Harmful Effects of Fake Social Media Accounts and Learning Platforms

Nadiia Pasieka[a], Myroslava Kulynych[b], Svitlana Chupakhina[a], Yulia Romanyshyn[c], and Mykola Pasieka[c]

[a] *Vasyl Stefanyk Precarpathian National University, 57 Shevchenko str., Ivano-Frankivsk, 76000, Ukraine*
[b] *Ukraine Academy of Printing, 19 Pid Goloskom str., Lviv, 79020, Ukraine*
[c] *National Technical University of Oil and Gas, 15 Karpatska str., Ivano-Frankivsk, 76068, Ukraine*

Abstract

Social networks are an integral part of the modern information society, have no age or professional restrictions. On the basis of social networks conduct various research, advertising, and marketing campaigns. Increasingly, Internet opportunities are actively used to shape public opinion, which, in turn, can affect the quality of life, political processes in the country and the world. Therefore, the issue of cybersecurity of information in social networks requires special attention. One of the main problems is social bots or harmful accounts. They can play both a negative and a positive role, but most often they are a tool for reducing trust in social networks, mass theft of personal data, organizing information injections, and so on. Therefore, the ever-growing threats of using malicious accounts make the task of their effective detection especially urgent and directly related to the cybersecurity of both users and critical data. An analysis of existing methods for detecting malicious accounts showed that most of them are based on classification. Each of the methods to some extent solves the problem of dividing data into groups with the same characteristics. Therefore, there is a need to develop an effective method for detecting malicious accounts. For these purposes, a systematic analysis of clustering methods was carried out in order to select the most effective for solving the set tasks. Based on these results, it was determined that cluster analysis is the most unified tool for solving the problem of detecting bots. And to obtain the most reliable results, it was proposed to use a combination of two methods of cluster analysis – hierarchical and fuzzy. In the systems operating today, the detection of social malicious bots is carried out using strict methods and conditions. Suspicious user accounts are either blocked or subjected to additional verification, which may cause users to leave the social network or switch to another alternative social network. In the developed algorithm, at first, the system is loyal to all new accounts, and additional analytical methods are used to classify suspicious users. This allows suspicious accounts with ambiguous indicators to be subjected to additional classification after a certain period, after collecting statistics and monitoring their behavior, and only after confirming that they belong to harmful accounts, to offer verification.

Keywords

Social networks, cluster analysis, social bot, accounts, learning platforms.

1. Introduction

Currently, public social networks are an integral part of most spheres of human life, integrating almost all existing Internet resources. They effectively structure users according to political or religious views, interests, and hobbies, affecting almost all segments of the population, and are a powerful tool for the self-organization of both individual groups and society as a whole. Social networks, which unite 45% of the

---

Cybersecurity Providing in Information and Telecommunication Systems, January 28, 2021, Kyiv, Ukraine
EMAIL: pasyekanm@gmail.com (A.1); kumyr@ukr.net (B.2); cvitlana2706@gmail.com (A.3); yulromanyshyn@gmail.com (C.4); pms.mykola@gmail.com (C.5)
ORCID: 0000-0002-4824-2370 (A.1); 0000-0002-9271-7855 (B.2); 0000-0003-1274-0826 (A.3); 0000-0001-7231-8040 (C.4); 0000-0002-3058-6650 (C.5)

CEUR Workshop Proceedings (CEUR-WS.org)

world’s population every day, have become not only a means of communication but also a great source of information, entertainment hosting, a commercial platform with a set of effective tools for distributing services and goods. Naturally, stakeholders tend to use such limitless potential for profit and achieve their far from noble goals. However, on of the main threats to social networks is the so-called “social bots”, which contribute to an increase in mistrust of interlocutors, doubts about their reality [3, 4, 6, 9]. These are malicious programs, fake accounts capable of imitating human behavior. At the moment, bots create a lot of problems, both for ordinary users and for those who use social networks to conduct a marketing campaign or conduct social research. By using bot profiles on social networks, information about the true advantages and interests of portal users is greatly distorted. One of the main goals of using bot programs by cybercriminals is to disseminate information, both positive and negative, to promote the idea, which in turn interferes with SMM analysis (social media marketing is the process of attracting user traffic, attention to a brand or product through social platforms), due to the “cheat” of the number of participants in the communities. Therefore, it is necessary to determine which users of the social network are programmed, and be able to share the data stream generated by bots, and which by a person [11, 13, 16, 25].

As a rule, the mass distribution of specially programmed false accounts is used to:

- As a vehicle for legal cybercrime business.
- Organization of stuffing of information flows.
- Mass theft of personal data.
- Creating conditions for the deterioration of trust in social networks.
- Creation of false news feeds and fake votes.
- Creating problems for social marketing.

So, for example, in the case of SMM promotion on “TikTok“ or “Facebook“, the share of the target audience is “dead” accounts or duplicate accounts. According to shadow market researchers, dozens of services are offered by bot readers, selling them in large quantities [23, 34]. At the same time, the market volume reaches $570 million. The threatening scale of the use of social bots requires the creation of effective algorithms for their detection. Internet platforms and social services themselves are not overly concerned with this problem. As a result, not only ordinary users who have formed various communities suffer, but also companies that promote goods, brands, services through social networks [14, 26, 36, 40]. At the same time, the generated statistical data: the number of likes (positive ratings), the growth of members of social communities, and the number of publications, which have little to do with the number of real buyers. Thus, the task of recognizing malicious accounts on social networks and combating them remains relevant in the issue of cybersecurity [38, 39, 47].

Cybersecurity is the protection of the vital interests of a person and citizen, society, and the state when using cyberspace, which ensures the sustainable development of the information society and the digital communication environment, timely detection, prevention, and neutralization of real and potential threats to the national security of Ukraine in cyberspace.

The solution will be the development of methods of network analysis, which are designed to detect and classify communities in social networks, assess their connectivity, the degree of trust, as well as the development of effective algorithms for detecting malicious accounts [41, 42, 46]. To solve the problems set in the work, the following methods were used: methods of analysis and synthesis (in the disclosure of theoretical provisions and clarification of categorical apparatus), logical method (in grouping theory), methods of visualization (for graphical representation of social networks), methods of data collection from various Internet services, mathematical methods of pattern recognition, methods of hierarchical clustering, methods of fuzzy clustering (for conducting distribution of users into groups with the same features), methods of probability theory and mathematical statistics, methods of algorithm development (for development of detection algorithm), simulation modeling (when verifying the obtained results).

2. Analysis of the Impact of Harmful Accounts on Social Networks

Society plays an important role in human life: since childhood, people who are around us influence us in one way or another, there is a continuous interaction in the social sphere (in kindergarten, school, university, home, work), without which the full development of personality is impossible. Since the middle of the 20th century, there have been many published works of social scientists, including domestic ones, which propose different approaches of analysis in the sphere of social relations [7, 10, 17, 44]. In their works, by the social network, they meant structures of people connected with each other by common relations or interests. Later, with the advent of the Internet, they began to refer to specialized electronic portals in this way. However, the concept of "social network" has a broader meaning. Sociologist G. Gradoselskaya proposed the following definition: "Social networks are a special reality and a special philosophy of data analysis, which allows integration of various mathematical approaches—statistical, systemic, simulation—with the modern social theory." With the development of information technology, there are electronic portals that can display certain aspects of human activity in society, to store and accumulate information. A special place is occupied by virtual social networks, such as "TikTok," "Facebook," "Twitter," and others. Thus, a different definition of social networking is popular today. It is also called a virtual social network [2, 5, 15, 18, 19]. By virtual (photo) social network is understood the social structure of the Internet environment, the nodes of which are organizations or individuals, and connections mean established interactions (political, corporate, service, family, friendship, interest). Immediately after the registration of a new participant in the network, a profile is created, which initially contains information on completed personal data: age, gender, marital status, interests, education, and more [12, 48, 49]. This profile is more "developed" compared to its counterparts in blogs and forums, as the latter can only be part of the means of communication in a social network. Various groups and communities of interest are formed within the social network, such as music lovers, car lovers, study lovers, and job lovers. The ties between the members of such associations are strong enough that they can be easily identified. Fig. 1 shows an example of the described situation. Within the group, there are more ties between members, and they are stronger than with other members of the social network. Subgroups and communities can appear within communities, thus forming a hierarchy.

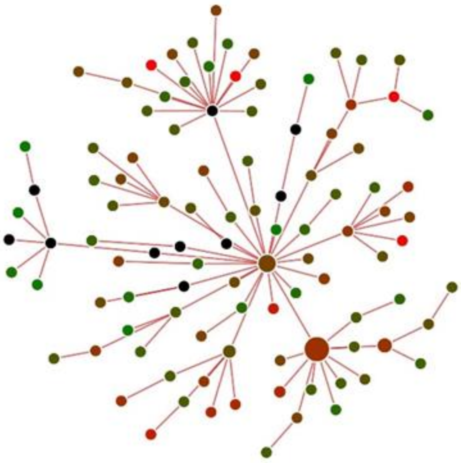

Figure 1: Hierarchy of groups in social networks

The tools by which communication takes place in virtual social networks are:

- A blog is an online journal in which posts are regularly added.
- Chat provides the ability to exchange text messages between several participants in the online mode.
- Forums, where a user can create a new topic that is accessible to others.

Other users can view it and leave comments. Social network analysis (SNA) is a branch of modern computer sociology that deals with the description and analysis of connections (networks) arising in the course of social interaction and communication. Big social data analysis is rapidly gaining popularity worldwide due to the emergence of online social networking services (SixDegrees, LiveJournal, Facebook, Twitter, Tik Tok, and others) in the early 1990s [20, 33]. Linked to this is the phenomenon of the socialization of personal data: facts of biography, correspondence, diaries, photo, video, audio materials, travel notes, etc. became publicly available. Thus, social networks are a unique source of data on the personal lives and interests of real people. This opens up unprecedented opportunities for solving research and business problems (many of which could not be solved efficiently before due to lack of data), as well as the creation of support services and applications for social network users. In addition, this explains the increased interest in social data collection and analysis on the part of companies and research centers. The analytical agency Gartner published a report titled “Cycle of Excitement for Technology, Fluttering.” According to the report, “Social Analytics” and “Big Data” technologies are currently at the “peak of inflated expectations”. In particular, Carnegie Mellon University, Stanford, Oxford, INRIA, as well as Facebook, Google, Yahoo!, LinkedIn, and many others are actively engaged in social data research. Companies-owners of online social networking services (Facebook, Twitter) actively invest in the development of improved infrastructure (Cassandra, Presto, Thrift) and algorithmic (new algorithms for search and recommendation of users, goods, and services) solutions for processing large amounts of user data. At the same time, when dealing with social data it is necessary to take into account such factors as the instability of the quality of user content (spam and false accounts), problems with ensuring the privacy of personal user data during storage and processing, and frequent updates of user model and functionality. All this requires constant improvement of algorithms for solving various analytical and business tasks.

For example, a team from the University of British Columbia in Vancouver, Canada, investigated the capabilities of so-called “social bots,” which are widely used by attackers and marketers for social networking activities. The two-month study was conducted to determine how vulnerable social networks and their users are to large-scale identity theft operations. The experimental portion of the study lasted two months. During that time the Canadian researchers obtained almost 350 gigabytes of information about the users of that social network with the help of “social bots” on Facebook alone. This alone makes it clear that today’s networks are very vulnerable to bots. According to the authors of the study, the “social” protection mechanisms in “Facebook” and other social networks exist, but they are not intelligent enough and cannot yet distinguish a real user from a bot, even if the latter acts entirely on automatic and without the participation of a live person. The study also notes that in the future, based on this or similar cyberbullies techniques, real campaigns to steal data from tens or even hundreds of thousands of people could be realized. The solution to this problem is the development of network analysis methods. Research laboratories on this topic have been established in many universities around the world. These methods can provide much more information—to identify and cluster communities, to assess the connectivity of such communities, the degree of mutual trust, opinion leaders. Together with content analysis to assess the tone of statements and more accurately outline the portrait of the client, his interests, and values. As a result, in 2012 another term appeared—social marketing analysis (SMM), which now offers several metrics based on network analysis. SMM is the process of attracting traffic or attention to a brand or product through social networks. This is a set of measures to use social media resources as channels to promote companies and solve other business problems. SMM tools based on network analysis, software products designed to improve the effectiveness of marketing campaigns carried out in social networks, development, and implementation of marketing campaigns, their analysis in real-time, are now appearing in the global market. When analyzing social networks, the focus should be on connections and not on the actors themselves. Typically, a social network is described by a graph or relationship matrix (Fig. 2).

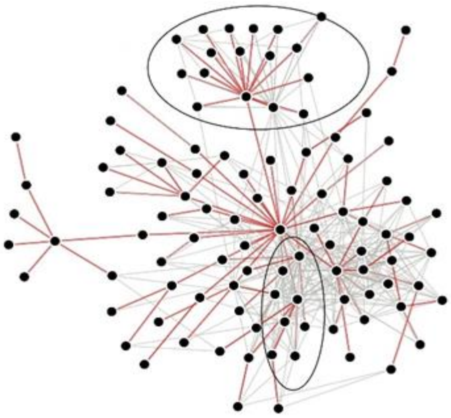

Figure 2: Example of a social network graph

One of the main modules of the social network analysis system is a module for detecting bots in social networks [22, 27, 31, 35, 43]. Many of the “TikTok” and “Facebook” groups consist of more than 45–50% of harmful accounts. The result is a general decrease in the effectiveness of actions in social media resources up to 30–40%. For example, when performing works on SMM promotion in the group instead of the target audience can also be attracted "dead" accounts and account doubles.

Classification of special bots for discrediting social and educational messages.

Social bots are a piece of software designed to mimic the behavior of a live person on a social or educational platform. Malware is used to impersonate a real user by stealing their data or to achieve other goals, such as promoting a false idea, brand, spreading anti-advertising educational platforms, creating interest in discussing some deliberately false topic for a large number of users. It is also a misconception that all bots are identical and created for the same malicious purposes. Ordinary users, as a rule, do not attempt to determine the sequence of their actions. Only by performing an in-depth system analysis of bot programs’ behavior can they be divided into two types [21]:

- Automatic: bots that perform simple, predetermined malicious actions.
- Automated: bots controlled by an operator who actively participates in semi-automatic mode (likes, reposts, etc.).

Thus, only advanced programs with elements of piece intelligence can automatically join communities, fill out a profile, even add users as friends. Such bot programs are most often used to distribute informational spam or to inflate scores. Managed malware also includes cloned pages from real news platforms, including educational ones.

Bots are also divided into the following categories:

- Avatars are profiles used by both real people and bots. This type of bot is also called cyborgs. Usually well-established social connections.
- News bots are bots that post the latest news on their profile. Needed mainly for the dissemination of this news. No harm is done if the news is not misinformation.
- Bots, which spread the message of real users. They do this in order to be similar to humans. One of the main purposes of their existence is information dumping.
- Gaming bots. Participate in games or other social networking applications, communicate with real people on behalf of real people. Most often used to scam app mentions.
- bots that participate in reputation scaling—increasing the number of friends and connections. Allow you to earn money for reposting.
- accounts that only publish reposts and other people’s news.

Next to the bots are the artificially created accounts, which are developed and maintained by real people. Depending on the purpose, time, and effort spent on maintaining the account, the pages of virtual personalities can be both robot-like and indistinguishable from real people. They can be divided into several types:

1. Accounts from likes exchanges. On profile exchanges you can buy almost everything—swear words, Sheri, comments, page signatures. Despite the fact that exchanges guarantee that they have real people working for them, many of them are aggregators of mini-bot farms.
2. Accounts with template pumping—according to their development corresponds to the script or a real person. Their quality is higher, but if you look closely, the posts in the feed appear at about the same time, in response, there are no interactions or there are lone comments.
3. Account with a non-template pumping—a bot with behavior similar to the actions of a real person. Different types of posts are used here, under which lone swear words or comments are encountered.
4. “Live person”—accounts that are created by a sophisticated program that convincingly mimics the actions of a person on a social network. Because of this, they do not arouse suspicion, such as on Facebook, because they are invisible to bot detection algorithms. “Live people” accumulate in friends and subscribed users an active audience and regularly interact with them.
5. Lomi bots are so-called opinion leaders. The top of the social bot chain. Hand-created and carefully maintained accounts. Each has its positioning and audience. Many of them have really

interesting and unique content. Such accounts are few and far between and are used to throw in the information a customer needs, not to scale it.

6. People and other public platforms. When accusing bots of information spam, we should not forget about real people who have nothing to do with bots, but work in social networks in whose interests. These can be experts, journalists, public figures, or just people whose accounts have found mass popularity. For the money, other forms of remuneration, or simply by the call of the heart, they write about what is interesting to the customer.

Because of this variety of bot programs, the main problem in developing methods to detect them is creating an algorithm to distinguish real users who resemble bots by their actions from the bots themselves.

3. Pattern Recognition Methods for Classifying Malicious Bots

To solve this problem, it is necessary to develop an algorithm for detecting a malicious account that is based on the principle of pattern recognition. This is a scientific field related to the development of principles and the construction of systems designed to determine whether an object belongs to one of the classes of objects. By objects in pattern recognition understand the different objects and phenomena, processes and situations, signals, etc.

Recognition systems have the following typical functional scheme: the input data to be recognized are fed to the system and subjected to pre-processing for their further transformation into the form required for the next stage or the extraction of the necessary signature features [32, 37, 45]. The next step is the mathematical calculations for the processing of large amounts of data, as a result of which the necessary response is formed.

Classification is based on precedent. A precedent is an image whose correct classification is known—a previously classified object that is taken as a model for solving classification problems. The idea of making decisions based on precedent is fundamental to the natural science worldview. Let us assume that all objects or phenomena are divided into a finite number of classes. For each class, a finite number of objects—precedents – are known and learned. The task of pattern recognition is to assign a new recognized object to a class. Recognition systems also provide for their tuning to the set of possible input data: this stage is called the learning stage of the system. The purpose of training the system is to form in its memory a set of information (information images) needed to recognize the expected class of input data. Depending on the specifics of the task, training may be expressed as a procedure of one-time manual parameter setting by the system designer, an automatic procedure for determining the optimal parameter values as a result of recognition training cycles, or a process of continuous parameter adjustment as a result of analysis of the responses produced by the system. As a rule, there is a combination of these approaches [8, 28, 29, 30].

Based on the presence or absence of precedent information, a distinction is made between recognition tasks with training and without training. In the case where there is a set of feature vectors obtained for some set of images, but the correct classification of these images is unknown, the problem arises of dividing these images into classes by the similarity of the corresponding feature vectors.

Clustering and classification of malicious objects. Clustering is the grouping of objects or observations into non-overlapping groups, called clusters, based on the proximity of their attribute values. As a result, each cluster will contain objects similar in their attributes to each other and different from those in other clusters. In this case, the greater the similarity of objects within the cluster and the stronger their dissimilarity to objects in other clusters, the better the clustering. The formal formulation of the clustering problem is as follows. Let the sets of objects $X = (x_1, x_2, \dots, x_n)$ and cluster numbers (names, labels) $Y = (y_1, y_2, \dots, y_k)$ be given. For *X* some distance function *D (x, x')*, such as the metric *L2*, is defined. In addition, there is a finite sample of training examples $X_m = (x_1, x_2, \dots, x_m)$ from the set *X*, which needs to be divided into X_m non-overlapping subsets (clusters) so that each of them consists only of elements close in metric *D*. Each object x_i from the set X_m is assigned a cluster number y_j. Then the problem is to find the function *f* which assigns to any object xx from the set *X* the cluster number *y* from the set *Y*, which itself is known in advance. However, in most cases, we have to determine the optimal number of clusters based on the peculiarities of the problem to be solved.

The task of dividing a set of objects or observations into a priori specified groups, called classes, within each of which they are supposed to be similar to each other, has approximately the same properties and features. In this case, the solution is obtained on the basis of analysis of values of attributes (signs). If the number of classes is limited to two, this is a binary classification, in which many complex problems can be combined. Many models are used for classification: neural networks, decision trees, support vector machines, *k*–nearest neighbors' method, coverage algorithms, etc., at the construction of which the training with the teacher is applied, when the initial variable (class label) is given for each observation. Formally, classification is based on partitioning the feature space into regions, within each of which multidimensional vectors are treated as identical. In other words, if an object falls into a region of the space associated with a certain class, it belongs to it.

Existing methods for detecting malicious accounts on social networks and learning platforms

Social networks and learning platforms themselves do not provide active resistance to malicious accounts. For the most part, all bot protection actions are prevention:

- Captcha is a fully automated public Turing test, a computer test used to determine whether a system user is a human or a computer.
- SMS verification is verification of user authenticity by sending him a confirmation code by SMS to his number, which he must then enter in the appropriate field when registering or logging in to the system.
- Rate limit (bandwidth limitation)—limiting the number of requests to the system for a certain time.

In [6] and [14] the problem of identifying spambots is investigated using a popular social network as an example. It is proposed to distinguish normal users from malicious programs using machine learning classification. For this purpose, traditional classification methods are used: decision trees, neural networks, support vector machines, naive Bayesian classifiers. The number of users who signed up and read and the graph-oriented relationships of users were taken as features. These classifiers were trained and tested on a large dataset of information (45,600 accounts) and the most productive method for their evaluation was identified. To evaluate the performance we first create a matrix, and then we measure the precision (accuracy) values $P=a/(a+c)$ *recall* $R=a/(a+b)$, *F-measure* $F=a/(a + c)$. As a result, we obtain a summary table, the analysis of which makes it possible to determine that the Bayesian naive algorithm classifies spam bots most accurately (the highest *F*-value).

Work [7] describes the design of a framework (system) [1] for detecting bots in social networking and training platforms. It considers a general approach for all social networks, which is as follows:

1. Data collection. The first problem that arises when solving the problem of bot detection is the collection of information about users.
2. Identifying the attributes (or metrics) on the basis of which the classification algorithm will work.
3. A sampling of already classified data to train the classifier. That is, we need a sample of bot profiles and a sample of non-bot profiles. The classifier will be trained on these data. The larger the sample, the more accurate the classifier will work in the future.
4. Training the classifier on this sample.

In [13] the problem of distinguishing between bots and ordinary people is considered - the classification of Twitter accounts into human, bot, and cyborg. A cyborg is defined as the average between a bot and a human. A database of 670,000 social network Twitter accounts was collected. The difference is determined by the following characteristics: behavior, the content of tweets, and profile characteristics. Based on the results of the research, a classification system was proposed, consisting of four parts:

- Entropy component (entropy component)—analyses the interval between tweets (messages) of an account to determine its behavior.
- Spam detection component (spam detection component)—analyses the content of tweets according to predefined templates to determine whether the content belongs to spam content.
- Profile analysis component (account properties component)—analyses attributes related to the profile.
- Classification component (decision marker component)—uses the results of the previous three components to classify an account using one of the classification algorithms.

Profile analysis components.

When analyzing the profile, the author uses eight account characteristics. Let us consider each of them:

- A number of unreduced tweets—bots use more unreduced links in their tweets.
- From which device the tweets are created from websites or APIs.
- A metric called account reputation. That is the ratio of readers to those who are read. Humans either have the same number of readers and those who are read, or more readers than those who are read. Bots, on the contrary, have more readers than readers.
- Users' Twitter contains links to blacklisted resources. There are services that can detect suspicious sites. For example, GoogleSafeBrowsing, PhishingTank, URIBL, SURBL, Spamhaus.

These services provide an API to use their features in their programs:

- Availability of verification from Twitter. This mostly concerns accounts of famous people.
- The ratio of the number of hashtags to the total number of tweets.
- The number of reposts from a given account.

The following account characteristics are also used:

- Number of readers and reads.
- Education, job.
- Gender.
- The number of characters in "About Me."
- Family status.
- The number of photos in which the user is tagged.
- The number of posts in the account.
- The number of uploaded photos.

4. Development of an Algorithm for Identifying Malicious Accounts

Consider the module algorithm for determining the optimal number of clusters. At the initial stage of the research all clusters in our case are single-element:

$$t \coloneqq 1;\ C_t \coloneqq \{x_1\},\{x_2\},\dots,\{x_n\};$$
$$R(\{x_i\},\{x_j\}) = p(x_i,x_j).$$

The next step is to select the initial value of the parameter δ.

$$P(\delta) \coloneqq \{(U,V)|U,V\} \in C_k, R(U,V) <= \delta$$

For all k = 2, ..., n (k- iteration number), if $P(\delta) = 0$, then increase δ so that $P(\delta) \neq 0$.

Find in $C_k - 1$ the two nearest clusters:

$$(U,V) \coloneqq argmin\, R(U,V)$$
$$(U,V) \in P(\delta)$$
$$R_k := R(U,V)$$

Entails clusters U and V, add merged cluster W = $U \cup V$:

$$C_k \coloneqq\ C_k - 1\ \cup \{\mathrm{W}\}\backslash\{\mathrm{UU,VV}\}$$

For all $S \in C_k$, calculate the distance R (W, S) by the Lance-Williams formula.

If $R(W,S)$) $<=\ \delta$ then P (δ) $\coloneqq$ P (δ) $\cup$ {(W, S)}

The rationale for the choice of clustering method to classify malicious accounts. Separating social network user accounts with a large number of unequal criteria into groups is difficult to represent by two degrees of affiliation of 0 or 1. It is more natural to use partial affiliation in the range from 0 to 1, which will allow users, whose characteristics are at the boundaries between several clusters to belong to them with different degrees. Therefore, the fuzzy clustering method—fuzzy c—averages—is chosen as the main method of partitioning user accounts into groups. Thus, the initial information for clustering malicious accounts is the observation matrix X, $m \times n$. Thus, we will get a matrix of observations with input information for detailed analysis and comparison with the base record, each row of which n are the values of features m of one of the S objects of clustering.

$$X = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1m} \\ S_{21} & S_{22} & \cdots & S_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ S_{n1} & S_{n2} & \cdots & S_{nm} \end{bmatrix}$$

where *m* is the number of accounts to analyse, *n* is the number of attributes of a particular account.

The task of clustering is to divide the set of objects into groups (clusters) of "similar" objects. The distance between two objects (the Euclidean distance) is considered as a degree of "similarity" in n-dimensional metric space of attributes. The number of clusters with is defined in the previous step of the detection algorithm. The cluster structure is given by the membership matrix Z ($i \times j$):

$$Z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1j} \\ z_{21} & z_{22} & \cdots & z_{2j} \\ \cdots & \cdots & \cdots & \cdots \\ z_{i1} & z_{i2} & \cdots & z_{ij} \end{bmatrix}$$

where Z_{ij} is the degree of affiliation, i is the degree of belonging of j cluster elements.

Note that the membership matrix must satisfy the following conditions:

$$z_{ij} \in [0,1], i = \overrightarrow{1,2,\dots,n},\ j = \overrightarrow{1,2,\dots,m}.$$

$$\sum_{i=1}^{n} z_{ij} = 1, j = \overrightarrow{1,2,\dots,n}.$$

Thus, each object must be distributed among all clusters.

$$0 < \sum_{j=1}^{m} z_{ij} < m, i = \overrightarrow{1,2,\dots,m}$$

Thus, no cluster should be empty or contain all elements. To assess the quality of the partitioning, a scatter criterion is used, which shows the sum of distances from objects to the centers of clusters with the corresponding degrees of affiliation:

$$J = \sum_{i=1}^{n}\sum_{j=1}^{m} \left(z_{ij}\right)^{w} d(v_i, v_j)$$

where $d(v_i, v_j)$ is the Euclidean distance between the j object $x_j = (x_{j1}, x_{j2}, \dots, x_{jn})$ and i cluster center $v_i = (v_{i1}, v_{i2}, \dots, v_{in})$, $w \in (1, \infty)$ is an exponential weight, which determines fuzzy or blurred clusters.

$$V = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1m} \\ v_{21} & v_{22} & \cdots & v_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ v_{n1} & v_{n2} & \cdots & v_{nm} \end{bmatrix}$$

where V ($n \times m$) is a matrix of coordinates of the cluster centers, the elements of which are calculated by the formula:

$$v_{ik} = \frac{\sum_{i=1}^{n} \left(z_{ij}\right)^{w} x_{jk}}{\sum_{j=1}^{m} \left(z_{ij}\right)^{w}}, k = \overrightarrow{1,2,\dots,n}(v).$$

The task is to find the matrix Z, which minimizes the criterion j. In the study for this purpose we use the algorithm of fuzzy c-means, which is based on the Lagrange multipliers method. It allows us to find the local optimum, so different results can be obtained for different runs.

When using the fuzzy c-means algorithm, in the first step the membership matrix MM, which satisfies the conditions, is generated randomly, and then an iterative process of calculating the cluster centers is run to mathematically recalculate the elements of the membership degrees.

$$z_{ij} = \frac{1}{(d_{ij})^{\frac{2}{w-1}} \sum_{k=1}^{n} \frac{1}{\left(d_{kj}\right)^{\frac{2}{w-1}}}}$$

$$d_{ij}{>}0;\ z_{ij} = \begin{cases} 1, k = i \\ 1, k \neq i \end{cases}\!\!;\ d_{ij} = 0, d_{ij} = d\left(v_i, v_j\right), i = \overrightarrow{1, 2, \dots, n},\ j = \overrightarrow{1, 2, \dots, m}.$$

So the computation continues until the change of matrix Z, which is characterized by the value $||M - M^*||^2$, where $||M - M^*||^2$ is the matrix at the previous iteration, is less than the predetermined stopping parameter ε. The convergence of the fuzzy c-means algorithm is proved, so let us stop at the choice of the value w exponential weight. The more this value is, the more the membership matrix is smeared and with $w \rightarrow \infty$ elements will look like $z_{ij} = \frac{1}{c}$, which is a bad solution, because all objects with the same degree are distributed throughout the cluster. Based on the statement that there is no theoretically valid rule for choosing the weight so far, so as a rule it is set equal to two (w = 2).

Identifying clusters with suspicious malicious accounts. As a result of the clustering, we have formed with clusters. The degree to which accounts belong to each cluster is reflected in the resulting membership matrix.

To identify clusters with suspicious accounts, a sample of reference users must be formed, which includes:

- Accounts with maximum characteristics of a real person—must have maximum profile fields that have passed all necessary checks and are not seen in spamming, etc.
- Accounts with signs of malware - minimal information about the user, have been blocked, etc.

The next step is to check whether all the reference accounts belong to the clusters obtained. Depending on the result, the clusters can be categorized into:

- Loyal, with accounts of real users.
- Suspicious—those that belong to accounts with signs of malware.
- Clusters that do not contain any of the reference users.

Algorithm of the management decision-making module.

The module inputs the membership matrix and the cluster classification obtained in the previous step.

The decision algorithm consists of the following steps:

1. For clusters in which none of the accounts from the benchmark sample belongs, a group of users with the maximum membership scores (located in the center of the cluster) should be selected. For such accounts an additional check is offered. Based on the results, a decision is made whether such information objects can be considered harmful accounts or not. The verified users are entered into the reference sample.
2. Checking accounts located on the edge of clusters:

- At the first step, accounts with membership coefficients close to 0.5 are determined, as well as the numbers of clusters, from the center of which they are equidistant (on the border of which they are located).
- The next step is to determine the time interval T, during which the system tracks the trajectory of the disputed objects—the change in the membership coefficients in the matrix. Also, the maximum distance of the object from the cluster center (affiliation limit), within which it is not disputable, is set.
- If the object being checked is located on the boundary of two non-suspicious clusters, go to the next step.
- If the object is located on the boundary of two suspicious clusters, it is disregarded.
- If the object is on the boundary between suspicious and “loyal” clusters, then the function of tracking its motion vector for T is started. For this purpose, after each execution of the algorithm, its indices of belonging to adjacent clusters in the matrix are compared.

If the index of objects belonging to the "loyal" cluster increases (the motion vector is directed towards the cluster that is not suspicious) and the value of the affiliation boundary is smaller, then the observation continues. And if the value of the affiliation index exceeds the value of the affiliation boundary, it ceases to be controversial. If the affiliation index of an object to a suspicious cluster increases (the motion vector is directed towards the suspicious cluster) then it is clearly suspicious.

5. Conclusions

This paper proposes a topical solution for cybersecurity of social networks or learning platforms, concerning the development of a new method for detecting malicious accounts (bots). The following analytical and practical results were obtained in the course of solving the task:

1. Analysis of modern methods of detection using different approaches of pattern recognition, mathematical models. A comparison of the results of testing the existing methods, which showed that most of them work with a certain amount of data, characterized by significant errors, need complex mathematical calculations, lack of variability in decision-making. In this regard, there was a need to develop a new algorithm for detecting malicious accounts, which can reduce the number of unnecessary checks on users, for this analysis of clustering methods, considering their advantages and disadvantages, in order to choose the most effective for solving the problems posed in the thesis.
2. Formulated the formal requirements for the choice of detection methods, defined the criteria by which to determine the degree of affiliation of an account in the social bots, and proposed a method for improving the quality of sorting of accounts, followed by analysis of the results and decision-making on individual groups of users.
3. A combined method for tracking the behavior of suspicious accounts in the network is proposed, which can significantly reduce the number of additional checks, as well as an algorithm for detecting malicious accounts based on the methods chosen to solve the problems;

From a practical point of view, the proposed algorithm can be used to integrate into the existing system for detecting malicious accounts in social networks to improve data analysis and recognition, and for a final conclusion on the effectiveness of the developed algorithm, a thorough testing is required.

6. References

[1] A. M. Vegni, V. Loscri, A. Benslimane, SOLVER: A Framework for the Integration of Online Social Networks with Vehicular Social Networks, in: IEEE Network 34(1), 2020, pp. 204-213, doi: 10.1109/MNET.001.1900259.

[2] A. U. Hassan, et al., Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression, in: 2017 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, 2017, pp. 138-140, doi: 10.1109/ICTC.2017.8190959.

[3] D. Ageyev, et al., Infocommunication Networks Design with Self-Similar Traffic, in: IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM), Polyana, Ukraine, 2019, pp. 24-27, doi: 10.1109/CADSM.2019.8779314.

[4] Z. Hu, et al., Development and Operation Analysis of Spectrum Monitoring Subsystem 2.4–2.5 GHz Range, Lecture Notes on Data Engineering and Communications Technologies, 2020, pp. 675–709. doi: 10.1007/978-3-030-43070-2_29

[5] D. Kumar Srivastava, B. Roychoudhury, H. Vardhan Samalia, Importance of User's Profile Attributes in: Identity Matching Across Multiple Online Social Networking Sites, in: 2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, 2018, pp. 14-15, doi: 10.1109/CONFLUENCE.2018.8442455.

[6] I. Dronyuk, M. Nazarkevych, O. Fedevych, Synthesis of Noise-Like Signal Based on Ateb-Functions, in: Distributed Computer and Communication Networks. DCCN 2015. Communications in Computer and Information Science 601 (2016), doi: 10.1007/978-3-319-30843-2_14.

[7] A. Dronyuk, I. Moiseienko, J. Gregus, Analysis of Creative Industries Activities in European Union Countries, in: International Workshop on Digitalization and Servitization within Factory-Free Economy (D&SwFFE 2019), 2019, pp. 479–484.

[8] E. Awad, et l., Pareto optimality and strategy-proofness in group argument evaluation, Journal of Logic and Computation 27(8) (2017) 2581–2609.

[9] G. Suryateja, S. Palani, Survey on efficient community detection in social networks, in: 2017 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, 2017, pp. 93-97, doi: 10.1109/ISS1.2017.8389304.

[10] R. Galkin, et al., Approaches for Safety-Critical Embedded Systems and Telecommunication Systems Design for Avionics Based on FPGA, in: 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), Ukraine, 2019, pp. 391-396, doi: 10.1109/PICST47496.2019.9061421.

[11] J. K. Patel, A. Sakadasariya, Survey on virtual reality in social network, in: 2018 2nd International Conference on Inventive Systems and Control (ICISC), Coimbatore, 2018, pp. 1341-1344, doi: 10.1109/ICISC.2018.8399026.

[12] J. Galarza, B. Sáenz, E. Mucha, Work in Progress: Academic and Technical Support for Improving Education in Engineering, 2019 IEEE World Conference on Engineering Education (EDUNINE), Lima, Peru, 2019, pp. 1-2, doi: 10.1109/EDUNINE.2019.8875785.

[13] K. Hameed, N. Rahman, Today’s social network sites: An analysis of emerging security risks and their counter measures, in: 2017 International Conference on Communication Technologies (ComTech), Rawalpindi, 2017, pp. 143-148, doi: 10.1109/COMTECH.2017.8065764.

[14] K. Kansara, B. Kadhiwala, Non-cryptographic Approaches for Collaborative Social Network Data Publishing - A Survey, in: 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2020, pp. 348-351, doi: 10.1109/I-SMAC49090.2020.9243431.

[15] K. P. Arjun, et al., PROvacy: Protecting image privacy in social networking sites using reversible data hiding, in: 2016 10th International Conference on Intelligent Systems and Control (ISCO), Coimbatore, 2016, pp. 1-4, doi: 10.1109/ISCO.2016.7726913.

[16] K. Roy, S. Singh, S. Ratra, Social-Network-Sites (SNS) & Its Impact on Students’ Academic Learning, 2018 IEEE Tenth International Conference on Technology for Education (T4E), Chennai, 2018, pp. 174-177, doi: 10.1109/T4E.2018.00045.

[17] L. Ezzouine, A. Amine, M. Oubrich, State of the Art and Trends of the Research on Social Media Use in Organization: Bibliometric Analysis from 2007-2017, 2018 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD), Marrakech, Morocco, 2018, pp. 119-124, doi: 10.1109/ITMC.2018.8691176.

[18] L. Minaeva, The Use of Social Networking Platform Twitter by Russian and British Politicians: Comparative Analysis, 2020 IEEE Communication Strategies in Digital Society Seminar (ComSDS), St. Petersburg, Russia, 2020, pp. 162-165, doi: 10.1109/ComSDS49898.2020.9101285.

[19] L. Parabhoi, N. Kumari, Awareness and Use of Academic Social Networking Sites by Faculty and Students of Indian Institute of Technology (Indian School of Mines), Dhanbad : A Case Study, 2018 5th International Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS), Noida, 2018, pp. 174-178, doi: 10.1109/ETTLIS.2018.8485201.

[20] M. Nakerekanti, V. B. Narasimha, Analysis on Malware Issues in Online Social Networking Sites (SNS), 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India, 2019, pp. 335-338, doi: 10.1109/ICACCS.2019.8728536.

[21] M. Pasyeka, et al., System Analysis of Caching Requests on Network Computing Nodes, in: 2019 3rd International Conference on Advanced Information and Communications Technologies (AICT), Lviv, Ukraine, 2019, pp. 216-222, doi: 10.1109/AIACT.2019.8847909.

[22] M. Nazarkevych, et al., Ateb-Gabor Filtering Simulation for Biometric Protection Systems. CPITS 2020, pp. 14-22.

[23] M. Medykovskyy, et al. Scientific research of life cycle perfomance of information technology. 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies, 2017, pp. 425-428. doi: 10.1109/STC-CSIT.2017.8098821.

[24] O. Mishchuk, R. Tkachenko, I. Izonin Missing Data Imputation through SGTM Neural-Like Structure for Environmental Monitoring Tasks. Advances in Intelligent Systems and Computing 938, 2020, pp. 142-151, doi: 10.1007/978-3-030-16621-2_13.

[25] M. A. J. Idrissi, et al., Genetic algorithm for neural network architecture optimization, 3 International Conference on Logistics Operations Management (GOL), 2016 pp. 1-4.

[26] H. Mykhailyshyn, et al., Designing network computing systems for intensive processing of information flows of data, 2021, doi: 10.1007/978-3-030-43070-2_18.

[27] N. Pasieka, et al., Models, Methods and Algorithms of Web System Architecture Optimization, in: IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), Kyiv, Ukraine, 2019, pp. 147-153, doi: 10.1109/PICST47496.2019.9061539.

[28] M. Nazarkevych, et al., Identification of biometric images using latent elements, in: CEUR Workshop Proceedings 2488, 2019, pp. 99-108.

[29] M. Nazarkevych, et al., The Ateb-Gabor Filter for Fingerprinting, in: International Conference on Computer Science and Information Technology, 2019, pp. 247-255.

[30] M. Nazarkevych, Complexity Evaluation of the Ateb-Gabor Filtration Algorithm in Biometric Security Systems, in: 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2019, pp. 961-964.

[31] M. Nazarkevych, et al., Biometric identification system with ateb-gabor filtering. Paper presented at the 2019 11th International Scientific and Practical Conference on Electronics and Information Technologies, ELIT, 2019, pp. 15-18, doi:10.1109/ELIT.2019.8892282.

[32] O. Riznyk, et al., Recovery schemes for distributed computing based on bib-schemes, in: First International Conference on Data Stream Mining & Processing (DSMP), 2016, pp.134-137.

[33] P. Chatzoglou, et al., Generation Z: Factors affecting the use of Social Networking Sites (SNSs), in: 2020 15th International Workshop on Semantic and Social Media Adaptation and Personalization (SMA, Zakynthos, Greece, 2020, pp. 1-6, doi: 10.1109/SMAP49528.2020. 9248473.

[34] P. Ravi, et al., Authentication Protocol for Secure Automotive Systems: Benchmarking Post-Quantum Cryptography, in: 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Sevilla, 2020, pp. 1-5, doi: 10.1109/ISCAS45731.2020.9180847.

[35] N. Pasieka, et al. Models, methods and algorithms of web system architecture optimization, in: 2019 IEEE International Scientific-Practical Conference: Problems of Infocommunications Science and Technology, 2019, pp. 147-152, doi: 10.1109/PICST47496.2019.9061539.

[36] M. Pasyeka, et al., System analysis of caching requests on network computing nodes, in: 3rd International Conference on Advanced Information and Communications Technologies, AICT2019, 2019, pp. 216-222, doi:10.1109/AIACT.2019.8847909.

[37] M. Pasyeka, T. Sviridova, I. Kozak, Mathematical model of adaptive knowledge testing, in: 5th International Conference on Perspective Technologies and Methods in MEMS Design, MEMSTECH, 2009, pp. 96-97.

[38] M. Pasyeka, et al., System analysis of caching requests on network computing nodes, in: 2019 3rd International Conference on Advanced Information and Communications Technologies, AICT, 2019, pp. 216-222. doi:10.1109/AIACT.2019.8847909.

[39] O. Riznyk, et al., Synthesis of non-equidistant location of sensors in sensor network, in: 14th International Conference on Perspective Technologies and Methods in MEMS Design, MEMSTECH, 2018, pp. 204-208. doi:10.1109/MEMSTECH.2018.8365734.

[40] S. AL-Kharji, Y. Tian, M. Al-Rodhaan, A Novel (K, X)-isomorphism Method for Protecting Privacy in Weighted Social Network, in: 2018 21st Saudi Computer Society National Computer Conference (NCC), Riyadh, 2018, pp. 1-6, doi: 10.1109/NCG.2018.8593107.

[41] S. M. Bhalerao, M. Dalal, Improved social network aided personalized spam filtering approach using RBF neural network, in: 2017 International Conference on Intelligent Computing and Control (I2C2), Coimbatore, 2017, pp. 1-5, doi: 10.1109/I2C2.2017.8321938.

[42] S. Rane, M. Ainapurkar, A. Wadekar, Detection of compromised accounts in online social network, in: 2018 Second International Conference on Computing Methodologies and Communication (ICCMC), Erode, 2018, pp. 612-614, doi: 10.1109/ICCMC.2018.8487486.

[43] L. Sikora, et al., Technologies of development laser based system for measuring the concentration of contaminants for ecological monitoring, in: 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT, 2018, pp. 93-96. doi:10.1109/STC-CSIT.2018.8526602.

[44] R. Tkachenko, et al., An approach towards increasing prediction accuracy for the recovery of missing iot data based on the grnn-sgtm ensemble, Sensors 20(9) (2020), doi:10.3390/s20092625

[45] R. Tkachenko, et al., Development of the non-iterative supervised learning predictor based on the ito decomposition and sgtm neural-like structure for managing medical insurance costs, Data 3(4) (2018), doi:10.3390/data3040046.

[46] V. Jagadishwari, S. Chakrabarty, Link Prediction using Influencer nodes of a Social Network, in: 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2020, pp. 927-930, doi: 10.1109/ICIRCA48905.2020.9183315.

[47] W. Delu, Enterprise Network Marketing Strategy Based on SNS Social Network, in: 2019 12th International Conference on Intelligent Computation Technology and Automation (ICICTA), Xiangtan, China, 2019, pp. 295-299, doi: 10.1109/ICICTA49267.2019.00069.

[48] Y. Romanyshyn, et al., Social-communication web technologies in the higher education as means of knowledge transfer, in: IEEE 2019 14th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), 2019, pp. 35–39.

[49] M. Zharikova, V. Sherstjuk, Academic integrity support system for educational institution, in: 2017 IEEE 1st Ukraine Conference on Electrical and Computer Engineering, UKRCON, 2017, pp. 1212-1215, doi: 10.1109/UKRCON.2017.8100445.