

## Development of an information base and analysis of existing systems for detecting illegal content in social networks

Olexandr Terentyev, Doctor of engineering, Associate Professor, Department of Information Technologies Design and Applied Mathematics. Kateryna Rokhvarg, student of SICS-41, Andrii Novak student of TSW-51

*KNUCA, 31, Povitroflotsky Avenue, Kyiv-037, 03680 Ukraine*

### ABSTRACT

Implementation of an information base or systems for detecting illegal content on social networks, which will undoubtedly help to simplify and improve the function of detecting illegal content on the Internet.

*Keywords: social media, security measures, illegal content, users, content blocking, TikTok, content blocking bypass.*

### 1. INTRODUCTION

Today social networks have millions of users. As a consequence, they should not be viewed solely as a means of communication or a waste of time. With the right approach, they have every chance of becoming a powerful tool. In some cases, even dangerous. Therefore, every time you should think about what you are putting on the network.

Social networks can easily have a negative impact on the unformed psyche. Children and adolescents are often affected by this influence.

### 2. PURPOSE OF THE WORK

Access to Internet content is blocked in many states, with the support of all kinds of technologies. Let's consider by what means the access to Internet services and content in them is limited. And just like modern users try to bypass the blocking.

TikTok is a short video creation and viewing service owned by ByteDance, a Beijing-based company. This service is becoming more and more popular in China and other countries, becoming one of the most popular and downloaded applications on the Internet and was founded in the summer of 2018.

But, unfortunately, TikTok is a fairly young social network, unlike Facebook, Twitter and many other competitors. Which during its time on the market has not yet become so advanced and cannot provide full control of the information transmitted in its service.

Access to a number of foreign sites from China is restricted under the Golden Shield project (the so-called Great Firewall of China). Web pages are filtered by national security related keywords as well as a blacklist of site addresses.

But this method is used only in China, which controls most of the transmitted information in its country between users. Other countries cannot afford to block most of the well-known social networks, websites and similar resources. Therefore, we have to use methods of blocking information and illegal content on the Internet.

As one of the means of avoiding censorship, including on Weibo, in recent years a special "coded" recording system known as the "Martian language" has been called. Which most of us saw in our news feed, but most likely did not think about what it is used for and how it works.

One of the options for the transmission of illegal content may be the replacement of some symbols with others, which is

called the "Martian language". He is also a threat of terrorism, the sale of drugs or alcoholic substances. This method of bypassing the rules for using Internet services makes it very difficult to control the incoming information to the masses.

For example, if you take the phrase "alcoholic substances", you can replace it with "@lc0h0lic\$ub\$t@nce\$". That is, we see that the character "a" from the English alphabet has been replaced by another character "@", which reads like "at". By the same analogy, the letter "o" becomes the number "0", the letter "S" can be replaced with the dollar symbol "\$". And this can be done with any letter of absolutely any alphabet, replacing them with other symbols.

Existing blocking methods cannot recognize this type of bypassing the rules, so this method is becoming popular and increasingly common on the Internet. Based on this, there is a need to create a new and more modern base of possible substitutions of symbols, that is, a list of symbols from the "Martian language".

For often this method of encrypting information is used in their videos by users of a social network called "TikTok". Since for advertising other applications; mate in the video, in the caption and comments; user account may be blocked. And since the test in the video content is superimposed on top of the videos themselves, this method greatly complicates the existing systems for blocking illegal content. Since artificial intelligence is engaged in blocking, which is not able to recognize non-existent words invented by a person.

There are millions of such examples on social networks, which, unfortunately, leads to a negative impact on the unformed psyche of adolescents and children who unconsciously find themselves in such situations.

In today's world, the influence of social networks is very great due to the universal availability, the possibility of simple exchange of information. High popularity among the younger generation, which uses social networks as the most accessible way of self-expression. A shortcoming that is immediately noticeable is the absence or change of truthful information, which leads to an incorrect or distorted perception of it. That in the future can negatively affect the psyche of a child or a teenager.

Teens with a fragile psyche can fall under the influence of other people's emotional reactions and inaccurate information.

And as you know, among the stages used and methods of instilling the necessary information through the media, the following stages can be cited:

- immersion stage.
- manipulation stage

- management stage.
- the stage of suppression.

These phenomena turn out to be even more harmful in the absence of available reliable information from official sources or in the absence of trust in official sources of information.

Therefore, it is required to improve the existing artificial intelligence, its program for implementing the search for illegal content on the Internet. For a faster and more correct response to similar problems and situations in social.

### 3. CONCLUSION

This article described in detail high-profile cases of bans on the Internet, namely, how to bypass the blocking of illegal content in modern social networks and instant messengers.

In the course of studying the possibility of blocking this content on social networks, many scientific and articles on these topics have been researched.

As a result of the above material in this article, IT giants are now developing intelligent content definition services. They are based on artificial intelligence and machine learning technologies.

In some countries, people are also the "tool" for blocking. As an example, in the already mentioned China, millions of employees are engaged in network monitoring. But in the blocking bureau of Saudi Arabia there are much fewer of them. Users themselves help them in their work - more than a thousand requests are received per day. Since artificial intelligence is not able to cope with all the methods of bypassing locks invented by mankind

Then the results of this article, it is obvious the need to create an information base for tracking illegal content on social networks and improving the existing artificial intelligence.

### References

- [1] «What the North Korean internet really looks like» By Saira Asher BBC News <https://www.bbc.com/news/world-asia-37426725>
- [2] «Zucked: Waking Up to the Facebook Catastrophe» By Roger McNamee