forming and testing hypotheses. The results obtained will contribute to the construction of more accurate ideas about the psychological characteristics of the user of the information system, the profile of his vulnerabilities and, as a consequence, the assessment of protection against social engineering attacks.

6. Acknowledgements

This work was carried out under the state assignment SPIIRAN No. 0073-2019-0003 and with the financial support of the Russian Foundation for Basic Research (grants No. 18-01-00626, No. 20-07-00839).

7. References

[1] V. Desnitsky, I. Kotenko, I. Parashchuk, Neural Network Based Classification of Attacks on Wireless Sensor Networks, in: Proceedings of 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, EIConRus 2020, IEEEXplore. St. Petersburg and Moscow, 2020. doi: 10.1109/EIConRus49466.2020.9039275.

[2] Verizon.com, Verizon Data Breach Investigations Report, 2019. URL: https://enterprise.verizon.com/resources/reports/dbir/2019/introduction.

[3] Kaspersky.ru, 7 ways to cyberattack vulnerability, 2020. URL: https://www.kaspersky.ru/resource-center/preemptive-safety/7-ways-to-cyberattack-vulnerability.

[4] Norton.com, 2019 data breaches: 4 billion records breached so far, 2019. URL: https://us.norton.com/internetsecurity-emerging-threats-2019-data-breaches.html.
[5] PurpleSec.us, The Ultimate List Of Cyber Security Statistics For 2019, 2019. URL: https://purplesec.us/resources/cyber-security-statistics.
[6] M. V. Abramov, T. V. Tulupyeva, A. L. Tulupyev, Social engineering attacks: social networks and user security assessments, SUAI, Saint-Petersburg, 2018.
[7] K. Mitnick, W. Simon, The Art of Deception, ITco., Moscow, 2004.
[8] K. Zetter, Teen Who Hacked CIA Director's Email Tells How He Did It, 2015. URL: https://www.wired.com/2015/10/hacker-who-broke-into-cia-director-john-brennan-email-tells-how-he-did-it.
[9] J. Du, C. Jiang, K. C. Chen, Y. Ren, H. V. Poor, Community structured evolutionary game for privacy protection in social networks, IEEE Transactions on Information Forensics and Security 13, 3 (2018) 574 - 589. doi: 10.1109/TIFS.2017.2758756.
[10] A. A. Korepanova, V. D. Oliseenko, M. V. Abramov, A. L. Tulupyev, Application of Machine Learning Methods in the Task of Identifying User Accounts in Two Social Networks, Computer tools in education, 3 (2019) 29-43. doi:10.32603/2071-2340-2019-3-29-43.
[11] M. V. Abramov, A. L. Tulupyev, T. V. Tulupyeva, Aggregating data from social networks to restore a fragment of a user's metaprofile, in: Proceedings of the Sixteenth Russian Conference on Artificial Intelligence, RCAI-2018, Moscow, 2018, 189-197.
[12] G. I. Bagretsov, N. A. Shindarev, M. V. Abramov, T. V. Tulupyeva, Approaches to development of models for text analysis of information in social network profiles in order to evaluate user's vulnerabilities profile, in: Proceedings of the XXI International Conference on Soft Computing and Measurement, SCM'2017, Saint-Petersburg, 1 (2017) 134-137.
[13] WeAreSocial.com, Digital in 2020, 2019. URL: https://wearesocial.com/digital-2020.
[14] M. V. Abramov, A. L. Tulupyev, A. A. Suleymanov. Tasks of analysis of user protection against social engineering attacks: building a social graph on information from social media, Scientific and Technical Journal of Information Technologies, Mechanics and Optics, issue 18, 2 (2018), 313-321. doi: 10.17586/2226-1494-2018-18-2-313-321.
[15] T. V. Tulupyeva, A. L. Tulupyev, A. Y. Pashchenko, A. A. Azarov, M. V. Stepashkin, Socio-psychological factors affecting the degree of vulnerability of users of automated information systems from the point of view of social engineering attacks, SPIIRAS Proceedings, 12 (2010) 200-214.
[16] N. Shindarev, G. Bagretsov, M. Abramov, T. Tulupyeva, A. Suvorova, Approach to identifying of employees profiles in websites of social networks aimed to analyze social engineering vulnerabilities, in: Proceedings of the 2nd International Scientific Conference on Intelligent Information Technologies for Industry, IITI 2017, 2017, 441-447. doi: 10.1007/978-3-319-68321-8_45.
[17] L. N. Sobchick, Method of color selection – modification of the eight-color Luscher test: Workshop on psychodiagnostics, Retch, Saint-Petersburg, 2007.
[18] D. Bora, A. Gupta, F. Khan, Comparing the Performance of L*A*B* and HSV Color Spaces with Respect to Color Image Segmentation, International Journal of Emerging Technology and Advanced Engineering, 5 (2015). URL: https://arxiv.org/abs/1506.01472.
[19] J. Lv, J. Fang, A Color Distance Model Based on Visual Recognition, Mathematical Problems in Engineering, 2018. doi:10.1155/2018/4652526.