

Conference Paper

Distribution Analysis of Information of an Extremist Nature in Open Sources over the Subjects of the Russian Federation

A. A. Aidaralieva and O. V. Alferova

National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Kashirskoe shosse 31, Moscow, 115409, Russia

Abstract

This article represents the distribution analysis of extremist information in the open sources of the subjects of the Russian Federation (RF) and for the first time the zoning of extremist information over the subjects of the Russian Federation are studied.

Keywords: extremism, data analysis, terrorism, open sources, media.

Corresponding Author:

A. A. Aidaralieva
djaella@mail.ru

Received: 11 December 2017

Accepted: 20 January 2018

Published: 13 February 2018

Publishing services provided by
Knowledge E

© A. A. Aidaralieva and O. V.

Alferova. This article is
distributed under the terms of
the [Creative Commons](#)[Attribution License](#), which
permits unrestricted use and
redistribution provided that the
original author and source are
credited.Selection and Peer-review
under the responsibility of the
FinTech and RegTech:
Possibilities, Threats and Risks
of Financial Technologies
Conference Committee.

1. Introduction

At the current time, the flow of information on the Internet, which can characterize the situation in this or that area, is very large. In particular, the media plays a very important role for politics. The media can be used for various purposes, often such as waging information warfare, manipulating the moods of the population, and so on. Often, people who perceive information from an open source are prone to fully believe such sources, without thinking about the possibilities of manipulating or distorting the information presented.

It is worth noting that in recent decades the threat of terrorist activity has increased dramatically. In this regard, the relevance of the activities associated with counter-acting the financing of terrorism and extremism becomes actual. Tracking the cash flows, analysis of activity in social networks and the Internet as a whole, analysis of the activities of terrorist organizations, and other areas of work.

A more detailed review of analytical programs for working with the text and their classification is presented in the article by N. V. Tanatara [3]. Also, more information on the importance of monitoring open sources and the media, as well as an overview of modern programs for implementing such monitoring is presented in the article by I. V. Yurik [4].

 OPEN ACCESS

A graduate of the Department of Financial Monitoring (NRNU MEPhI), V.D. Stetsurin, in his thesis on "Developing a methodology for monitoring the signs of activities related to extremism and terrorism in the Internet through IPAK" Avalanche 2.7 "looks at the importance of the Internet and open sources for extremist and terrorist activities. The author also reviews Russian open source monitoring and analysis systems, develops methods for keywords searching and sources. According to the results of the thesis, Stetsurin analyzes the collected information: sources of publication of banned materials were identified, resources were analyzed for extremist activity and propagation of radical sentiments.

An automatic information search is carried out with the help of keywords, from which search queries assigning information sources are formed.

The formation of the keywords list can occur as follows:

- Use of the Federal List of Extremist Materials. The name of the material, its author, the source of the publication of the material, specific fragments of the text (most characteristic for this document) can be used;
- Use of the list of organizations and individuals for whom there is evidence of their involvement in extremist activities and terrorism. This list is available on the website of the Federal Financial Monitoring Service.
- Use of slang and slang keywords. To do this, you must first manually collect the relevant information, from which to generate a set of keywords;
- View the meta-tag of the relevant publication's tag and update the keyword list with those that correspond to the topic. The meta-tag contains words that are used to search for this source by search engines.

2. Materials and methods

The purpose of this paper is to analyze the information distribution of an extremist nature in the mass media located on the Internet.

As information sources were taken the media of the subjects of the Russian Federation, in particular news sites. Collection of sources is carried out manually through search engines. Also there are directories of mass media in certain regions of the Russian Federation.

With the help of popular search engines like *Google* (www.google.ru) and *Yandex* (www.yandex.ru) a manual search of the necessary media sources for each subject

of the Russian Federation was carried out (560 sources were prepared for collecting information).

Using the methods described in [4] the search for relevant keywords was performed to generate search queries. As the key words were used words such as terrorism, terrorist attack, terrorism financing, terrorism sponsorship, organized group, criminal group, terrorist, terror, IGIL, bomb, fighter, fascism, fascists, genocide, extremism, extremist, OCG, combat operation, hostage, to collect money, collecting of money. From these words, using the language of search queries, more universal forms have been formed (for example, terrorism *).

At present, the use of automatic programs for searching and collecting information is quite common. Among such programs are popular: Avalanche, Babkee, SiteSputnik, CyberMode, Web Parcer and others. However, almost all of these programs are limited in search functions, often parsing pages of specified sites is carried out, or searching for keywords among the sources which have been already installed in the program. For the present work, it is necessary to combine both functions. Also, none of the above programs performs the required type of search in the provided demo version. In this regard, information search was implemented manually through a search engine.

In the Russian-speaking part of the Internet, the most widely used: *Yandex*, *Google* и *Rambler* (www.rambler.ru). Currently *Rambler* uses the searching techniques as *Yandex* uses, that is why the comparison of *Google* and *Yandex* has been made. According to the information received, where the amount of documents found by *Google* is almost 17 times greater than that of *Yandex*, the relevance of *Google's* results is 85.8% (*Yandex* 77.7%), and the quality of the thematic search from *Google* is about 23.8% (*Yandex* 20.1%), *Google* was chosen as the most optimal system for searching.

To search for information, we used queries from various combinations of keywords and a team of search queries to restrict the search medium to a given media source (for example, *terrorism * or extremism * site: Smolensk-i.ru*).

The search was carried out over publications in 2016. Based on the results of the search, 1317 publications related to terrorism and extremism were found on all the media sources collected.

3. Results

First of all, it is necessary to mention the nature of the sources under consideration. During the search for information, several types of media in the subjects of the Russian Federation were identified:



Figure 1: Top 10 subjects by the number of publications of an extremist nature.

- Mass media, publishing information about events that occur exclusively in this subject;
- The media, which publishes information about events occurring both in this subject, and in the country and the world, as a whole.

Accordingly, the amount of relevant information that was found depends on the type of media in the subject considered. It is also important to understand that the news resources of especially large subjects often integrate the information of geographically close subjects.

In the analysis it is necessary to distinguish what the data was collected for. An analysis of crime related to terrorism can be carried out, then one stream of information is considered. In the case of an analysis of media policy in the field of news related to extremism and terrorism, the range of information sought is significantly increased. In this paper we considered the second case.

The list of subjects of the Russian Federation with the largest number of publications on these subjects was made up by Moscow (133 publications, 10.15% of the total number of publications) and the Republic of Dagestan (129 publications, 9.84% of the total number of publications). The third place was taken by the Chechen Republic (83 publications, 6.33% of the total number of publications).

The zoning according to the results of the analysis is shown in Fig. 1.

Further, comparing these indicators with registered crimes [2] of a terrorist nature, it was found that Dagestan (907 crimes) and Chechnya (156 crimes) were in the top

three lists. Moscow fell into the top ten of the subjects with the highest number of registered crimes (58).

The media often publish news of high-profile lawsuits. A correlation coefficient between the number of publications and the number of crimes with a terrorist nature comprises 0.57, which indicates the slight relationship between these two indicators. In addition, in these regions there is a large number of publications related to the prevention and fight against terrorism and extremism, for example, news about the National Guard of the Russian Federation, news about the commissions on combating terrorism, etc.

Also, such subjects as Chechnya, Dagestan, Tatarstan, Bashkortostan have a large part of the population with a Muslim faith. Also a fairly large number of Muslims live in Moscow and St. Petersburg.

If we talk about the content of the publications in Dagestan and Chechnya, in these subjects the vast majority of publications are connected with the prevention and struggle of terrorism. Dagestan and Chechnya are located on the border of the Russian Federation, close to countries with a high terrorist activity (Iraq, Syria, Iran, etc.).

Subjects of the trio with the least number of publications were Ivanovo Region, Kirov Region, Kostroma Region. Comparing this sample with the subjects with the lowest number of recorded crimes of a terrorist nature, only Murmansk region was the intersection of these two samples.

Also, the average number of publications per subject of the Russian Federation was considered - approximately 15 publications per year. That is a little more than one publication for a month. 24 subjects of the Russian Federation a month, on average, have more than one publication on topics related to terrorism and extremism. This is a fairly small number, but it must be remembered that in reality the distribution of publications is very uneven (Fig. 2).

It is difficult to make assumptions about the reasons for such an uneven distribution as the number of sources per subject is rather small. In the following papers, it is possible to increase the number of sources considered and to carry out a more detailed analysis of the distribution of the publications number.

4. Conclusion

At present, there are practically no people who would not use the Internet for various purposes. In this regard, the value of open sources in the network is significantly increased, but at the same time, there is a growing threat of using such sources for

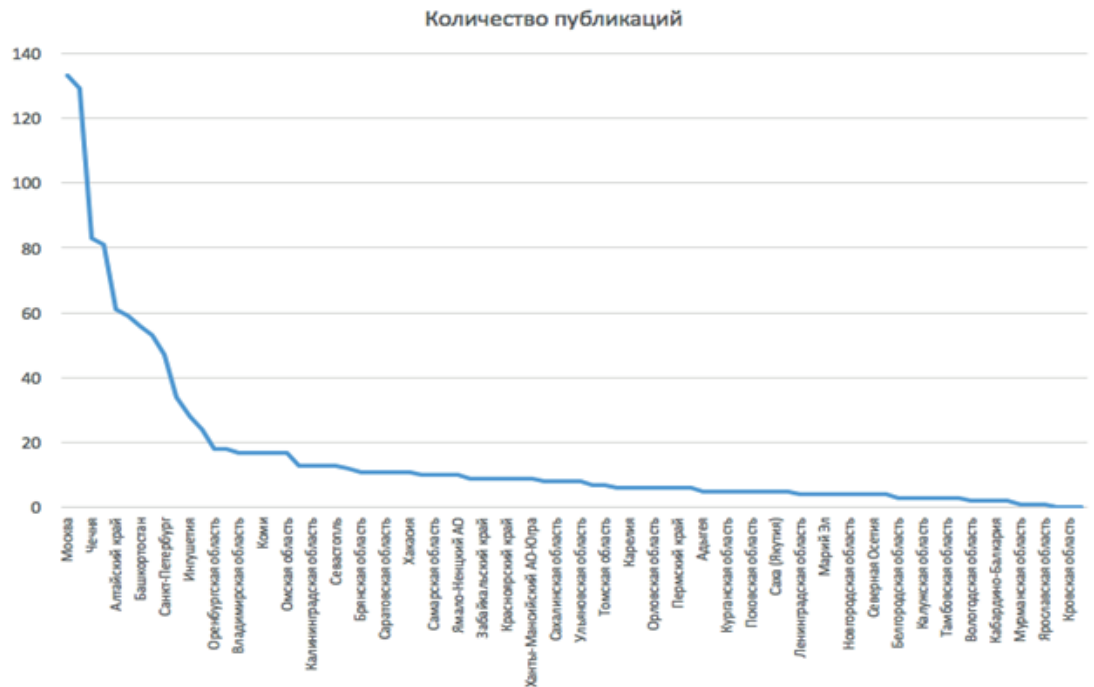


Figure 2: Distribution of the number of publications over subjects.

illegal purposes. Regarding the media, located on the Internet, the importance of publications available on the Internet versions of publications increases.

In the course of this work, media values for extremist and terrorist activities were considered, lists of media for each subject of the Russian Federation were formed, and a list of relevant keywords for search queries was compiled.

The largest number of publications was placed in such subjects as Moscow, the Republic of Dagestan, the Republic of Chechnya, the Amur Region, the Altai Territory, the Arkhangelsk Region, the Republic of Bashkortostan, the Astrakhan Region, St. Petersburg, the Republic of Tatarstan.

The reasons for such indicators are the location of the subjects, the religion of the population, the specifics of the material published by the media.

In the future, it is possible to increase the characteristics by which data analysis is carried out, the number of media under monitoring is increased for each subject, and a program is created for automatic search of information with the required specified parameters.

References

- [1] Stetsurin V.D. Developing a methodology for monitoring the signs of activities related to extremism and terrorism in the Internet through IPAK «Avalanche 2.7» // Diploma project/ NRNU MEPhI. – Moscow, 2013 – C. 1-131.
- [2] Portal of legal statistics of the General Prosecutor's Office of the Russian Federation [Electronic resource]: Access mode: <http://crimestat.ru/> (date of reference 25.12.2016).
- [3] Tanatar N.V. Intellectual search and analytical systems for monitoring the media / N.V. Tanatar, A.G. Fedorchuk // Libraries of national academies of sciences: problems of functioning, development tendencies. - 2008. - Issue. 6th – C. 205-219.
- [4] Yurik I. V. Monitoring media in the practice of information services: from history to the present / I. V. Yurik, A.V. Shkutova // Library Light. – 2014. – №1 – C. 28-29.