

PAPER • OPEN ACCESS

## Automation of network monitoring: methodology of destructive risk assessment

To cite this article: Y Goncharov *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **537** 052022

View the [article online](#) for updates and enhancements.



**IOP | ebooks™**

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

# Automation of network monitoring: methodology of destructive risk assessment

Y Goncharov, V Zarubin, A Kalashnikov, N Tolstykh and A M Nuzhnyy

Voronezh State Technical University, Moscow ave., 14, Voronezh, 394026, Russia

E-mail: cyberjober@gmail.com

**Abstract.** This paper explores the parameters of network and content in the social network for sharing media, which affect its popularity. To assessment the danger of content and its distribution channel, emotions are used, which the content transmits and the distribution channel parameters: the number and length of content, audience size, likes, dislikes, intersection of audiences, the number of content created in the past, keywords and tags. The paper also presents the results of content risk assessment based on selected parameters. The proposed parameters and the method of their assessment are applicable in systems with a high level of automation to identify the most harmful sources of information on the social network for sharing media.

## 1. Introduction

Networks for sharing media content increase the number of their users every day, potential audience that is exposed to destructive content (DCs) growing and, consequently, the number of DCs is increasing (the larger the audience of the resource, the more potential advertisers it has) [1, 2].

DC in social networks for sharing media content includes such materials as texts, audio, images, video files containing pornographic materials, aggression, obscene language, illegal information, inciting racial hatred, propaganda of anorexia and bulimia, suicide, gambling, and narcotic drugs. Illegal content includes any calls for incitement of national discord, dissemination and suggestion of information about narcotic substances via the Internet, information propagandizing extremist actions [1,2].

Through popular social networks for sharing media content YouTube and Instagram, a large number of people have been recruited into illegal organizations. Social media content has been an instrument of “color revolutions”, the success of which largely depended on the human factor.

## 2. Methodology of destructive risk assessment

The key parameters of the analysis [1-2] are the risk of user involvement and popularity of the destructive content (DC).

The risk of user involvement can be estimated as follows:

$$Risk(i)_{post} = \frac{(C_t + V_t + S_t + L_t)S}{S} * 100\%,$$

where  $C_t$  – comments count for the destructive content during its existence;  $L_t$  – the number of likes for the destructive content during its existence;  $S_t$  – the number of reposts for the destructive content



during its existence;  $V_{\text{DC}}$  – the number of views for the destructive content during its existence;  $S$  – the number of subscribers of channel or community.

The popularity of DC in the social network YouTube can be defined as the number of views of this content. For evaluation, we will use the following content and channel parameters [3]:

- Expressions transmitted by this content.
- The parameters of the channel that hosts the content.
- The length of the content.

### 3. Emotion and duration of content analysis

For the analysis and identification of the parameters affecting the popularity of content in a given social network, there were collected 500 content for each of the content categories. Once obtained, from the samples where the number of views is extremely small ( $V < 500$ ) were removed from the content and content without subtitles, since emotion analysis will hereinafter be made based on the textual representation of the content.

To determine the emotions that the content transmits, the IBM Watson Text analyzer was used, which determines the presence of emotions in the content, such as:

- Joy.
- Anger.
- Fear.
- Sadness.
- Confidence.
- Analytical.
- Positivity.

Emotion score is the score for the emotion in the range of 0 to 1. A score greater than 0.5 indicates a high likelihood that the tone is perceived in the content.

The data obtained from the analysis are presented in table 1.

**Table 1.** Average values of emotion sampling parameters.

Emotions	mean	$\sigma^2$
Joy	0.3539	0.1019
Anger	0.0694	0.0358
Fear	0.0638	0.0340
Sadness	0.2374	0.0879
Analytical	0.1133	0.0612
Tentative	0.3506	0.1270
Confidence	0.0216	0.0147
Positivity	0.0141	0.0789

Analysis of emotions affecting the popularity of content through logistic regression:

$$viral = \frac{1}{1 + \exp\left(-\left(\alpha_t + \beta_1 * pos_t + \beta_2 * joy_t + \beta_3 * fear_t + \beta_4 * anger_t + \beta_5 * analytical_t + \beta_6 * sadness_t\right)\right)},$$

where viral = 1: content that is popular, and 0 if otherwise,  $\alpha_t$  – some constant, a  $pos_t$ ,  $joy_t$ ,  $fear_t$ ,  $anger_t$ ,  $analytical_t$ ,  $sadness_t$  – content emotions, and  $\beta$  – regression results (table 2).

**Table 2.** The impact of emotions on the popularity of content.

Emotion	$\beta$	Emotion	$\beta$
Joy	0,002921	Tentative	0,000503
Anger	0,001043	Positivity	-0,000007
Fear	0,003835	Analytical	-0,008661
Sadness	0,002453	Self-confidence	-0,003116

Based on the regression results presented in table 2, it can be concluded that users prefer the content that is conveyed by such emotions as fear or joy.

The risk of involvement from the emotions that the content conveys was also assessed. The results are shown in table 3.

**Table 3.** Correlation of risk between involvement and emotions.

Emotion	Risk of involvement
Joy	57%
Anger	64%
Fear	33%
Sadness	48%
Analytical	12%
Confidence	9%
Tentative	21%

Further, the content duration was considered according to the following categories:

- short (up to 4 minutes);
- medium (4 to 15 minutes);
- long (more than 15 minutes).

Table 4 presents the average values of risk of involvement for content of various destructive directions.

**Table 4.** The dependence of the risk of involvement on the length of the content (%).

Category	Drug promotion	Denial of traditional values	Propaganda of terrorism	Extremism	Inciting Ethnic Hatred	Social destabilization
Short	9,8	15	17,6	17,9	33,5	13,7
Medium	4,8	9,8	23,3	28,8	41,5	48,2
Long	0,1	5,5	10,6	25,3	24,9	21,7

#### 4. The parameters of the channel analysis

Now onto the parameters of the channel on which the content will be placed. The target audience for content consists of individuals who are directly connected to the content creator. Thus, the size of a network determines the total number of users (NumSeeds).

For a homogeneous network (Homogeneity), this study focused on some terms interest, which (in the context of YouTube) can be assessed through general subscriptions. Consumers who share common interests are likely to have many overlapping subscriptions. To get a measure of interest in some terms, the idea of an affiliate network is used to analyze social networks [4]:

Let be  $Sub_{jg}$ -binary variable indicating the subscription status of an individual user  $j$  on channel  $g$ . We can define the level of human's interest of some terms between users  $j$  and  $k$  as [4]:

$$Terms_{jk} = \sum_{g=1}^G Sub_{jg} Sub_{kg},$$

where  $G$  represents a complete set of subscriptions of all consumers of the original audience.

Based on this human interest [4] of some terms, it is possible to estimate the homogeneity of the network from equation [4]. It represents the degree of subscription overlap among initial video  $I$  consumers as part of all possible overlaps across the entire subscription set [4]:

$$Homogeneity_i = \frac{\sum_{j=1}^{NumSeeds_i} \sum_{k \neq j}^{NumSeeds_i} Terms_{jk}}{G * (G - 1)}.$$

The analysis takes into account the influence of other demographic, historical and related factors:

- the volume of previous contribution in the number of videos posted in the past (Vol);
- the popularity of a past contribution is measured by the average number of views of past videos (AvgView).

Also, since the correctly selected title and tags affect the visibility of the video, we introduce two variables:

- the presence of keywords in the title of a video (Keyword);
- the number of tags affixed to a video (Tags).

Since the number of views of a viral video probably contains several views of the same user (and these events may not be systematic), this work used a multiplicative formulation similar to proportional risk models, which provides a mechanism for outputting repeated events [5].

Let be  $View(t)$  – the number of video views  $i$  before time  $t$ , and  $dView_i(t)$  there will be an increase in views over a short period of time  $[t, t+dt]$ . Speed function is defined as the expectation  $dView_i(t)$ . Similar to the proportional formulation of risk, the proportional rating model represents the repetition rate in a multiplicative form as [5]:

$$dR_i(t) = \exp(\beta' X_i(t)) dR_0(t),$$

where  $R_0(t)$  is an unspecified continuous function,  $X_i(t)$  is a vector of time-independent or time-varying parameters described above.

To evaluate this model, 100 contents were used, and 25 were used as a test sample. Compared with the model, which does not have any independent variables, the proposed model showed good agreement with real data ( $\chi^2=537.63$ ,  $P<0.001$ ).

Since the parameters were entered into the model as exponents, the percentage change in the diffusion rate due to a single parameter change was designated as  $100 \times [\exp(\beta) - 1]$ . Table 5 presents the estimation of model parameters.

**Table 5.** Estimation of model parameters.

Parameter	$\beta$	$\sigma^2$	$\rho$
NumSeeds <sub>i</sub>	0.0017	0.0001	<0.001
TieStrenght	0.176	0.083	0.034
SeedConnection	-0.00004	0.00001	<0.001
Homogenety	0.518	0.168	0.002
Homogenety <sup>2</sup>	-1.182	0.324	<0.001
Vol	-0.0008	0.003	
AvgView	0.0003	0.00002	<0.001
Keyword	0.291	0.00003	<0.001
Tags	0.00002	0.00001	<0.001

## 5. Conclusion

Based on the data obtained above, it can be argued that the popularity of content depends on many parameters such as emotions - users like content that conveys joyful emotions more than content duration - users are not ready to spend too much time on content on a social network.

Big influence on the popularity of content have the parameters of the channel on which it is located. Every 100 additional consumers of content on the channel contributed to an increase in the rate of diffusion by 17.01 percent.

Thus, in the course of the study, content risk assessment parameters were proposed based on the YouTube social network, which includes the study and analysis of 2 aspects: the emotions that the content conveys and the content distribution channel parameters. As a result of taking these two aspects into account, we revealed the parameters of the content and the channel as a whole, which to a greater degree influence the popularity of the content. It should be noted that the proposed parameters of the content and its distribution channel will allow more accurate determination of potential destructive content than previously proposed, based only on the analysis of emotions caused by content, positivity, the presence of keywords in the title [6-8] makes them applicable in the systems automated monitoring of the social network to identify the most dangerous sources of distribution of harmful information.

## References

- [1] Gubanov D A, Novikov D A and Chkhartishvili A G 2010 *Social network YouTube: models information influence, management and confrontation* (Moscow)
- [2] Moss E 2016 *What content is popular on YouTube* (Mainstream)
- [3] Wojnicki A C and Godes D 2006 Word-of-Mouth as Self-Enhancement *HBS Marketing Research* 06-01
- [4] Wasserman S and Faust K 1994 *Social Network Analysis: Methods and Applications* (Cambridge University Press)
- [5] Lin D Y, Wei L J, Yang I and Ying Z 2000 Semiparametric Regression for the Mean and Rate Functions of Recurrent Events *R. Statist. Soc. B* **62** 711-30
- [6] Ostapenko G A, Parinova L V, Belonozhkin V I, Bataronov I L and Simonov K L 2013 Analytical models of information-psychological impact of social information networks on users *World Applied Sciences Journal* **25** 410-5
- [7] Kalashnikov A O, Yermilov Y V, Choporov O N, Razinkin K A and Barannikov N I 2013 Ensuring the security of critically important objects and trends in the development of information technology *World Applied Sciences Journal* **25** 399-403
- [8] Kravets O J and Choporov O N 2018 The Problems and Peculiarities of Modelling Integrated

Systems of Heterogeneous Traffic Services *Journal of Siberian Federal University - Mathematics & Physics* **11** 581-7