Methodology for Countering Malicious Information on Social Networks

Svitlana Popereshnyak[1], Viktoriia Zhebka[2] and Anastasiya Vecherkovskaya[3]

[1] National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 37, Prospect Beresteiskyi, Kyiv, 03056, Ukraine
[2] State University of Information and Communication Technologies, str. Solomyanska, Kyiv, 03110, 7, Ukraine
[3] Taras Shevchenko National University of Kyiv, 24, Bohdana Gavrylyshyn Street, Kyiv, 02000, Ukraine

Abstract

This work is devoted to an urgent problem, namely, the fight against malicious information on social networks. The goal of this work is to increase the effectiveness of countering malicious information in social networks by analyzing the sources of malicious information and automating the process of selecting countermeasures. The theoretical significance of this work lies in its contribution to the development of the theory and methodology of information security. The proposed approach allows us to determine scientifically based requirements for solving problems related to the analysis of sources of malicious information on social networks and countering both the message itself and its source. In addition, the developed models, algorithms, methods and architecture can be included in the operator's decision support system in order to combat malicious information. The proposed models, algorithms, methods and architecture, as well as their practical implementation, together provide a solution to the current scientific and technical problem of increasing the effectiveness of countering the spread of malicious information on social networks.

Keywords

Model, algorithm, methodology, malicious information, social networks, information security

1. Introduction

The deep implementation of social media in daily life is huge, and its advantage is that the participants of communication can quickly express their opinions to a large audience and share media files. Nowadays, social networks play not only the role of a means of communication, but also a tool for information dissemination. One of the obvious problems of information security in modern society is the spread of malicious information. Terrorist and criminal groups are increasingly using the means of information influence, developing strategies to expand their influence and attract new supporters through social networks. Therefore, one of the key elements of information security is to control, analyze and actively counteract malicious information in social networks. The concept of "malicious information" is considered by experts from various sciences, but no consensus has yet been reached.

Currently, the problem of combating malicious information has an insufficient number of scientific and technical solutions. The known means of detecting and counteracting malicious information in social networks do not meet the requirements for speed, completeness, accuracy and adequacy of decisions. This is due to several reasons, including the division of systems into two independent modules (Figure 1): monitoring and counteraction.

In between these modules is the operator, which plays a central role. In addition, social networks have a complex structure and contain many different messages, which is often not taken into account when defining countermeasure targets, such as message type, message source and

COLINS-2024: 8th International Conference on Computational Linguistics and Intelligent Systems, April 12–13, 2024, Lviv, Ukraine

spopereshnyak@gmail.com (S. Popereshnyak); viktoria.zhebka@ukr.net (V. Zhebka); vecherkovskaia90@gmail.com (A. Vecherkovskaya

0000-0002-0531-9809 (S. Popereshnyak) ; 0000-0003-4051-1190 (V. Zhebka); 0000-0003-2054-2715 (A. Vecherkovskaya)

other parameters. It is important to note that huge volumes of messages need to be processed in real time and targets for countermeasures need to be identified quickly. Manually, a countermeasure operator cannot completely stop the spread of malicious information.

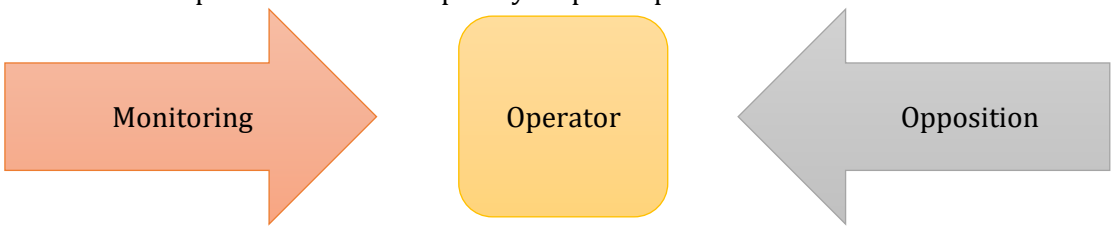

Figure 1: System modules for detecting and countering malicious information on a social network

Consequently, the main problem in combating malicious information in social networks is directly related to the current trends in the development of the information sphere, namely: (1) increasing the volume of messages containing malicious information; (2) increasing the speed of malicious information dissemination; (3) increasing the speed of message replication; (4) increasing the speed of new sources of information dissemination in social networks; (5) increasing the number of ways to attract the attention of the audience; (6) increasing the level of heterogeneity; (7) increasing the number of ways to attract the audience's attention; and (8) increasing the level of information dissemination in social networks. This requires improved effectiveness in combating malicious information on social networks, including more rapid and well-founded countermeasures.

Thus, the task that this study addresses - developing models, algorithms and methods to combat malicious information in social networks - is highly relevant.

The goal of this work is to increase the effectiveness of countering malicious information in social networks by analyzing the sources of malicious information and automating the process of selecting countermeasures.

The theoretical significance of this work lies in its contribution to the development of information security theory and methodology. The proposed approach makes it possible to define scientifically sound requirements for solving problems related to analyzing the sources of malicious information in social networks and counteracting both the message itself and its source. In addition, the developed models, algorithms, methodology and architecture can be included in the operator's decision support system to combat malicious information.

2. Related Works

The emergence of online social networking platforms has completely transformed how students interact with information and one another. With the internet's ascent and the interactive nature of online spaces, a global phenomenon known as social networking has emerged. This encompasses a wide array of activities, ranging from the creation of virtual communities to casual conversations and blogging. A study [1] explores the extensive utilization of Social Networking Sites and their influence, particularly on student. The primary objective of the research [2] was to explore the dynamics of knowledge sharing within academic social networks among university students. The results of the study revealed a notable correlation between perceived personal outcome expectations, perceived social expectations, and knowledge sharing among students.

As social media usage escalates, individuals using these platforms become increasingly susceptible to their negative impacts. Detecting cyberbullying on social media platforms poses a significant challenge, particularly due to the constant evolution of slang. Nonetheless, in the paper [3] proposes a practical solution—an application designed to identify cyberbullying across various social media platforms, leveraging data from Twitter and Wikipedia. The paper utilizes Deep Learning techniques to accomplish this task effectively.

In the study [4], two prevalent algorithms for network community detection were examined: Agglomerative Hierarchical Clustering and the Louvain Method. The research delved into their

mechanisms, investigating and contrasting their implementation nuances and the outcomes of their clustering behavior on a standardized dataset.

Advancements in technology have resulted in the accumulation of vast amounts of data from diverse sources, such as biological and social networking data. Consequently, there has been significant interest in social network analysis, given the abundance of raw datasets that can be conceptualized using a network framework. The majority of these datasets can be represented as social networks, characterized by a graph structure comprising actors and their relationships. Numerous tools have been developed for social network analysis, aimed at extracting insights from these networks. In [5], an enhanced version of NetDriller is presented, incorporating new essential features, including the construction of social networks through data collected from platforms like Twitter, IEEE, and DBLP.

As machine learning techniques increasingly intersect with real-world scenarios, the application contexts for these algorithms become progressively complex. Various domains across different fields have embraced and profited from the implementation of diverse machine learning algorithms. This complexity is particularly pronounced within the realm of social networks [6].

Unfortunately, there are limited studies that have explored the integration of convolutional neural networks for automating opinion discretization. In their paper [7], the authors introduce a novel distributed architecture aimed at addressing the challenge of opinion classification mining. With experimental results yielding high accuracy (72.99% ± 3.64), it can be inferred that implementing the authors' proposed distributed framework for opinion discretization on Facebook is indeed viable.

The study outlined in reference [8, 9] explores primary categories of social networks and their respective analytical methodologies. It delves into various types of connections and scrutinizes issues pertaining to ties within social networks. Additionally, it investigates and confirms the correlation between graph theory principles and the analysis of social networks.

Social networks have experienced significant success in facilitating online social interaction. However, malicious users exploit these platforms to disseminate rumors. Recent studies indicate that integrating social applications can enhance efficiency. Regrettably, new security challenges arise as malicious users exploit this integration to spread rumors across multiple social networks. In paper [10], which addresses cross propagation in multilayer social networks, the S2IR2 model is introduced to analyze the dynamics of rumor spreading.

The research [11] examines several network metrics, including modularity-based algorithms for community detection and dsynamics within and between groups. Additionally, it explores network measures such as Degree centrality, betweenness centrality, closeness centrality, authority, and hub, which could correspond to essential leadership qualities such as influence, attentiveness, communication, adaptability, dissemination of information, and social adeptness.

Social network analysis proves to be a valuable tool in addressing challenges such as money laundering, identity theft, network fraud, cyberattacks, and similar issues. Numerous researchers have dedicated their efforts to investigating the dynamics of social networks [12-13]. The works [14-15] is dedicated to exploring methods for detecting and combating malicious accounts and spammers within online social networks. The paper [16] explores countering misinformation campaigns on social media using social network analysis, addressing challenges in identifying and attributing campaigns, tracing information flows, and understanding spheres of influence, ultimately proposing tactical approaches for mitigation.

3. Models and algorithm for source analysis and ranking of countermeasures

Based on the conducted research, a set of functional and non-functional properties of countermeasures against malicious information in social networks and requirements for countermeasures methodology are identified.

The following properties of countering malicious information in a social network are highlighted:

- Responsiveness - the time it takes to counter malicious information on social media;
- validity - a set of considered parameters for the selected objects of influence and countermeasures in the process of counteraction;
- resource consumption - the probability that the amount of resources used will not exceed an acceptable value.

The input and output parameters for the study were determined.

Given:

$$\text{DATASET} \subseteq \{\text{messages}, \text{sources}\}, \quad (1)$$

where $messages$ – is the set of messages containing malicious information, $sources$ – is the set of sources of these messages.

$$\text{MESSAGE} = < \text{messageURL}, \text{source}, \text{activity}, \text{messageType} >, \quad (2)$$

where $messageURL$ is the address of the post on the social network, $source$ is the source of the post, $messageType$ is the post type (post, comment, or reply to a comment), and $activityy$ is the characteristics of the post.

$$\text{SOURCE} = < \text{sourceID}, \text{sourceURL} >, \quad (3)$$

where $sourceID$ is the source's unique identifier, $sourceURL$ is the source's social media address.

$$\text{ACTIVITY} =< \text{countLike}, \text{countRepost}, \text{countView}, \text{countComment} >, \quad (4)$$

where $countLike$ is the number of "like" marks, $countRepost$ is the number of "reposts" (copies with a link to the source), $countView$ is the number of views, and $countComment$ is the number of comments.

Required Finding:

$$\text{DATASET\_MAX} \subseteq \{\text{messages\_max}, \text{sources\_max}\}, \quad (5)$$

where $messages\_max$ is the set of messages ($messages$) that will have the highest $activity$ characteristics compared to other messages in the set $messages$, and $sources\_max$ is the set of sources ($sources$) that are associated with the maximum number of messages ($messages$) in the set $messages\_max$.

The objective of the study is to develop:

1. malicious information models based on the social network model and source.;
2. a set of algorithms for analyzing sources of malicious information in social networks and ranking countermeasures.
3. techniques for countering malicious information on social media;
4. architecture and software prototypes of the components of a system for countering malicious information in social networks.

The aim of the research is to improve the effectiveness of countering malicious information in social networks. In this paper, the effectiveness indicator is defined through the indicator of validity, as well as considering the requirements for efficiency and resource consumption.

Based on the models of social network and malicious information source, a theoretical-multiple model of malicious information in a social network is developed, which includes such basic elements as:

- information object IO (from English information object),
- T information threat attribute (from the English threat),
- MIO malicious information object,
- An information threat attribute contained in a malicious information object $Token$ (token),
- discrete feature of an information object $Feeature$ (from English feature)
- connections between objects.

The set-theoretic model is formally represented as follows:

$$IO = \{io\};\ MIO = \{io\};\ MIOi = \{io\}, \quad (6)$$
$$MIO \subset IO;\ \forall io \in MIO: io \in IO,$$
$$MIOi \subseteq MIO;\ \forall io \in MIOi : io \in MIO,$$
$$Token_{mio_i} \subset T;\ Token_{mio_i} = \{t\},$$
$$CheckFeature(io, t) = \{True;\ False\},$$

$io \in MIO_i \Leftrightarrow \exists\, Token_{mio_i}: checkFeature\ (io, t) = True,$

where IO – is a set of information objects, io_1 – is a single information object, T – is a set of all possible attributes of an information threat, t_n – is a single attribute, MIO – is a set of malicious information (a set of malicious information objects), MIO_i – is a separate class of malicious information, $Token_{mio_i}$ – a set of attributes characterizing MIO.

To form a set of attributes of malicious information, consider an information and attribute model that includes the following elements:

1. information threat - specified by the countermeasure system operator;
2. malicious information in the social network - specified by the countermeasure system operator by forming a set of keywords;
3. information features forming the set of all possible features.

The developed set of models of social network, source and malicious information contains new classes and attributes of objects, new relations between them, and also allows to form requirements to algorithms for analyzing and evaluating sources and choosing countermeasures.

The set of algorithms for analyzing malicious information sources and ranking countermeasures (Fig. 1) consists of:

1. an algorithm for ranking sources by potential,
2. of the source estimation algorithm,
3. an algorithm for sorting the objects of influence,
4. a ranking algorithm for countermeasures.

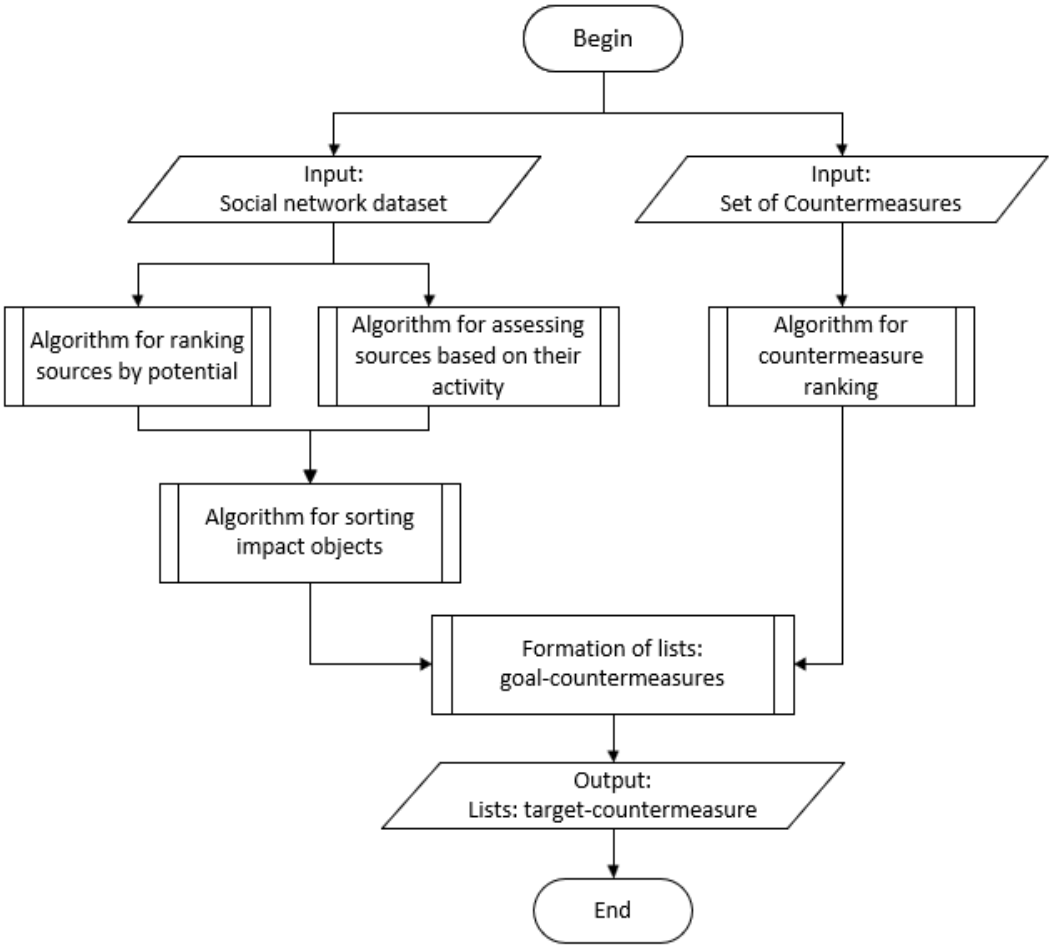

Figure 2: Schematic of a set of algorithms for analyzing sources and ranking countermeasures

The formal notation of the complex of source analysis and countermeasure ranking is as follows:

$$Z = SC \rightarrow \max, \tag{7}$$

$$f_1(S) \to I_p^S = \{0,1,2\},$$
$$f_2(S) \to I_i^S[0,1,2], \left(I_i = \frac{I_i}{\max I + 1}\right),$$
$$f_3(S) \to I_{pr}^S = I_p^S + I_i^S = [0,4], \quad (8)$$
$$f(C) \to complexity(c_x) = \frac{cw_x \cdot \sum_{i=1}^{|KC|} w_i \cdot \left(\sum_{j=1}^{|KC_i|} (cp_{x,i,j} \cdot lc_{x,i,j})\right)}{100 \cdot |KC|}, f(c) \to (0;1]$$

where: S - source, C - countermeasure, $f_1(S)$ - source potential index (I_p^S) is equal to 0, 1, 2 depending on the number of messages in the analyzed dataset belonging to the source. It is calculated using the "source potential ranking algorithm". $f_2(S)$- source influence index (I_i^S), whose value is between 0 and 2. The calculation of the inferentiality index follows the "source evaluation algorithm". $f_3(S)$ is the priority of the source (I_{pr}^S) as an influence object in the analyzed dataset. The "impact object sorting algorithm" is applied to obtain the value. $f(C)$ - ranked countermeasures based on their complexity. The ranking is done according to the countermeasure ranking algorithm.

At the output of the set of algorithms, lists of target-countermeasure pairs are generated, with the following rules for selecting objects of influence as targets ($target$):

$$\{source \in TARGET \mid I_{pr}^s \cong max\}, \quad (9)$$
$$\{messege \in TARGET \mid I_{pr}^s \cong min\}, \quad (10)$$

where $TARGET$ is the set of objects of influence.

The developed set of algorithms for analyzing sources of malicious information and ranking countermeasures differs from existing analogues by taking into account such attributes as source potential, user activity on the source page, and the number of views of messages with malicious information. The algorithm for ranking countermeasures differs from analogs by taking into account coefficients and complexity levels for each countermeasure. At the same time, the developed set of algorithms allows to form the requirements to the methodology of counteraction to malicious information and is the basis for the counteraction system.

4. Methodology for countering malicious information in social networks

Let's consider the methodology of countering malicious information in a social network. The methodology for countering malicious information in a social network consists of two stages: (1) the customization stage and (2) the exploitation stage. The operation stage of the technique consists of 3 steps and is presented in Figure 2.

At the same time, the customization stage of the technique consists of two steps:

Step 1. "Query system customization", in which the operator defines information threats and their attributes, and the countermeasure system generates and stores lists of threats and their attributes.

Step 2. "Countermeasure ranking", in which the operator selects available implementation agents, and the system generates and saves lists of available implementation agents.

Next, the operator selects the available countermeasures, the system generates a list of countermeasures and selects the complexity coefficients of countermeasures based on expert judgment, then the system generates and saves the list of ranked countermeasures.

The outputs of the methodology are:

1. possible information threats, attributes, countermeasures and their coefficients, available agents of countermeasure realization;
2. different parameters of impact objects, according to which the operator distributes his attention and the order of decision making on countermeasures;
3. formed target-countermeasure pairs to counter malicious information in social networks through available realization agents.

The developed methodology differs from the known ones by using the author's algorithms for analyzing sources and ranking countermeasures, which increases the validity of decision making

about countering the target and choosing a countermeasure and reduces the operator's work time in the process of countering malicious information in the social network.

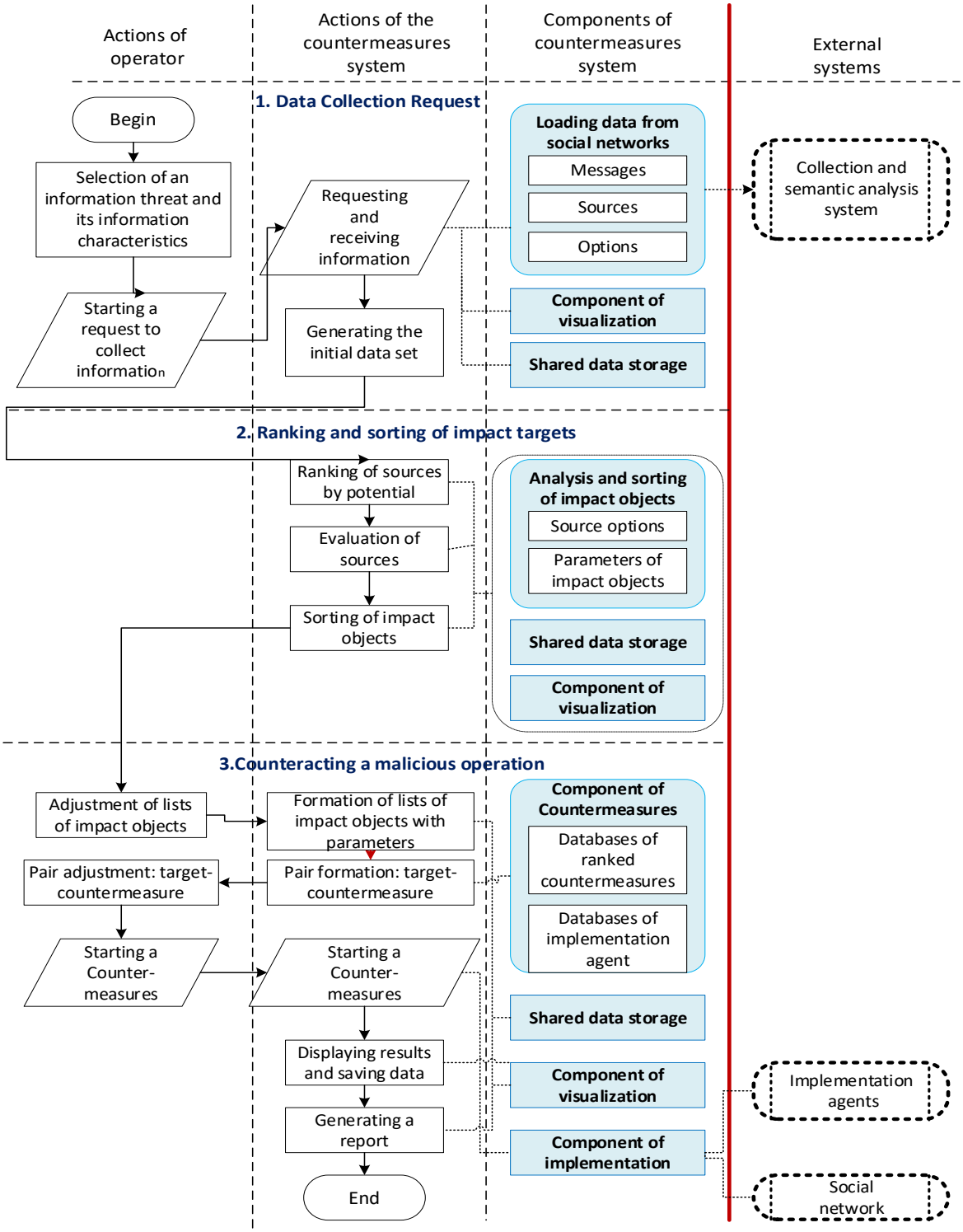

Figure 3: Representation of a methodology for countering malicious information at the operational stage

5. Architecture of a system for countering malicious information in social networks

The architecture and software prototypes of the components of the anti-malware system are presented in Figure 3.

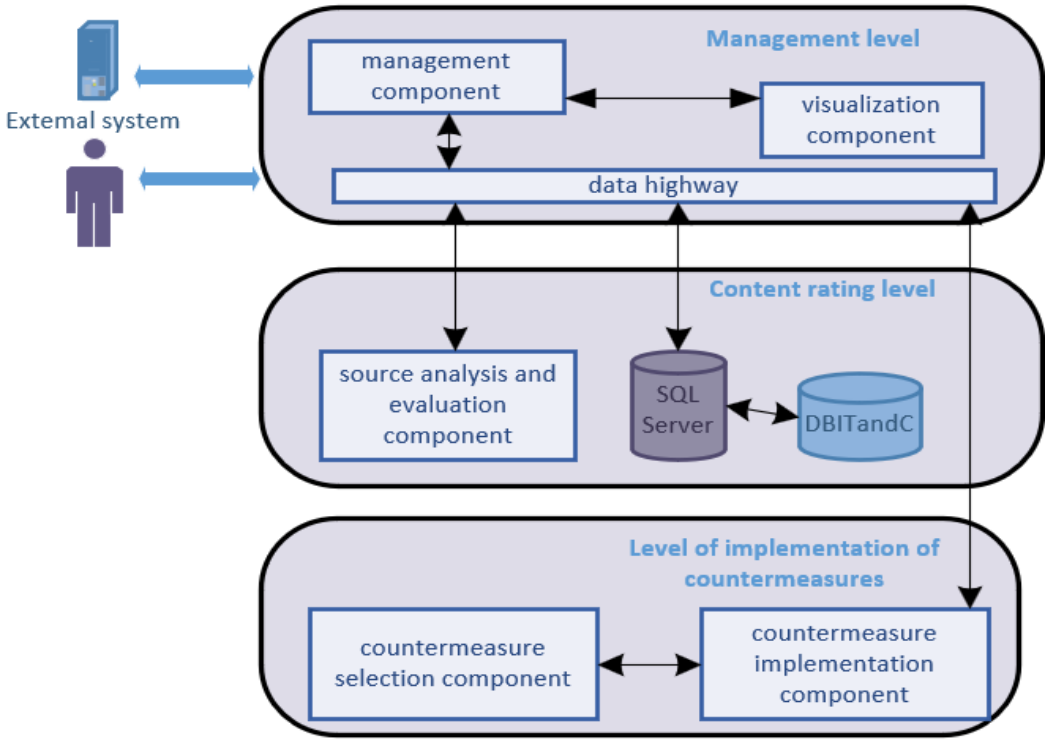

Figure 4: Architecture of the system of counteraction to malicious information in social networks stage

The architecture includes three levels:

1. management level (management component (flow and request management), visualization component (reporting));
2. content evaluation level (source analysis and evaluation component, SQL server and database);
3. the level of countermeasure implementation (countermeasure selection component and countermeasure implementation component).

Elements of the architecture are implemented as software prototypes:

1. a software prototype of a social network source analysis and evaluation component that includes a source ranking algorithm, a source evaluation algorithm, and an impact object sorting algorithm;
2. a software prototype of a countermeasure selection component that includes a countermeasure ranking algorithm, an expert judgment algorithm for generating coefficients;
3. a software prototype of the Information Threat and Countermeasure Database (DBITandC), which contains information on countermeasures against malicious information in the social network, the types of information objects to which countermeasures may be applicable, and the implementation agents through which countermeasures may be implemented.

6. Experiment

The experimental evaluation was carried out in several stages. First, the developed software prototypes and components were evaluated, then the operator's work time when countering malicious information without using a technique, and the operator's work time using a technique, were experimentally assessed. Resource consumption was assessed based on the data obtained in the two previous stages.

For experimental evaluation, data from the social network were collected, time measurements of the complex algorithms of source analysis and countermeasure ranking were made, CPU and RAM load indicators were obtained.

The information threat was chosen the dataset that contained to one of 19 categories: Adult English, Beer, Casino, Cigarette, Cigars, Cults, Dating, Religious, Marijuana, Occults, Prescription drugs, Racist groups, Religion, Spirits, Sport betting, Violence, Wine, Weapon, Other.

15,132 messages were collected from social networks, including posts, comments, and replies to comments. For each message, information was collected on the number of likes, comments, reposts, views, and information with the name of the source was obtained (Fig. 5). The data was obtained in csv format and converted into an excel workbook.

	A	B	C	D	E	F	G
1	SourseId	messageURL	messageType	likesCount	commentCount	RepostCount	viewCount
2			post	183	160	4	25926
3			comment	1	0	0	37
4			reply to comment	1	0	0	227
5			post	1	1	0	288
6			reply to comment	2	0	0	30
7			reply to comment	0	0	0	136
8			comment	3	1	1	298
9			comment	0	0	0	36

Figure 5: Example of experimental data set

Next, the large data set was divided into 10 small sets of 1000 messages each. Each small set was analyzed and ranked using a software prototype component for analysis and evaluation of sources in social networks, the following results and characteristics were obtained (Table 1, Table 2).

Table 1

Results of analysis and sorting of impact objects

Data set	Target1 (Source)	Target2 (Objects for the operator)	Target3 (MessageURL)
1	11	96	570
2	14	87	553
3	10	81	598
4	9	58	588
5	5	82	617
6	4	55	627
7	2	12	661
8	2	21	631
9	1	32	673
10	13	105	568

In the Target1 column, Source is recommended as the object of influence and shows the number of sources with high priority for counteraction, which own 334 messages out of 1000 for 1 data set, 360 messages out of 1000 for the second, etc. The Target2 column shows the number of messages that have medium priority and require additional evaluation by the operator. 118 Column Target3 recommends MessageURL as the target and shows the number of such low priority messages for each set. Thus, the sequence of the operator's work according to the results obtained is as follows (for the 1st data set): 1) the operator needs to agree on 11 objects of influence (sources) with high priority to counter them; 2) the operator needs to analyze 96 objects of influence, taking into account all characteristics (number of comments, likes, views, reposts, activity index, viewability index, potential, influence index); 3) check 570 targets last,

due to their low priority for counteraction. An experimental evaluation of a set of algorithms showed the efficiency of the approach to analyzing and sorting objects of influence.

Table 2

Results of an experimental assessment of the performance characteristics of a software prototype of a component for analysis and evaluation of sources in social networks

Data set	Time in seconds for the algorithm	Additional load on the CPU	Additional memory load
1	42.53	25 %	512 Mb
2	40.86	22 %	512 Mb
3	41.07	28 %	128 Mb
4	41.19	24 %	300 Mb
5	41.71	29 %	100 Mb
6	41.11	22 %	128 Mb
7	40.63	28 %	212 Mb
8	40.68	22 %	300 Mb
9	42.49	21 %	410 Mb
10	41.07	28 %	512 Mb

Next, an experimental evaluation of the countermeasure ranking algorithm was carried out. At the beginning, a list of countermeasures dependent on implementation agents was compiled. Then 10 experts were invited to participate in the experiment and were sent a voting questionnaire completed in the Google Forms service.

At the first stage of voting, experts assessed the possibility of using countermeasures to counter malicious information on social networks. Then a summary table was sent to the experts for the next vote, in which for each specified value the experts gave difficulty ratings from 1 to 10.

The following results were obtained at the output (Table 3, 10 lines out of 35).

Table 3

Result of expert assessment of countermeasures and their subsequent ranking

Countermeasure	Method of impact			Type of impact			Comp-lexity
	Positive	Negative	Neutral	Auto	Automated	Manual	
	2	1	3	1	2	3	
Message Notification	1	0	0	0	0	1	4
Source Notice	1	0	0	0	0	1	4
Blocking a message in the browser	0	1	0	1	0	0	6
Blocking the source in the browser	0	1	0	1	0	0	6
Blocking a message via antivirus	0	1	0	1	0	0	6
Blocking the source via antivirus	0	1	0	1	0	0	6
Blocking a message through the operating system	0	1	0	1	0	0	6
Blocking the source through the operating system	0	1	0	1	0	0	6
Blocking a message via a social network	0	1	0	1	1	0	8
Blocking a source via a social network	0	1	0	1	1	0	8

As a result of the experiment, countermeasures were ranked taking into account the coefficients and levels of complexity for each countermeasure.

7. Results of experimental and theoretical evaluation of the methodology

For experimental evaluation, data from the social network were collected, time measurements of the complex algorithms of source analysis and countermeasure ranking were made, CPU and RAM load indicators were obtained. Further, research and experiments were conducted to form the initial data, it was found out that the most costly process in terms of operability is the operator work time at the stage of setting up the methodology at the 1st, 4th step at the stage of operation of the methodology. To evaluate the indicator of operator's work time to make a decision on counteraction with and without the methodology, experiments were conducted, in which 10 experts participated. According to the results of the experimental evaluation of operability, the probability of performing the technique in a given time was calculated, which is $P_{\text{operability}}\left(T_m \leq T^{additional}\right) = 0{,}9942$, which meets the requirements ($P_{\text{operability}}^{\text{acceptable}} = 0.99$) for responsiveness.

Resource consumption was assessed using a number of specific indicators typical for step 2 of the operation phase of the social network anti-malware technique. CPU load, RAM utilization, and operator work time were considered. It is shown that the resource utilization estimate meets the requirements $P_{\text{res}}\left(r \leq R^{\text{acceptable}}\right) \geq P_{\text{res}}^{\text{acceptable}}$, where P_{res} is the probability that the resources r, spent on countering malicious information according to the methodology do not exceed the acceptable value $R^{\text{acceptable}} = 75\%$, $P_{\text{res}}^{\text{acceptable}}$is the acceptable probability value.

As part of the theoretical evaluation, the validity indicators for the developed methodology were compared with analogs, such as the solutions of Zerofox, ESET, Ithreat Cyber Group Inc. and others. It is shown that the developed methodology considers a larger number of parameters for the selected objects of influence and countermeasures in the course of countering malicious information in the social network, while meeting the requirements for other properties. Compared to analogs, the number of parameters taken into account when using the technique is larger, such that $N_{param}^{M} > maxN_{param}^{S}$, where N_{param}^{M} – is the number of considered parameters for the technique, $maxN_{param}^{S}$ – is the maximum number of considered parameters for analogs. At that $N_{param}^{M} = 12$, $maxN_{param}^{S} = 8$.

8. Discussions

A comparative analysis of the developed methodology with known methods in terms of the functionalities used, such as:

- A - possibility to form tasks of message collection and analysis for the monitoring system;
- B - the ability to customize the available countermeasures in the system;
- B- the ability to analyze the sources of messages in the resulting dataset;
- D - possibility of ranking and sorting the objects of influence in the obtained dataset;
- E - the ability to rank and sort the available countermeasures from the countermeasure database for each dataset;
- G - the ability to select the target of influence for counteraction.

The results are shown in Table 4 (the following designations and scores are used: "+" - presence of the parameter in the work (1 point); "+/-" - partial compliance with the parameter (0.5 points); "-" - absence of the parameter (0 points).

The analysis of the results of comparing the methodology for countering malicious information in social networks with analogs allows us to draw the following conclusions. First, none of the techniques, except for the proposed one, satisfies all functional requirements at the same time. Second, all techniques allow ranking countermeasures to a greater or lesser extent. Third, the parameters of messages, sources, countermeasures are considered only in the

proposed methodology and in the solution from Creopoint Inc. Fourth, the lag of the closest analogs from the proposed methodology ranges from 1.5 points to 4 points. That is, the proposed method wins over the closest analogs.

Table 4

Comparison of the developed method with known analogues

Methodology for countering malicious information on social networks	Parameters						Rating
	A	B	C	D	E	F	
Zerofox Inc. "Brand Protection"	+	+	+	-	+/-	+/-	4
ESET Internet Security	-	-	+	-	+/-	+/-	2
Ithreat Cyber Group Inc	-	+	-	-	+/-	-	1,5
Creopoint Inc.	+	+	+	+/-	+/-	+/-	4,5
AVG Internet Security	+	+	-	-	+/-	+/-	3
Developed methodology	+	+	+	+	+	+	6

Thus, the results obtained in the work allow us to assert the achievement of higher efficiency of the developed methodology compared to the known ones, which proves the realization of the final goal of the study - to increase the effectiveness of countermeasures against malicious information by analyzing the sources of malicious information and automating the choice of countermeasures.

9. Conclusions

The rise of the Internet poses a substantial risk to both personal and state data security. Consequently, the detection and mitigation of unsuitable content circulating on the worldwide web emerge as a matter of national significance.

The proposed models, algorithms, methodology and architecture, as well as their practical implementation together provide a solution to the actual scientific and technical problem of improving the effectiveness of countering the spread of malicious information in social networks. The results of the work constitute the following research outcomes:

1. A set of models of social network, source and malicious information is proposed, which differs from the existing analogs by the possibility of simultaneous consideration of the structure of information exchange in the social network, sources and malicious information.

2. A set of algorithms for analyzing malicious information sources and ranking countermeasures has been developed, which, unlike existing algorithms, takes into account connections and dependent attributes of objects in the social network, such as source potential, user activity on the page, number of message views, etc. The algorithms for ranking countermeasures take into account coefficients and complexity levels of each countermeasure. Countermeasure ranking algorithms take into account coefficients and difficulty levels for each countermeasure.

3. A methodology of countermeasures against malicious information in a social network is proposed, focused on automatic and automated selection of objects of influence and countermeasures against malicious information from a list of ranked countermeasures.

4. The architecture and program components of the system of countermeasures against malicious information are developed, which differs from existing architectures in that it supports ranking and selection of countermeasures available to the operator in the system for malicious information specified by the operator. The architecture contains original components for analyzing and evaluating the source of malicious information, a database with information on countermeasures for malicious information in social networks.

As recommendations for further development of the topic are to expand the class of algorithms for analyzing the behavior of sources and authors of messages, algorithms for analyzing the dissemination of information in a social network, integration of automatic and automated countermeasures mechanisms into existing architectures and systems.

References

[1] Z. A. Shanaa, K. M. Naser, M. Abualrish, E. A. Zaitoun and J. Yousef. "Effects of Social Networking Sites on the Academic Performance of Graduate Students," Tenth International Conference on Social Networks Analysis, Management and Security (SNAMS), Abu Dhabi, United Arab Emirates(2023): 1-6. doi: 10.1109/SNAMS60348.2023.10375420

[2] A. Kaba, S. Eletter and G. A. E. Refae, "Knowledge Sharing Through Academic Social Networking: The Impact of Personal and Social Outcome Expectations," Tenth International Conference on Social Networks Analysis, Management and Security (SNAMS), Abu Dhabi, United 3Arab Emirates (2023): 1-7. doi: 10.1109/SNAMS60348.2023.10375418

[3] M. Mahat. "Detecting Cyberbullying Across Multiple Social Media Platforms Using Deep Learning," International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India(2021): 299-301. doi: 10.1109/ICACITE51222.2021.9404736

[4] N. Motschnig, A. Ramharter, O. Schweiger, P. Zabka and K. -T. Foerster. "On Comparing and Enhancing Two Common Approaches to Network Community Detection," 2021 IEEE Global Communications Conference (GLOBECOM), Madrid, Spain(2021): 1-6. doi: 10.1109/GLOBECOM46510.2021.9685248

[5] S. Afra, T. Özyer, J. Rokne and R. Alhajj. "NetDriller-V3: A Powerful Social Network Analysis Tool," 2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Istanbul, Turkey(2022): 570-574. doi: 10.1109/ASONAM55673.2022.10068570

[6] D. Oreški, I. Pihir and D. Višnjiű. "Comparative Analysis of Machine Learning Algorithms on Data Sets of Different Characteristics for Digital Transformation," 2023 46th MIPRO ICT and Electronics Convention (MIPRO), Opatija, Croatia(2023): 1428-1433, doi: 10.23919/MIPRO57284.2023.10159910

[7] H. Xuan Huynh, V. T. Nguyen, N. Duong-Trung, V. -H. Pham and C. T. Phan. "Distributed Framework for Automating Opinion Discretization From Text Corpora on Facebook," in IEEE Access, vol. 7(2019): 78675-78684. doi: 10.1109/ACCESS.2019.2922427

[8] S. Popereshnyak, O. Suprun, Tools and methods for intersubjective relationships in cyberspace forecasting, Proceedings of the 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT ’2017, Lviv, Ukraine, 2017, pp. 244-247. doi:10.1109/STC-CSIT.2017.8098779.

[9] S. Popereshnyak, I. Yurchuk, Social networks: Analysis, algorithms and their implementation, CEUR Workshop Proceedings, 2870, 2021, pp. 811–821.

[10] Q. Han, M. Gu, L. You, F. Miao, Rumor, Spreading with Cross Propagation in Multilayer Social Networks, Proceedings of the Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking, ISPA/BDCloud/SocialCom/SustainCom ’2019, Xiamen, China, 2019, pp. 1641-1645. doi: 10.1109/ISPA-BDCloud-SustainCom-SocialCom48970.2019.00242.

[11] S. Garg, T. Gandhi, B. Panigrahi, Social Network measures association with social and intelligent behaviors in Dolphin network, Proceedings of the 11th International Conference on Cloud Computing, Data Science & Engineering, Confluence ’2021, Noida, India, 2021, pp. 655-659. doi: 10.1109/Confluence51648.2021.9377088.

[12] F. Long, N. Ning, C. Song, B. Wu, Strengthening Social Networks Analysis by Networks Fusion, Proceedings of the ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM ’2019, Vancouver, BC, Canada, 2019, pp. 460-463. doi: 10.1145/3341161.3342939.

[13] S. A. Bitaghsir, S. Kashipazha, A. Dadlani, A. Khonsari, Social-aware Mobile Road Side Unit for Content Distribution in Vehicular Social Networks, in: Proceedings of the 2019 IEEE Symposium on Computers and Communications, ISCC ’2019, Barcelona, Spain, 2019, pp. 1-6, doi: 10.1109/ISCC47284.2019.8969669.

[14] Kayode Sakariyah Adewole, Nor Badrul Anuar, Amirrudin Kamsin, Kasturi Dewi Varathan, Syed Abdul Razak. ”Malicious accounts: Dark of the social networks.”Journal of Network and Computer Applications, Volume 79(2017): 41-67. https://doi.org/10.1016/j.jnca.2016.11.030.

[15] Gianluca Stringhini, Christopher Kruegel, and Giovanni Vigna, Detecting spammers on social networks, in Proceedings of the 26th Annual Computer Security Applications Conference (ACSAC '10). Association for Computing Machinery, New York, NY, USA, 2010, pp. 1–9. https://doi.org/10.1145/1920261.1920263

[16] A. Bargar, S. Pitts, J. Butkevics and I. McCulloh, "Challenges and Opportunities to Counter Information Operations Through Social Network Analysis and Theory," 11th International Conference on Cyber Conflict (CyCon), Tallinn, Estonia, 2019, pp. 1-18. doi: 10.23919/CYCON.2019.8756832.

[17] Jain, A.K., Sahoo, S.R. & Kaubiyal, J. “Online social networks security and privacy: comprehensive review and analysis”. Complex Intell. Syst. 7, (2021): 2157–2177. https://doi.org/10.1007/s40747-021-00409-7

[18] Umit Can, Bilal Alatas. ”A new direction in social network analysis: Online social network analysis problems and applications.” Physica A: Statistical Mechanics and its Applications, Volume 535(2019). https://doi.org/10.1016/j.physa.2019.122372.