

METHOD AND COMPUTING DEVICE FOR IDENTIFYING SUSPICIOUS USERS IN MESSAGE EXCHANGE SYSTEMS

FIELD

5 This technique relates to the field of information security, in particular to a method and computing device for identifying suspicious users in messaging systems.

BACKGROUND

10 Currently, the most popular means of communication on the Internet, corporate and local networks are various instant messaging systems (also called instant messengers or IMS systems), which allow users to organize a real-time or near-real-time message exchange, using any user computing device, such as a desktop computer, laptop, netbook, tablet, communicator, cell phone, smartphone, etc., on which the necessary client software (also called IM-client) is installed and which has configured to
15 connect to the appropriate information network. The most popular messengers in the world are, in particular, Viber, WhatsApp, Facebook Messenger, Telegram, Slack, Skype, Signal, ICQ, VideoChat and other popular messengers. In addition, offline communication systems are still popular communication tools, examples of which are electronic mail systems, forums, guest books on a standalone-web resource, Internet
20 blogs, etc.

 It shall be noted that each of the above messengers usually has separate servers and protocols, and also differs from other messengers by its rules and features, wherein there is no direct connection between different messengers, which, however, does not prevent any user from simultaneously using several different instant messengers for
25 communication purposes.

 Malicious users often use these popular instant messengers to send victim users messages comprising information about fake email addresses , fake account information in financial institutions or payment systems and / or malicious links that allow malicious users to deliver files with malicious content to the computing device of the victim users,
30 to install malicious software on the computers of the victims users, to redirect victim users to a fake website, etc. in order to further carry out various malicious activities, such as spamming, cyber blackmail, phishing, anonymous access to the network, theft of confidential data, etc., depending on the goals of these malicious users.

In order to identify and block malicious users based on the results of the analysis of their messages sent to instant messengers, the so-called intelligent chat bots embedded in such instant messengers or connected to them are widely used.

One of the illustrative examples of such an intelligent chat bot is described in
5 Russian patent No. RU2670906 C1 (publ. on 25.10.2018; H04L 29/00).

In particular, patent RU 2670906 describes a self-adjusting interactive message and / or call exchange system between users of different websites using client-server technology, comprising: an access module that provides the user access to the specified system if it is successfully authenticated by at least one social network and / or email
10 authentication; a user device associated with the access module and configured to provide navigation to at least one website, enter messages and / or make calls designed for users of the websites; a user profile reliability assessment module associated with a user device and adapted to calculate a user rating reflecting the level of trust in messages that users exchange; a message and / or call exchange module between users
15 of at least one web site associated with a user profile reliability assessment module and configured to exchange messages and / or calls between the users of at least one web site associated with a user profile reliability assessment module and configured to organize the exchange of messages and / or calls in a group of users depending at least on the user's rating and the possibility of automatic blocking of the possibility of
20 exchanging personal and / or group messages and / or calls for at least one user depending at least on the user's rating; a data protection module that checks any content which users exchange using a messaging module and configured to check comments and data messages for their belonging to the data sent by bots, by estimating the frequency of the user sending comments and / or messages based on
25 the IP address of these users to ensure users are blocked if the frequency of sending messages and / or comments exceeds a specified threshold for a given period of time, and configured to perform an anti-virus check of content in the said system.

It shall be noted that the protection module of the messenger users from the
30 actions of intruders in the messaging system and / or calls between users, described in the patent RU 2670906, uses an extremely limited set of tools to detect and block suspicious users, and also allows identifying malicious users only within a single messenger, that is, does not use any means or mechanisms to identify possible

accomplices of these malicious users who may carry out similar malicious activities, not only in other messengers, but also in the same messenger.

Thus, there is an obvious need for further improvement of protection against targeted attacks, for the organization of which the attackers use messages with malicious content sent to the users, for example, in various instant messengers, in particular, to more effectively detect malicious users in instant messengers and groups of interconnected malicious users who perform their malicious actions in the same instant messenger or in several different instant messengers.

Consequently, the technical problem solved by this technique is to create improved means for detecting suspicious users in the messengers, in which the above-mentioned disadvantage of known remedies is at least partially eliminated, which consists in the low effectiveness of identifying malicious users in messengers and in the impossibility of identifying a group of intruders carrying out similar malicious activities both in the same messenger and in several different instant messengers.

SUMMARY

The said technical problem is solved in one aspect of this technique, wherein a method for identifying suspicious users in messaging systems is claimed, performed on a computing device, and according to this method: receiving multiple user messages from at least one messaging system; analyzing each of the received user messages to identify at least one sign of suspiciousness of a message from a given set of signs of suspiciousness of messages; if at least one sign of suspiciousness of the message in the analyzed user message, identifying at least one user, associated with the analyzed user message, in the specified messaging systems; assigning each of the identified users an assessment of suspiciousness depending on the identified signs of suspiciousness of the message, each contributing to the assigned assessment of suspiciousness, ensuring the change in the reputation indicator of the specified user based on the assigned assessment of suspiciousness; attributing users under at least one of the specified messaging systems to suspicious users in the event of the changed value of their reputation score exceeds a predetermined limit of reputation threshold.

In one of the embodiments of this technique, the claimed method may include an additional operation, wherein the description format of the received user messages is identified, if the identified message description format does not conform to the unified

data description format suitable for computing devices, then the received messages will be converted into the specified unified format .

In another embodiment of this technique, the claimed method may include an additional operation, wherein voice messages and videos are detected in said received user messages and converted into text.

In some embodiments of this technique, when analyzing each of the received user messages in the claimed method, at least one of the following sub-steps is performed, wherein: at least one malicious link is detected in the analyzed user message; in the analyzed user message at least one malicious account of the payment system is detected; in the analyzed user message at least one malicious email address is detected; in the analyzed user message information about at least one malicious account in a financial organization is revealed.

In other embodiments of this technique, if at least one malicious link is detected in the analyzed user message in the claimed method, at least one link is extracted from the analyzed user message, ensuring at least one of the following operations, wherein: the domain name is analyzed for the extracted references for harmfulness using at least one analysis technique; at least one file located on the extracted link is obtained, followed by its analysis for maliciousness using at least one analysis method; and the html-code of the web resource located on the extracted link is received, followed by its analysis for harmfulness using at least one analysis technique.

In some other embodiments of this technique, when detecting at least one malicious link in the analyzed user message in the claimed method, it is further established whether the extracted link coincides at least partially with one of the known malicious links.

In other embodiments of this technique, when analyzing a domain name for harmfulness, it is additionally checked whether the analyzed domain name coincides at least partially with one of the known malicious domain names.

In some other embodiments of this technique, when analyzing the file obtained by the extracted reference, the hash sum is additionally calculated for harmfulness and it is determined whether the calculated hash sum of the analyzed file matches the hash sum of one of the known malicious files.

In one of the embodiments of this technique, when analyzing the obtained html-code of a web resource, search is conducted in the specified html-code for specified keywords indicating the harmful nature of the web resource.

5 In another embodiment of this technique, the sub-step of detecting at least one malicious payment system account in the analyzed user message includes extracting from the analyzed user message at least one payment system account with the provision of verification, whether the extracted payment system account matches one of the known malicious payment system accounts.

10 In yet another embodiment of this technique, the sub-step of detecting at least one malicious email address in the analyzed user message involves extracting at least one email address from the analyzed user message, ensuring that the extracted email address matches one of the known malicious e-mail addresses.

15 In some other embodiment of this technique, the identification data of the identified users is additionally analyzed in order to detect at least one sign of user suspicion from a given set of signs of user suspicion with the provision of allocation of an additional assessment of suspicion to each of the indicated users identified depending on the detected signs of suspicion of the message, each making its own assigned contribution to the assigned additional assessment of suspicion, with the provision of changes in the specified user's reputation based on the assigned additional suspicion
20 rating.

According to one of the embodiments of this technique, additional analysis of user identification data is carried out with respect to only those identified users whose reputation exceeds the specified minimum reputation threshold, which is less than the specified limit of reputation threshold.

25 According to another embodiment of this technique, when identifying at least one suspiciousness sign of each of the identified users: it is determined whether the avatar, full name and / or nickname of the identified user match the corresponding information about the administrator in at least one of the messaging systems; it is detected whether the identified user is a program that imitates the user's behavior in the
30 corresponding messaging system; it is detected whether the persons invited by the identified user to exchange messages within at least one message exchange channel in the corresponding messaging system are a program that simulates a person's messaging activity in the specified messaging system, and / or it is identified whether there is in

any messaging systems at least one user with credentials matching the credentials of the identified user.

According to another embodiment of this technique, when identifying whether the identified user is a program that simulates the user's behavior in the corresponding
5 messaging system, the activity of sending messages by the specified user is analyzed in the specified messaging system for a specified period of time, wherein if detecting that the activity of sending messages of the user corresponds to a certain time scheme, the specified user is classified as one of the programs simulating user behavior.

According to some other embodiment of this technique, when identifying at least
10 one user associated with a suspicious user message in messaging systems, messages that are similar to the specified suspicious user message are detected in these messaging systems, ensuring the association of users who sent such similar messages to a group of related users, in this case, when the reputation indicator of at least one of the users in the specified group exceeds the specified limit of reputation threshold, all users from the
15 specified group are classified as suspicious users.

According to yet another embodiment of this technique, when identifying at least one user associated with a suspicious user message in messaging systems, users with similar identification data are detected in these messaging systems by ensuring that such users are combined into a group of related users; wherein if the reputation indicator
20 of at least one of the users in the specified group exceeds the specified limit of reputation threshold, all users from this group are classified as suspicious users.

According to other embodiments of this technique, when identifying at least one user associated with a suspicious user message in messaging systems, a possible link is established between the references extracted from the messages being analyzed in
25 messaging systems with other references extracted from the messages being analyzed in these messaging systems, with the provision of the integration of such users into a group of interconnected users when establishing such communication, and wherein if the reputation indicator of one of the users in the specified group exceeds a predetermined limit of reputation threshold, , all users from this group are classified as suspicious
30 users.

According to some embodiments of this technique, when establishing a connection between references extracted from analyzed messages in messaging systems, at least one of the following is established for each pair of references: whether the

domain names have a similar spelling; whether the domain names are registered to the same person; whether the same contact information is provided for registered domain names; whether the domain names are located at the same IP address; and whether the extracted references have the same single web resource index (“URL”, for example,
5 <https://ru.wikipedia.org/wiki/URL>).

According to some other embodiments of this technique, the claimed method further sends information about each of the suspicious users to the appropriate messaging system.

10 In one of the embodiments of this technique in the claimed method, an additional request is sent to block each of the suspicious users to the corresponding data exchange system.

In another embodiment of this technique, in the claimed method, the analyzed user messages is additionally stored in the database of user messages located on the computing device.

15 In another embodiment of this technique, in the claimed method, for at least one of the stored user messages, a request is sent to the appropriate messaging system associated with the specified stored user message to establish whether the user has changed the user message in the specified messaging system, wherein if it is established that a user message has changed, the specified user message is received from the
20 specified messaging system with provision of its re-analysis for signs of suspicion.

In some other embodiments of this technique, the request is sent to the messaging system for at least one stored user message associated with a user whose reputation has exceeded the minimum reputation threshold.

In some embodiments of this technique, the access to at least one of the stored
25 user messages is received, via a computing device, to the corresponding messaging system, ensuring that a user message is retrieved from its message base, the attributes of which match the attributes of the specified user message, and the hash sum is calculate of the specified saved user message and the extracted user message, in this case in the event of a mismatch of the calculated hash-sums, the repeated analysis of the specified
30 modified user message is carried out for signs of suspicion.

In other embodiments of this technique, access to the message base of the messaging system, extracting the necessary user message from it, is obtained for at least

one saved user message associated with a user whose reputation has exceeded the minimum reputation threshold.

In some embodiments of this technique, in the claimed method, identified suspicious users in the respective data exchange systems are additionally blocked by means of a computing device.

The said technical problem is also solved in another aspect of this technique, wherein a computing device for detecting suspicious users in messaging systems is claimed, which is capable of communicating with messaging systems and comprises a memory for storing machine-readable instructions and at least one computing processor configured to execute machine-readable instructions with the implementation of the above described method of detecting suspicious users in messaging systems.

The objectives and advantages of the claimed method and computing device for identifying suspicious users in messaging systems will be apparent from the description below, attached drawings and the attached claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Accompanying drawings, which are provided for a better understanding of the essence of this technique, are included in this document to illustrate the following embodiments of this technique. Accompanying drawings, in conjunction with the description below, serve to explain the essence of this technique. In the drawings:

Fig. 1 shows schematically a system for detecting suspicious users in messaging systems;

Fig. 2 illustrates one implementation of a computing device for detecting suspicious users in messaging systems;

Fig. 3 shows a flowchart of a method for detecting suspicious users in messaging systems according to this technique.

DETAILED DESCRIPTION

Some examples of possible embodiments of this technique are described below, and it shall not be assumed that the following description defines or limits the scope of this technique.

Messaging systems (messengers) are usually implemented on the basis of a central computer or a management server, which, among other things, is connected to a

message database and a user database with the provision of managing the update of these databases and controlling access to these databases. Depending on the specific implementation of the messaging system, its users can send and / or receive text messages, voice messages, images and video.

5 In order to start working in any messaging system, the user needs to install an IM client on his device and to receive an identifier that is unique within a particular messaging system. In messaging systems that require an account, depending on their implementation, a login is used as a unique identifier, which can be an email address (e-mail), a set of numbers and / or characters (a nickname), an integer (for example,
10 “UIN” in the “ICQ” messenger or “user_id” in the “Telegram” messenger) or any combination of them. In the same messaging systems that do not require such an account (for example, blogs that allow anonymous commenting of published user posts), the IP address of the user's device can be used as a unique user ID, a unique “fingerprint” of the user type user-agent ", the MAC address of the user's device, the"
15 IMEI "of the mobile device, etc.

Each messaging system has one or more messaging channels (also called chats) used by users to exchange messages with each other.

The messaging channel in various messaging systems may include the following:

- 20 - web chat running on a standalone web resource, for example, on the chat.ru web resource;
- a separate so-called channel of “Telegram” messenger, available on both mobile devices and personal computers;
- text dialogue of at least two users in instant messengers such as “ICQ”,
25 “Skype”, “WhatsApp”, etc., available both on mobile devices and on personal computers;
- voice dialogue (accompanied by video and without it) in messengers like “Skype”, “WhatsApp”, etc .;
- exchange of images, text and / or video in the system of virtual or augmented
30 reality in instant messengers like “VideoChat”, in which each of the interlocutors must use special equipment, for example, VR helmets or VR glasses;

- any other text messaging channel used on the Internet, equipped with its own API-interface, for example, e-mail, a forum or a guest book on a standalone web resource, comments to a blog entry, etc.

Each messaging system has its own external connection interface (API), through which any external device that is connected to the specified API and knows the syntax of its commands can interact with the managing server on the basis of which this particular messaging system is implemented, for sending to this control server, for example, of requests to perform this server control actions and / or requests to send the entire message flow to the specified external device from all users through all messaging channels within this particular messaging system, including all additional information that attributes such user messages. Thus, the API of each messaging system is bidirectional, that is, it provides not only the transfer of requests from external devices to the management server, but also the transfer of necessary information from the management server to such external devices in response to requests received from them.

The user message transmitted via any of the above channels in the messaging system, each having its own unique logical address under which it can be accessed, is essentially formed from two main parts, namely, from the message body, which is that content (text, images, sound recordings, videos, files, etc.) which one user wants to transfer to another user within the established messaging channel in the messaging system, and the message header, which is the information used by the messaging system to describe the transmitted user message, including the date and time when it was sent, the sender's identification data, the recipient's identification data, the messaging system identification data, text attributes such as font, size, leading, etc. Thus, all user messages have the same structure: the body of each user message comprises data (content) that must be transmitted in accordance with the information placed in the header of the user message.

Terminology

In the context of this document, unless expressly stated otherwise, the “reputation indicator” is a number, such as an integer, and this number is one of the required database fields or databases in which user data is stored, in particular the local data storage described below. When a new record associated with a new user is created

in such a database or databases, the reputation indicator of this new user is assigned a predetermined numerical value, for example, a zero value. In case of detection of malicious content, for example, a malicious link, in a message sent by a specific user in the messaging system, the value of this user's reputation indicator is changed, in particular, the value of this user's reputation indicator is increased by the value calculated by the method described below, preserving in the corresponding database or databases the specified changed value of the reputation indicator.

In the context of this document, unless expressly stated otherwise, the “minimum reputation threshold” is the value of the reputation indicator, the excess of which for a particular user leads to the attribution of this user to potentially suspicious users (i.e., among the users suspected of spreading malicious messaging systems) for which the additional checks described below can be carried out, the negative results of which can lead to increase of reputation indicator of the specified user.

In the context of this document, unless expressly stated otherwise, the “limit reputation threshold” is the value of the reputation indicator, the excess of which for a particular user leads to the attribution of this user to suspicious users, that is, to the users who spread malicious messages in messaging systems followed by at least temporarily blocking this user in at least one of the messaging systems.

System for detection of suspicious users in messaging systems

Fig. 1 schematically shows the system 100 for identifying suspicious users, comprising the messaging system 110, corresponding, for example, to the WhatsApp messenger, the messaging system 120, corresponding, for example, to the Telegram messenger, and the computing device 200 for the detection of suspicious users according to this technique, shown on Fig. 2. The messaging system 110 includes, among other things, a control server 112, an API-interface 114 that provides configured to interact with the control server 112, a message base 116 and a user data base 118 managed by the control server 112, the same time, all messages transmitted by the users in the messaging system 110 are stored in the message database 116, and all users who use the messaging system 110 to communicate with other users of this messaging system 110 are stored in the user database 118. The messaging system 120 also includes, among other things, a control server 122, an API-interface 124 that provides configured to interact with the control server 122, a message base 126 and a user data

base 128 managed by the control server 122, wherein all messages transmitted by the users in the messaging system 120 are stored in the message database 126, and the information on all users who use the messaging system 120 for communication with other users of this messaging system 120 is stored in the user database 128. It shall be noted that each of the user messages in the message database 116 and each of the user messages in the message database 126 are assigned to at least one unique identifier corresponding to one of the users, and information about users in the user database 118 and information about users in the user database 128 also associated with at least the unique identifier of one of the relevant users. In addition, each of the messaging systems 110, 120 may additionally comprise its own database of suspicious users (not shown).

Thus, system 100 shown on Fig. 1 is designed to detect suspicious users in the messaging system 110 and the messaging system 120.

A computing device 200 for detecting suspicious users according to this technique is connected to the messaging system 110 and the messaging system 120 via a parser 130, respectively, connected to the API-interface 114 of the messaging system 110 and preconfigured to work with it, and the parser 140 connected to API-interface 124 of the messaging system 120 and preconfigured to work with it, wherein the parser 130 is connected configured to exchange messages with the computing device 200 using the Internet network 150, and the parser 140 is connected with the possibility of exchanging messages directly with the computing device 200 itself. It shall be noted that each of the API-interface 114 and the API-interface 124 has its own specific command syntax, so the parser 130, working with the API-interface 114, must be preprogrammed to understand the command syntax of this API-interface 114, and the parser 140, working with API-interface 124, shall be pre-programmed to understand the syntax of the API-interface 124 commands in the same way, wherein setting of the parser 130 and the parser 140 to work respectively with the API-interface 114 and the API-interface 124 is carried out when the initial connection of the computing unit 200 to the messages exchange systems 110, 120 is made.

The computing device 200 is configured to send requests to each of the messaging system 110 and the messaging system 120, using the parsers 130, 140 that are preconfigured to work respectively with the API-interfaces 114, 124, through which the directed requests are received respectively in systems 110, 120, these requests will comprise commands that are understood respectively by the control servers 112, 122,

after which these control servers 112, 122 will be able to properly handle the specified received requests and adequately respond to them.

It shall be noted that the messaging system 110 and the messaging system 120 are shown on Fig. 1 solely as an example, i.e. it shall not be considered that the possible embodiment of system 100 is limited to the example shown on Fig. 1, it shall be clear to those skilled in the art that system 100 can comprise both a single messaging system, corresponding, for example, to any one of the known messengers, and more than two messaging systems, each corresponding to, for example, any one of known messengers, and the interaction between any such messaging system and computing device 200 will occur as described below in the interaction processes between computing device 200 and messaging systems 110, 120

Any of the parsers, including any of the above parsers 130, 140 used in system 100 to detect suspicious users, can be connected to the computing device 200 indirectly, that is, using a wired or wireless connection, for example, using the Internet 150 network, and directly, that is, without using, for example, the Internet network.

According to this technique, computing device 200 is configured to send requests to each of messaging system 110 and messaging system 120 to receive user messages from them sent by users through all messaging channels within the respective messaging system, wherein in response to accepted requests servers 112, 122 receive access respectively to the messages database 116 and the messages database 126 with the provision of extracting the requested user messages from them and sending the extracted user messages, respectively, through the API-interfaces 114, 124 to the parsers 130, 140, wherein the parser 130 provides configured to process the output data stream from the API-interface 114 to extract user messages from it, requested from the messaging system 110, and the parser 140 provides configured to process the output data stream from the API-interface 124 to retrieve from it user messages requested by the messaging system 120. In particular, the output data stream from any API-interface comprises the text of the user messages themselves and the identification data describing the transmitted user messages, for example, the date and time of submission, the sender's identification data, the recipient's identification data, the messaging system identification data and / or other necessary attributes of these transmitted messages. The output data stream from any API-interface is usually a set of characters as a string of characters with a specified description format, which is divided into structural elements,

which can be implemented, in particular, using a predefined symbol, such as the "#" character (grid) , wherein the description format of such a string of characters from the API-interface of a particular messaging system shall be known to the parser, which is configured to work with this API-interface. As an example, the output data stream from the API-interface can be a character string of the following form: ...message_id: id736353672383#from: @user1#date: 11.10.2018 15:155#text: Hello world. <https://phishingwebsite.com/login#>. Upon receipt of such character strings from API-interfaces 114, 124, parsers 130, 140, respectively, extract from these received strings, each divided into a known set of structural elements, the user message text, and at least some of the above described identification data, describing these user messages, with the provision of transfer of the specified information to computing device 200 for further processing, the features of which are described below. For the above example of a character string that can be retrieved by any of the parsers 130, 140 from the corresponding API-interface, the text of the user message that will be retrieved by the specified parser will look, for example, as follows: "Hello world. <https://phishingwebsite.com/login> ", wherein in the specified parser the extracted text "Hello world. <https://phishingwebsite.com/login> " will be associated with at least some of the extracted identification data describing this user message, and transmitted to computing device 200. In particular, it shall also be noted that in order to correctly extract all necessary information, in particular the text of the user message and identification data describing this user message, from the output data streams from API-interfaces 114, 124, the parsers 130, 140, respectively, shall be preprogrammed or configured to handle the output data streams, presented in the form of character strings, respectively, from the API-interfaces 114, 124, wherein each of the parsers 130, 140 shall at least know recording format of the processed data streams. In particular, for the example above, the parsers 130, 140 shall be aware that the keyword in the resulting character string pointing to the beginning of the text of the user message is the word "text". It shall also be noted that in order to process the output data stream from the corresponding API-interface in the parser, it could additionally (in addition to the extracted text of the user message, as described earlier in this document) retrieve all the necessary identification data describing the extracted user message, for example, date and time of sending, identification data of the sender, identification data of the recipient, identification data of the messaging system itself and / or other identifiers (that is, all

the information that is usually comprised in the headers of user messages transmitted within a particular messaging system, and which allows to describe the body of these user messages), the specified parser shall also know other standard keywords commonly used in the received strings of characters to indicate the presence of certain identifying information following this keyword in these character strings.

In one of the embodiments of this technique, each of the parsers 130, 140 may be further programmed or configured to recognize or identify the description format of the output data stream, respectively, from API-interfaces 114, 124. In addition, each of the parsers 130, 140 is pre-programmed in such a way that it knows the unified data description format with which the computing device 200 operates, wherein the text format of data description is used as the unified format. In this embodiment of the this technique, the system 100 may additionally comprise a transforming module (not shown) connected configured to exchange data to the parser 130, and another conversion module (not shown) connected configured to exchange data to the parser 140, each of these converting modules are designed to convert voice / sound into text (for example, using known temporal dynamic algorithms, known methods of context-dependent classification based on the extraction from the speech of individual lexical elements, such as phonemes and allophones, with their subsequent integration into syllables and morphemes, known methods of discriminant analysis based on Bayesian discrimination, methods based on hidden Markov models, methods based on neural networks, etc.) converting a video / video stream into text (for example, using the “Virtual Audio Cable” software product, which is one of the known transcription tools), configured to recognize text in a video / video stream, such as a subtitle in (for example, using the ABBY Real-time Recognition SDK software), configured to recognize objects in a video / video-stream, ensuring their conversion into text (for example, using the cloud service Cloud Video Intelligence IP) and / or configured to recognize actions in a video / video-stream ensuring its conversion into text (for example, using the same cloud service “Cloud Video Intelligence IP”).

In one of the embodiments of this technique, at least each of the conversion modules may comprise several conversion submodules, each configured to implement at least one of the above-described functionality. In yet another embodiment of this technique, each of the above-described conversion modules can be divided into several separate modules, each configured to implement at least one of the above-described

functionality. In the case when any of the parsers 130, 140 revealed that the output data stream from the corresponding API-interface is described using a format different from the unified data description format suitable for the computing device 200, this parser issues this output data stream to the appropriate external converting module for converting it from the identified format to the specified unified data description format using at least one conversion method known to it, for which this converting module was pre-programmed, wherein the selection of said conversion method depends on the identified data description format. Thus, if any of the parsers 130, 140 revealed that the output data stream from the corresponding API-interface is essentially a voice / audio message and / or video / video stream, this parser sends them to the appropriate external module for converting the specified voice / audio message and / or video / video stream to text. Further, the parser, which revealed the difference between the format of the description of the output data stream of the API-interface and the unified format of data description, accepts the text extracted in the corresponding conversion module (in this case such extracted text received from the conversion module is essentially a text message), and starts the above process of extracting structural elements from this text. In case when any of the parsers 130, 140 revealed that the output data stream from the corresponding API-interface is described using a format that matches the unified data description format, that is, it is presented in a textual data description format, this parser immediately starts to execute the above process of extracting structural elements from this output data stream.

In some embodiments of this technique, the system 100 may additionally comprise a separate user message base that is external to the computing device 200 to detect suspicious users, and the parsers 130, 140 can be configured to access this user message base with a possibility of recording an information to it, extracted from the output data streams, respectively, from the API-interfaces 114, 124, respectively, as a result of which this user message base comprises user message text associated with supporting identification data describing these user messages, such as the date and time the user message was sent, the sender ID, the recipient ID, the messaging system ID, and / or at least one other identifier. Computing device 200 is configured to obtain access to such a database of user messages with the possibility of extracting the necessary information from it for subsequent processing, the features of which are described below.

In other embodiments of this technique, the system 100 for identifying suspicious users may comprise only of the computing device 200 for identifying suspicious users and a structured database of user messages that is external to the computing device 200. In this embodiment, the user message database comprises the user message text complying with supporting identification data describing these user messages, for example, the date and time the user message was sent, the sender ID, the recipient ID, the messaging system ID, and so on, and computing device 200 for identification of suspicious users is configured to gain access to such a database user messages with the possibility of extracting from it the necessary information for further processing, the features of which are described below.

In one of the additional embodiments of this technique, the computing device 200 for detecting suspicious users may be configured to connect directly to the messaging systems 110, 120 with direct access to their message bases 116, 126 to extract user messages from them for their post-processing, the features of which are described below.

In one of the embodiments of this technique, in response to a request for receiving user messages sent by computing device 200 to detect suspicious users to any of the messaging systems 110, 120, this messaging system issues only those user messages to the computing device 200 as described above, which were saved in the message database 116 or 126 for a given period of time, for example, all recent user messages starting from the time of sending the said request by the computing device 200 (the reading of the system clock of the computing device 200 comprised in the request itself) or from the moment the request was received by the corresponding management server via the API-interface (reading of the system clock in this particular messaging system), or all user messages starting from user messages saved in the last few minutes, hours, days, weeks, months, etc. depending on the tasks.

Computing device for detection of suspicious users in messaging systems

The computing device 200 for detection of suspicious users according to this technique, shown on Fig. 2 is designed to detect suspicious users in messaging systems 110, 120 and is essentially a hardware-software complex implemented as a general-purpose computer, having the structure described below, well known to those skilled in the art.

In particular, a general-purpose computer usually comprises a central processor, system memory, and a system bus, which in turn comprises various system components, including memory associated with the central processor. A system bus in such a general-purpose computer comprises a memory bus and a memory bus controller, a peripheral bus and a local bus, configured with a possibility of interaction with any other bus architecture. System memory comprises read-only memory (ROM) and random access memory (RAM). The Basic Input / Output System (BIOS) comprises basic procedures that ensure the transfer of information between the elements of such a general-purpose computer, for example, at the time of loading the operating system using the ROM. In addition, a general purpose computer comprises a hard disk for reading and recording data, a magnetic disk drive for reading and recording to removable magnetic disks, and an optical drive for reading and recording on removable optical disks such as CD-ROM, DVD-ROM and other optical storage media, but other types of computer storage media can be used to store data in machine-readable form, such as solid-state drives, flash cards, digital disks, etc., and connected to the system bus via controller. At a general-purpose computer, a hard disk, a magnetic disk drive and an optical drive are connected to the system bus via a hard disk interface, a magnetic disk interface and an optical drive interface, respectively. Drives and associated computer storage media are non-volatile means of storing computer instructions, data structures, program modules and other general-purpose computer data. A general purpose computer has a file system that stores a recorded operating system, as well as additional software applications, other software modules and program data. The user can enter commands and information into a general-purpose computer using known input devices, such as a keyboard, mouse, microphone, joystick, game console, scanner, etc., wherein these input devices are usually connected to a general-purpose computer via a serial port, which is in turn connected to the system bus, but they can also be connected in some other way, for example, using a parallel port, a game port, or a universal serial USB bus. A monitor or other type of display device is also connected to the system bus via an interface, such as a video adapter. In addition to the monitor, a personal computer may be equipped with other peripheral output devices, such as speakers, a printer, etc. A general purpose computer can work in a network environment, and a network connection

can be used to connect to one or more remote computers. Network connections can form a local area network (LAN) and wide area network (WAN). Such networks are usually used in corporate computer networks and internal networks of companies, wherein they have access to the Internet. In a LAN or WAN network, a general purpose computer is connected to the local network via a network adapter or network interface. When using networks, a general purpose computer may use a modem, network card, adapter or other means of providing connection with a global computer network, such as the Internet, and these means of communication are connected to the system bus via a serial port. It shall be noted that in the ROM of the general purpose computer or at least in any of the above computer-readable media that can be used in a general-purpose computer, the computer-readable instructions can be stored which can be accessed by the CPU of the general-purpose computer, wherein execution of these machine-readable instructions on a general-purpose computer may cause the central processor to execute various procedures or operations described later in this document.

In one of the embodiments of this technique, the computing device 200 may be implemented as a single computer server, such as a Dell TM PowerEdge TM server using the Ubuntu Server 18.04 operating system. Besides, in other embodiments of this technique, the computing device 200 may be presented in the form of a desktop personal computer, laptop, netbook, smartphone, tablet, and other electronic computing device suitable for solving the set tasks.

In other embodiments, the computing device 200 may be manufactured in the form of any other combination of hardware, software or software and hardware complex, suitable for solving tasks.

In some embodiments of this technique, the system 100 for identifying suspicious users may comprise at least two computing devices, similar to computing device 200, and the functionality described below for computing device 200 can be in any appropriate way divided between specified at least two computing devices, each of them, for example, can be configured as a separate computer server.

The computing device 200 shown on Fig. 2, has a communication module 210, an analysis module 220 and a local data storage 230, each connected to a communication bus 240, wherein each of the communication module 210 and the analyzing module 220 being capable of exchanging data via the communication bus 240

with a local data storage 230, and the communication module 210 is also configured to exchange data with the analyzing module 220.

In one embodiment of the this technique, the above-described parsers 130, 140 can each be configured as a separate preprocessing module embedded in the computing device 200 to detect suspicious users (i.e., included in this computing device 200) and having the above-described functional capabilities of one of the corresponding parsers 130, 140, in particular the functionality for providing interaction between the computing device 200 of one of the corresponding messaging systems 110, 120 through one of the corresponding API-interfaces 114, 124 (i.e., each of these separate data preprocessing modules shall be preprogrammed to work with one of the corresponding API-interfaces 114, 124) and processing output data streams from one of the corresponding API interfaces 114, 124. In one form of this embodiment, the communication module 210 of the computing device 200 may be configured, for example, as multi-channel, with each of the communication channels in such a communication module 210 being pre-configured to exchange data via the communication bus 240 with one of the above described separate modules of data pre-processing and data exchange with one of the corresponding API-interfaces 114, 124. In another version of this embodiment, computing device 200 may be equipped with two communication modules, similar to each communication module 210, each of these communication modules being pre-configured to exchange data via communication bus 240 with one of the above described separate data processing modules and to exchange data with one of the corresponding API-interfaces 114, 124. In this embodiment, the separate data preprocessing modules (not shown) are also each configured to interact, via the communication bus 240, with the analyzing module 220 to process the corresponding requests (described later in this document), which can be generated by this analyzing module 220, followed by their direction from the computing device 200 to one of the corresponding messaging systems 110, 120. It shall also be noted that when processing the output data streams from one of the corresponding API-interfaces 114, 124, each of these separate preprocessing data modules (not shown) can, among other things, identify or recognize the description format of the output data stream, respectively received from the specified API-interface, and if the identified data description format does not conform to the unified data description format relevant to the computing device 200, then this module of data preprocessing may be additionally configured to

convert this received output data stream into the specified unified format, each of these data preprocessing modules may be additionally configured to communicate, via the communication bus 240, with the local data storage 230 to provide receipt from it of data on the unified data description format (as described below), understandable to the computing device 200, and configured to compare the specified identified and unified data formats for a decision on their conformity or non-conformity with each other. Thus, if any of the above described separate data preprocessing modules reveals that, among the output data streams received from one of the corresponding API-interfaces 114, 124, there are, for example, voice / audio messages and / or video / video streams, then such separate data preprocessing module converts the specified voice / audio messages or video / video streams into text using at least one of the known conversion methods (see the description above), that is, to such data description format that is understood by computing device 200. In one version of this embodiment of this technique, the functionality of each of the above-described preprocessing modules for converting voice / sound messages or video / video streams into text can be implemented in a separate conversion module embedded in the computing device 200 and connected configured to exchange data with one of the corresponding data preprocessing modules.

In another embodiment of the this technique, the above-described parsers 130, 140 can be implemented as a single data preprocessing module (not shown) embedded in the computing device 200 to detect suspicious users (i.e., included in the computing device 200), and having the above-described functionality of both parsers 130, 140, in particular the functionality for providing interaction between computing device 200 and messaging systems 110, 120 respectively, through their API-interfaces 114, 124 (i.e., such a single data preprocessing module shall be preprogrammed to work with both API-interfaces 114, 124) and for processing output data streams from API-interfaces 114, 124. In this embodiment, a single data processing module (not shown) shall also be connected in the computing device 200 to the communication bus 240 to enable data exchange with the communication module 210, which provides communication between the API-interfaces 114, 124 and the computing device 200, wherein this communication module 210 of the computing device 200 in this case can be configured, for example, multichannel, and each of the communication channels in this communication module 210 can be pre-configured to exchange data with one of the

corresponding API-interfaces 114, 124. In this embodiment, a single data preprocessing module (not shown) is also adapted to interact, via the communication bus 240, with the analyzing module 220 to process the above requests (described earlier in this document), which can be generated by this analyzing module 220, with their subsequent
5 direction from the computing device 200 to the above messaging systems 110, 120. It shall also be noted that when processing output data streams from API-interfaces 114, 124, a single data preprocessing module (not shown) can, among other things, identify or recognize the description format of output data streams received from API-interfaces 114, 124, wherein, if the identified data description format does not match the unified
10 data description format appropriate for the computing device 200, then it may additionally be configured to convert these received output data streams to the specified unified format, wherein this single data preprocessing module can be additionally configured to communicate, via the communication bus 240, with the local data storage 230, ensuring that data about the unified data description format is received (as
15 described below), understandable to the computing device 200, and with the possibility of comparing the indicated identified and unified data formats for making a decision on their compliance or non-compliance with each other. Thus, if the above described single data preprocessing module reveals that, among the output data streams received from API-interfaces 114, 124, there are, for example, voice messages or video, then such a
20 single data preprocessing module converts the specified voice messages or video in the text, that is, in the format of the data description, which is understandable to the computing device 200.

Local data storage

25 The local data storage 230 is also designed to store executable program instructions that can control the operation of the communication module 210 and the analyzing module 220, as well as various data used in the operation of the computing device 200, in particular, data on a unified data description format understood by the computing device 200, data on the reputation threshold limit used in the computing
30 device 200 to make a decision about classifying the analyzed user as the suspicious user, data on known malicious links, data on known malicious domain names, data on hash sums of known malicious files, data on keywords indicating the malicious nature

of a web resource, data on known malicious accounts of payment systems, data on known malicious email addresses and minimum threshold reputation data.

In addition, auxiliary data used in the work of the analyzing module 220 can also be stored in the local data storage 230, for example, data on language dictionaries and a predetermined threshold value used in the method of analyzing domain names based on the correctness of their spelling; virtual machine image files and a set of rules for analyzing changes in virtual machine state parameters used in suspicious file analysis methods based on changes in virtual machine state parameters, a set of regular expressions used to extract user messages analyzed in analyzing module 220, references, accounts of the payment systems, email addresses and accounts at financial institutions, etc.

In the computing device 200 shown on Fig. 2, the communication module 210 is adapted to receive data transmitted by the parsers 130, 140 to the computing device 200 in response to the corresponding request of this computing device 200 sent to at least one of the messaging systems 110, 120, and then storing the received data in the local data storage 230, to which this received data can be transmitted via the communication bus 240. Thus, in the local data storage 230, user messages from the messaging systems 110, 120 can be stored, in particular, the auxiliary user identification data (such as avatars, full name and / or nickname) extracted from their message bases 116, 126 from the messaging systems 110, 120, in particular from their user databases 118, 128, and the auxiliary identification data of each of the administrators (such as, for example, avatar, full name and / or nickname) of messaging systems 110, 120, in particular, retrieved from their user databases 118, 128.

In some embodiments of this technique, the local data storage 220 of the computing device 200 may comprise one or more databases, each configured to store at least one separate group of the above listed groups of data.

In other embodiments of the computing device 200 at least one separate remote data storage can be used (not shown), to which the analyzing module 220 of the computing device 200 can gain access using the communication module 210, to store therein at least part of the above groups of data or all these groups of data, in particular for storing received user messages in a unified data description format.

In some other embodiments of this technique, computing device 200 may comprise at least one local data storage and at least one remote data storage (not

shown), each designed to store at least one of the above described data groups, wherein the local data storage each are connected to the analyzing module 220 via the communication bus 240, and remote data storages each are connected to the analyzing module 220 via the communication module 120. Thus, for example, an embodiment of this technique is possible, in which the computing device 200 comprises a single local data storage 220, for example, storing only received user messages in a unified data description format, and several remote data storages storing other data groups used in the operation of the computing device 200.

In one embodiments of this technique, at least one of the above groups of data, in particular, received user messages in a unified data description format, can be stored in a corresponding separate local data storage (not shown) different from the local data storage 230 and connected, via the communication bus 240, with the analyzing module 220, which in turn is configured to be connected to any of these separate local data storages with provision of extraction of the necessary data from them.

The analyzing module 220 may be implemented as a single processor, such as a general-purpose processor or a special-purpose processor (for example, processors for digital signal processing, specialized integrated circuits, etc.) and configured to execute software instructions stored in local data storage 230, ensuring the implementation of the following functionality of the analyzing module 220.

Local data storage 230 may be implemented, for example, in the form of one or more known physical computer-readable media for long-term data storage. In some embodiments of this technique, local data storage 230 may be implemented using a single physical device (for example, a single optical storage device, a magnetic storage device, an organic storage device, a storage device on disks, or a different type of storage device), and in other embodiments, local data storage 230 may be implemented using two or more known storage devices.

Communication module

The communication module 210 used in the computing device 200 shown on Fig. 2 has a wireless connection to the parser 130 and has a wired connection to the parser 140, as described earlier in this document, with the possibility of receiving data from them.

In one embodiment of this technique, the communication module 210 may be connected to both parsers 130, 140 in a wired manner, for example using a coaxial cable, twisted pair, fiber optic cable or other physical connection, with the possibility of receiving data from them. In this embodiment, the communication module 210 may be implemented as a network adapter equipped with the necessary connectors for connecting the necessary types of physical cables to them depending on the types of physical connections used to provide communication with the parsers 130, 140.

In another embodiment of this technique, the communication module 210 may be connected to both parsers 130, 140 wirelessly, for example using a communication line based on WiFi technology, a communication line based on 3G technology, a communication line based on LTE technology, etc. In this embodiment, the communication module 210 may be implemented as a network adapter in the form of a WiFi adapter, a 3G adapter, an LTE adapter, or another wireless communication adapter, depending on the type of wireless link used to provide communication with the parsers 130, 140.

Communication module 210 may also be a known communication device, such as a transmitter, receiver, transceiver, modem, and / or network interface card for exchanging data with external devices of any type via a wired or wireless communication network, for example, using an Ethernet network connection. , digital subscriber line (DSL), telephone line, coaxial cable, cellular telephone system, etc.

Analyzing module

The analyzing module 220 included in the computing device 200 shown on Fig. 2, may be implemented as a single processor, such as a general-purpose processor or a special-purpose processor (for example, processors for digital signal processing, specialized integrated circuits, etc.), for example, as a central processor of the above-described general-purpose computer, in the form of which computing device 200 may be implemented.

Analyzing module 220 is configured to access local data storage 230 (separate local data storage or remote data storage, depending on the embodiment, as described earlier in this document) or to communicate with it using communication bus 240 with provision of extraction from it of the user messages, the format of which corresponds to

the unified data description format, understandable to the computing device 200, for their subsequent analysis, as described below.

In one embodiment of this technique, the analyzing module 220 may be configured to communicate, via the communication bus 240, with the communication module 210, enabling it to receive user messages from it, the format of which corresponds to the unified data description format comprehensible to the computing device 200 for subsequent analysis as described below. Thus, in this embodiment, the analyzing module 220 may receive user messages in a unified data description format directly from the communication module 210 after receiving the specified user messages by the communication module 210.

In the embodiments of this technique, in which received user messages in a unified data description format are stored in a separate local storage other than the local data storage 230, or in a remote data storage, the analyzing module 220 may be configured to access such a separate or remote data storage or to communicate with it using the communication bus 240, ensuring that stored user messages are retrieved from it for their subsequent analysis as described below.

The analyzing module 220 is configured to analyze each of the extracted user messages to detect at least one sign of suspiciousness of that user message from a predetermined set of signs of suspicious messages.

When analyzing each of the extracted user messages, the analyzing module 220 performs at least one of the following operations: 1) detects in each analyzed user message at least one malicious link; 2) identifies at least one malicious payment system account in each analyzed user message; 3) detects in each analyzed user message at least one malicious email address; 4) in each analyzed user message, detects at least one malicious account in a financial institution or payment system.

In particular, to detect the presence of malicious references in each analyzed user message (the sign have their specific recording format) of the malicious payment system accounts (the accounts also have their specific recording format), malicious email addresses (email addresses also have their specific recording format) and / or malicious accounts in a financial institution (accounts also have their own specific recording format); the analyzing module 220 uses a given set of regular expressions known to it, allowing to derive from the user message all of the above information, necessary for the analyzing module 220 for further detection of suspicious signs of such a user message.

For example, a regular expression in the form «(https?|ftp)://(-\.)?([^\s/?\.\#-][^\s/?\.\#-]+)?» allows to extract from any user message any web address, a regular expression like «A[13][a-km-zA-HJ-NP-Z1-9]{25,34}\$» allows to extract from any user message any bitcoin wallet, and a regular expression in the form «^(\S+)@([a-z0-9-]+\.)?([a-z]{2,4})(\.)?([a-z]{0,4})+\$» allows to extract from any user message any email address.

In order to detect at least one malicious reference in any analyzed user message, the analyzing module 220 (i) extracts, by means of a corresponding regular expression from the set of regular expressions known to it, from this user message at least one reference; gains access to local data storage 230 (separate local data storage or remote data storage, depending on the embodiment, as described earlier in this document) or communicates with it using the communication bus 240 to ensure that hash sum data of the known malicious files is received from it, and (iii) establishes, by character-by-character comparison of the extracted reference with known malicious references from the indicated obtained data, the fact of at least partial coincidence of the extracted reference with one of the known malicious references. If the analyzing module 220 has determined that the extracted reference does not at least partially coincide with any of the known malicious references, then it additionally performs at least one of the following operations: 1) analyzes the domain name of the extracted reference for harmfulness using at least one domain name analysis technique known to it; 2) receives or downloads at least one file located at the extracted reference, followed by its analysis for ma using at least one file analysis method known to it; and / or 3) obtains the html-code of the web resource located at the extracted reference, followed by its analysis for maliciousness using at least one html-code analysis method known to it.

When analyzing a domain name of any extracted malicious reference, the analyzing module 220 (i) accesses or communicates with local data storage 230 (separate local data storage or remote data storage, depending on the embodiment, as described earlier in this document), using the communication bus 240 to ensure that data about known malicious domain names is obtained from it, (ii) establishes, by character-by-character comparison of the analyzed domain name with known malicious domain names from the specified obtained data, the fact of at least partial coincidence of the analyzed domain name with one of the known malicious domain names. If the analyzing module 220 has determined that the domain name being analyzed does not at

least partially coincide with any of the known malicious domain names, then it may additionally apply suspiciousness of the domain name analysis to such analyzed domain name, for example, a domain name analysis method based on its length (the longer the domain name is, the more suspicious it is), a method for analyzing a domain name based on its entropy (in this case, the higher the informational entropy calculated for a given domain name using the well-known Shannon formula, the more suspicious is the given domain name), a technique for a domain name based on its meaningfulness and / or analyzing technique of the domain name based on the correctness of its spelling. As an example, when the analyzing module 220 analyzes a domain name for harmfulness using a domain name analysis technique based on the correctness of its spelling, the analyzing module 220 performs at least the following operations: (i) establishes a connection with the local data storage 230 (separate local or remote data storage, depending on the embodiment) for obtaining data about language dictionaries from it, (ii) extracts at least one word from each of the obtained domain names n, (iii) determines the Levenshtein distance between each of the specified extracted words and the corresponding one of the words in the language dictionaries of the specified obtained data, and (iv) compares a certain Levenshtein distance with a specified threshold value, which, for example, can be used as equal to two (2), ensuring the assignment of the analyzed domain name to malicious domain names, if a certain Levenshtein distance exceeds a specified threshold value, for example, equal to two (2).

Thus, if the analyzing module 220 has established, by at least one of the above described analysis methods, that the domain name of the analyzed extracted reference belongs to malicious domain names, this indicates that the analyzing module 220 has detected a malicious link in the analyzed user message, which corresponds to the presence in this analyzed user message, of one corresponding sign of suspicion from a given set of signs of suspicion of messages, known to the analyzing module 220.

When analyzing a file located on the extracted reference, the analyzing module 220 (i) receives a file located at the extracted reference; (ii) calculates the hash sum of the resulting file; (iii) gains access to local data storage 230 (separate local data storage or remote data storage, depending on the embodiment, as described earlier in this document) or communicates with it using the communication bus 240 to ensure that hash sum data of the known malicious files is received from it; (iv) establishes, by comparing the calculated hash sum of the file with the hash sums of known malicious

files from the specified data, the fact that the calculated hash sum of the file matches one of the hash sums of known malicious files. If the analyzing module 220 has determined that the hash sum of the received file does not match any of the hash sums of known malicious files, then it may additionally apply at least one suspicious file analysis method to the received file, for example files for suspicion based on changes in the state parameters of virtual machines, wherein the analyzing module 220 performs at least the following operations: (i) runs each received file at least one virtual machine characterized by a given set of state parameters, (ii) records changes in a given set of state parameters of the specified at least one virtual machine over a specified period of time, (iii) analyzes the resulting changes in state parameters using a specified set of analysis rules ensuring that the specified launched file is assigned to malicious files, if the analyzed changes of state parameters are that of malicious files.

Thus, if the analyzing module 220 has established, through at least one of the above analysis methods, that the analyzed file located at the analyzed extracted link refers to malicious files, this also indicates that the analyzing module 220 has detected a malicious link in the analyzed user message, which corresponds to the presence in this analyzed user message of the corresponding one sign of suspicion from a given set of signs of suspicion known to the analyzing module 220.

When analyzing the html-code of a web resource located at the extracted reference for maliciousness, the analyzing module 220 performs at least the following operations: (i) loads the html-code of the web page located at the extracted link; (ii) analyzes downloaded html-code for maliciousness using at least one well-known analysis method, for example, methods based on keywords indicating the harmful nature of a web page, depending on the malicious content (corresponds to some specific malicious elements for example, a text inviting a user to perform an action, a file of a specific format, scripts, substituted logos, etc.) identified in the loaded html-code, pre-programmed classifier can classify or identify the type of threat posed by the loaded html-code, for example to identify whether the loaded html-code refers to a type of 'phishing' threats.

In addition, when analyzing downloaded html-code for maliciousness, analyzing module 220 can also download all images and / or other files associated with the loaded web page, for example, graphic design elements (* .JPG, * .PNG, etc.) , style sheets (* .css), JS scripts, etc., based on the lists of such images and / or other files obtained by

the analyzing module 220 from the extracted html-code, ensuring the verification of the so-called screen-signatures, that is, search for similar images and analyze related websites, wherein search for the similar images may be performed, e.g., using techniques similar to image search on the basis of a well-known method of searching for the nearest neighbors. During such a search, the analyzing module 220 determines whether, for example, the images placed on the analyzed web page correspond to the domain name and registration data of the web resource, wherein the analyzing module 220 can also additionally calculate the hash sums of all the images present on the analyzed a web page, and determine whether each calculated image hash sum matches one of the hash sums of known malicious elements that can be stored, for example, in the local data storage 230. In addition, the analyzing module 220 can additionally check the so-called resource signatures, for which it can calculate the hash sums of all previously loaded resources of the analyzed web page, for example, images, cascading style sheets (CSS), JS files, fonts, etc., and to establish whether each calculated hash sum of a resource matches one of the hash sums of known malicious resources that can be stored, for example, in local data storage 230.

Thus, if the analyzing module 220 has established, through at least one of the above methods of analysis, that the analyzed web page located at the analyzed extracted reference has malicious content, this also indicates that the analyzing module 220 has detected a malicious link in the analyzed user message that corresponds to the presence in this analyzed user message one of the corresponding signs of suspiciousness of a message from a given set of indications of suspiciousness of messages known to the analyzing module 220.

To detect at least one malicious payment system account in any analyzed user message, the analyzing module 220 (i) extracts at least one payment system account from the user message through the corresponding regular expression from the set of regular expressions known to it; (ii) gains access to local data storage 230 (separate local data storage or remote data storage, depending on the embodiment, as described earlier in this document) or communicates with it using communication bus 240 to receive the information about known malicious payment system accounts from it and (iii) establishes, by, for example, a character-by-character comparison of the extracted payment system account with known malicious payment system accounts from the

specified received data, the fact of coincidence of the extracted account of the payment system with one of the known malicious accounts of payment systems.

Thus, if the analyzing module 220, by means of, for example, the above-described method of analyzing payment system accounts, determines that the payment system account being analyzed is related to malicious payment system accounts, this fact corresponds to the presence of another corresponding sign of suspiciousness from the specified message in this analyzed user message from a set of signs of suspicious messages, known to the analyzing module 220.

In order to detect at least one malicious email address in any user message being analyzed, the analyzing module 220 (i) extracts, by means of a corresponding regular expression from the set of regular expressions known to it, from this user message at least one email address; (ii) gains access to local data storage 230 (separate local data storage or remote data storage, depending on the embodiment, as described earlier in this document) or communicates with it using communication bus 240 to receive the information about known malicious email addresses and (iii) establishes, for example, by character-by-character comparison of the extracted email address with known malicious email addresses from the specified received data, fact of matching of the extracted email address with one of the known malicious email addresses.

Thus, if the analyzing module 220 has determined, for example, by the above described method of analyzing email addresses, that the analyzed email address refers to malicious email addresses, then this fact corresponds to the presence in the analyzed user message of the corresponding yet another sign of suspiciousness from the specified a set of signs of suspicious messages, known to the analyzing module 220.

The analyzing module 220 is also configured to identify the user who sent the analyzed user message in the corresponding messaging system. In order to identify the user who sent a specific user message, the analyzing module 220 extracts from the identification data describing the specified user message and usually comprised in the header of such a user message, the ID of user who sent the user message, for example login, which can be specified email address (e-mail), a set of numbers and / or characters (nickname), an integer (for example, "UIN" in the messenger "ICQ" or "user_id" in the messenger "Telegram") or any combination thereof, or the user's device's IP address, a unique user-agent user's fingerprint, the user's device's MAC address, mobile device's IMEI, or any combination of them, and retrieves the exchange

system identifier message, in which the specified user message was sent, such as the IP address of the messaging server, and the extracted message system identifier uniquely identifies the messaging system used by the user to send this message, and the extracted user identifier uniquely identifies the specific user in the specified messaging system.

5 Thus, the analyzing module 220 can identify the user associated with the user message, for which the analyzing module 220 revealed that it has at least one sign of suspiciousness of the message, for example, at least one of the above described signs of suspiciousness of the message.

10 In addition, the analyzing module 220 is additionally configured to write data about the identified user, set in accordance with the unique user identifier and unique identifier of the messaging system, into the appropriate user database, which can be placed, for example, in the local data storage 230 (or in a separate local storage of user data to which the analyzing module 220 can gain access or with which it can communicate using the communication bus 240, or in a remote user data store, to which
15 the analyzing module 220 can gain access or with which it can communicate using the communication module 210 connected to the analyzing module 220 via the communication bus 240, depending on the embodiment of this technique), wherein each such identified user has his own specific reputation indicator, the value of which is initially set by the analyzing module 220 to zero and further adds of suspicion scores
20 assigned to this user by the analyzing module 220 as a result of analyzing his messages in the messaging system, and the data on the values of users' reputation indicators are stored in a specified user database in which they are assigned to a unique user identifier.

Analyzing module 220 is also configured to assign a grade of suspiciousness to each of the identified users, depending on the detected signs of suspiciousness of the
25 message, each making its own specified contribution to the assigned suspiciousness grade, ensuring that the specified user's reputation is changed based on the assigned suspiciousness grade, wherein the analyzing module 220 is configured to assign a given numerical value (for example, a score) to each of the identified indications of the suspiciousness of a message depending on, for example, the specific nature of the threat
30 corresponding to that particular sign of the suspiciousness of the message and / or the degree of its harmfulness (in particular, the nature of the threat and / or the degree of its harmfulness can be assessed by the analyzing module 220 using, for example, special pre-programmed classifier based on the results of the analysis of user messages for

harmfulness), so that the numerical value (grade) assigned to specific suspiciousness sign of a message determines the contribution of this suspiciousness characteristic of a message to the assessment of suspiciousness assigned to a specific user. Thus, an assessment of suspiciousness assigned to the identified user by the analyzing module 220 is obtained by summing the specified numerical values (grades) assigned to the suspiciousness signs of the message detected by the analyzing module 220 for that identified user. It shall also be noted that if the analyzing module 220 did not reveal any signs of suspicion in the analyzed user message, then the assessment of suspicion assigned to a particular user associated with this user message is zero, which does not result in change of value by the analyzing module 220 of this user's reputation indicator originally set to zero. Otherwise, as a result of analyzing a user message to a specific user associated with that user message being analyzed, a suspicious rating was assigned to a value other than zero, the analyzing module 220 (i) accesses or communicates with the local data storage 230 using communication bus 240 (either gets access to a separate local user data storage or a remote user data storage or communicates with them depending on the embodiment of this technique), (ii) finds in the user's database the current value of this user's reputation based on his unique identifier; and (iii) updates this current value of the reputation indicator by increasing it by the suspicion rating grade assigned to this user.

The analyzing module 220 may be further configured to perform the following operations, wherein it: (i) analyzes the identification data of each identified user to identify at least one suspicious user sign from a given set of suspicious user signs; and (ii) assigns an additional suspicion rating to each of the identified users, depending on the identified suspicious user signs, each contributing to the assigned additional suspiciousness rating, ensuring that the specified user's reputation score changes based on the assigned additional suspiciousness rating.

It shall be noted that when identifying at least one sign of the use suspiciousness for each identified user, the analyzing module 220 performs at least one of the following operations, wherein it: (i) determines whether the avatar, full name and / or nickname of the identified user are corresponding with the administrative support credentials in the corresponding one of messaging systems 110, 120; (ii) identifies whether the identified user is a program that simulates user behavior in one of the corresponding messaging systems 110, 120; (iii) identifies whether the persons invited by the identified user to

exchange messages within the established message exchange channel in one of the corresponding messaging systems 110, 120 are programs that simulate a person's messaging activity in the specified messaging system; and / or (iv) identifies whether there are users with the same identification data in the messaging systems 110, 120.

5 To establish the coincidence of the avatar, full name and / or nickname of the identified user with the corresponding auxiliary administrator identification data, the analyzing module 220 performs the following operations, wherein it: (i); sends a request to one of the appropriate messaging systems 110, 120 based on the unique identifier of the identified user and the unique identifier of the messaging system to obtain the
10 necessary supporting identification data of this identified user, in particular, avatar, full name and / or nickname of the identified user, from one of the corresponding databases 118, 128 ; (ii) sends another request to the specified messaging system, from which the necessary auxiliary identification data of the identified user was obtained, in order to obtain the necessary auxiliary identification information about the administrator of this
15 messaging system, in particular, the avatar, name and / or the nickname of the administrator, from its user database (it shall be noted that the messaging system administrator is in fact also one of its users; however, it has a wider scope equal to other users of this messaging system, in particular, by comparison with the identified user); and (iii) determines whether the avatar, full name and / or nickname of the identified
20 user are comprised in the received auxiliary identification data of the identified user, respectively, with the avatar, name and / or administrator nickname comprised in the obtained auxiliary identification data of the administrator, in this case , if at least one element of auxiliary identification data from the avatar, the full name and nickname of the identified user matches at least one corresponding element of the administrator
25 identification data, it indicates that the user is identified with a certain degree of probability simulates the behavior of an administrator in the messaging system.

 Thus, if the analyzing module 220 has determined that the avatar, full name and / or nickname of the identified user matches the avatar, full name and / or nickname of the administrator, respectively, then this fact indicates that the identified user is
30 characterized by one of the corresponding signs of the user suspiciousness from the specified set indications of user suspiciousness known to analyzing module 220.

 In order to determine whether the identified user is a program that simulates the user behavior, the analyzing module 220 analyzes the message sending activity of the

specified user in one of the corresponding messaging systems 110, 120 (in particular, in each messaging channel to which the identified user was connected within this messaging system) for a specified period of time based on the unique identifier of the identified user and the unique identifier of the messaging system, in this case, if the analyzing module 220 detects, by means of, for example, a pre-programmed classifier embedded in the analyzing module 220, that the activity of sending messages by a specified user in the specified messaging system corresponds to one of the specified timing schemes known to the analyzing module 220, this will result in the identification of this identified user as a program simulating the user's behavior in the messaging system (bot programs).

Thus, if the analyzing module 220 has determined that the identified user is a bot program, this fact also indicates that the identified user is characterized by another relevant indication of the user's suspiciousness from a given set of signs of suspicious messages known to the analyzing module 220.

If it identifies whether the persons invited by the identified user (in particular, such invited persons are also users who use the messaging system to communicate with other users of this messaging system) to take part in the messaging (that is, in the process of communication between users of the system by messaging, such as, for example, chatting, writing, multiuser chat, private chat or the like) within at least one messaging channel according to one of the messaging systems 110, 120, programs that simulate the user behavior, the analyzing module 220 (i) sends a request to one of the corresponding messaging systems 110, 120 (this request uses the syntax of the API command of this messaging system)) based on the unique identifier of the identified user and the unique identifier of the messaging system to obtain from it the necessary data about persons invited by the identified user; (ii) for each of the invited persons who were communicated by the specified messaging system to the analyzing module 220 in response to the specified request, proceeds to perform operations identical to the above described operations performed by the analyzing module 220 to determine if the identified user is a program that simulates user behavior (see the above description).

Thus, if the analyzing module 220 has determined that at least one messaging user invited by the identified user is a bot program, this fact also indicates that the identified user is characterized by another relevant sign of suspiciousness by a user from a given set of indications of suspicious messages known to analyzing module 220.

If it will be established, whether there is at least one user in the messaging system with identification data that matches the identification data of the identified user, the analyzing module 220 (i) sends requests (each request uses the syntax of the API command of one of the corresponding 110, 120 messaging systems) to the messaging systems 110, 120 based on the unique identifier of each of these messaging systems 110, 120 and a unique identifier of the identified user to identify in each of these messaging systems 110, 120 the users having the same identification data as the identified user.

Thus, if the analyzing module 220 has determined that there are users in the messaging systems 110, 120, at least one identity element of which matches the corresponding at least one identity element of the identified user, this fact also indicates that an identified user is characterized by another relevant sign of user suspiciousness from a given set of signs of suspicious messages, known to analyzing module 220.

The analyzing module 220 is also configured to assign a predetermined numerical value (for example, a score) to each of the identified signs of user suspiciousness depending on, for example, the number of elements of auxiliary identification data from among the avatar, full name and nickname of the identified user and administrator, the number of users that are the bot programs, among users of a particular messaging system, invited by an identified user, the number of identity elements, such as e-mail, that match the identified user and the users of all messaging systems, and the number of such users, etc. Thus, the numerical value (score) assigned to a specific user suspicion sign determines the contribution of this user suspicion sign to the additional grade of suspiciousness assigned to a specific user. Thus, an additional assessment of suspiciousness attributed to the identified user by the analyzing module 220 is obtained by summing the specified numerical values (grades) assigned to the suspiciousness signs of the user identified by the analyzing module 220 for that identified user. It shall also be noted that if the analyzing module 220 did not reveal any of the above user suspicious signs for the identified user, then the additional suspicious rating assigned to this particular user is zero, which does not change the reputation indicator value by the analyzing module 220 for this user. Otherwise, if the identified user was assigned, by the analyzing module 220, an additional grade of suspiciousness with a value other than zero, then this analyzing module 220 (i) accesses the local data storage 230 or communicates with it using the communication bus 240 (or gains access

to a separate local storage of user data or remote storage of user data, or establishes communication with them depending on the embodiment of this technique), (ii) finds in the user's database the current value of this user's reputation indicator based on its unique identifier; and (iii) updates this current value of the reputation indicator by increasing it by the value of the additional suspicion rating assigned to this user.

Thus, the value of each user's reputation indicator in each of the messaging systems 110, 120 is influenced not only by the assessment of suspiciousness assigned to this user by the analyzing module 220 based on the detected signs of suspiciousness of messages sent by this user, but also by the additional assessment of suspiciousness assigned to this user based on the identified suspicious signs of this user, as described earlier in this document.

In one of the embodiments of this technique, the analyzing module 220 may perform the above-described analysis of user identification data only for those identified users whose reputation exceeds the specified minimum reputation threshold that is less than the reputation threshold limit. Thus, if the assessment of the suspiciousness of the identified user, assigned based on the results of the above analysis of the user message sent by this user in the messaging system, the presence of signs of suspiciousness of the message led to a change in the reputation indicator of this user to a value that exceeds the minimum reputation threshold, the analyzing module 220 proceeds to perform the above described analysis of identification data of the specified user, wherein the analyzing module 220 obtains the set value of the reputation threshold limit and the set value of the minimum reputation threshold from the local data storage 230 (separate local storage of data on threshold values for messaging systems or remote storage of data on threshold values for messaging systems depending on the embodiment of this technique).

The analyzing module 220 is also configured to classify the identified users at least in one of the messaging systems 110, 120 as suspicious users in the case if the changed value of their reputation score exceeds set reputation threshold limit, retrieved from local data storage 230 (a separate local storage of data on reputation threshold values for messaging systems or a remote storage of data on reputation threshold values for messaging systems, depending on the embodiment of this technique).

The analyzing module 220 is also configured with the possibility to send information about each of the identified suspicious users to one of the corresponding

messaging systems 110, 120 based on the unique identifier of this suspicious user and the unique identifier of this messaging system associated with this suspicious user. Upon receipt of the sent information about a suspicious identified user, associated with at least the unique identifier of that suspicious user and the unique identifier of the corresponding messaging system, the managing server of that messaging system may, for example, store the received information in a separate database of suspicious users hosted in this messaging system, with the provision of prevention or prohibition, for example, to view any messages that suspicious user stored in the appropriate message base, to view any information about the suspicious user stored in the appropriate user database, to send any messages to this suspicious user, and to exchange messages with that suspicious user in this messaging system.

In one of the embodiments of this technique, the analyzing module 220 may be configured to send a request to block each identified suspicious user to the appropriate messaging system. Upon receipt of such a request that is made at least with the unique identifier of this suspicious user and the unique identifier of the corresponding messaging system, the managing server of this messaging system may, for example, prohibit, for example, viewing any messages of this suspicious user stored in the corresponding database. messages, viewing any suspicious user information stored in the relevant user database, sending any messages to this suspicious user and messaging with that suspicious user in this messaging system.

In one of the embodiments of this technique, if the analyzing module 220 has detected a suspicious user in one of the corresponding messaging systems 110, 120, then it can save information about such suspicious user, set in accordance with the unique identifier of that suspicious user and unique identifier of this messaging system, in a separate database of suspicious users, located, for example, in the local data storage 230 (a separate suspicious user data storage to which the analyzing module 220 can have an access or with which it can communicate using the communication bus 240, or in a separate remote storage of suspicious user data to which the analyzing module 220 can have an access or with which it can communicate using communication module 210 connected to analyzing module 220 via communication bus 240, depending on the embodiment of this technique). Thus, in this embodiment, when receiving a user message from any messaging system, in particular from any of the messaging systems 110, 120, the analyzing module 220 may first (that is, before performing the above

described analysis of this user message for maliciousness) identify the user who sent the received message, on the basis of this user's unique identifier comprised in this received message, and then check whether this identified user has previously been identified as suspicious user in any other messaging system that is part of system 100 in order to
5 detect suspicious users and interact with device 200 to detect suspicious users, by accessing or linking to the suspicious user database described above, as a result, the analyzing module 220 may send information about the identified suspicious user to this messaging system, from which the specified user message has been received, for subsequent blocking of this suspicious user in it (see the above description) or to send a
10 request to block this identified suspicious user to the specified messaging system (see the description above).

In another embodiment of this technique, the above-described database of suspicious users may be further interacted (via communication module 210 in conjunction with an appropriate parser or preprocessing module, depending on the
15 embodiment of this technique) with all messaging systems, in particular messaging systems 110, 120, configured to periodically update their databases of suspicious users, in particular, to amend database of suspicious users with the information about new suspicious users identified in any of the messaging systems that make up the system 100 to identify suspicious users and interacting with the device 200 to detect suspicious
20 users.

In another embodiment of this technique, the analyzing module 220 may interact (via the communication module 210 in combination with an appropriate parser or preprocessing module depending on the embodiment of this technique) with control servers of messaging systems, in particular with control servers 112, 122 of messaging
25 systems 110, 120, configured to control their work, so that in the above case of identifying a suspicious user, the analyzing module 220 can be configured to send control instructions to one of the appropriate control servers 112, 122 to perform the above blocking operation of this suspicious user in one of the corresponding messaging systems 110, 120 based on the unique identifier of the messaging system and the unique
30 identifier of the suspicious user.

In some embodiments of this technique, for at least one of the identified user messages from the messaging systems 110, 120 stored in the user message base in the local data storage 230, the analyzing module 220 can also be configured to submit to the

corresponding one of the systems 110, 120 messaging associated with the specified stored user message, a request to establish whether the user message was changed in the specified messaging system by the user himself, if, in response to the specified request, the specified messaging system determined that the specified user message was changed, the analyzing module 220 receives, through the communication module 210, this modified user message from the specified messaging system for reanalyzing for signs of suspicion, as described earlier in this document.

In other embodiments of this technique, the above-described request for determining whether a user message has changed can be sent by the analyzing module 220 to one of the corresponding messaging systems 110, 120 only under those user messages for which the analyzing module 220 has determined that the users who sent these user messages have each reputation value exceeding the minimum reputation threshold, wherein the information about the value of the minimum reputation threshold is received by the analyzing module 220 from the local data storage 230, and information about the current value of the user reputation indicator is received by module 220 from the user database, located in the local data storage 230, based on at least a unique user identifier.

According to one embodiment of this technique, for at least one of the identified user messages from messaging systems 110, 120 stored in the user message base in the local data storage 230, the analyzing module 220 can also be configured to gain access by computing device to one of the corresponding messaging systems 110, 120 based on at least unique messaging system identifier for retrieving from one of the corresponding user message bases 116, 126, the attributes of which match the attributes of the specified stored user message, and configured to calculate hash sums of the specified stored user message and the extracted user message, wherein in case of inconsistency of the calculated hash sums, the analyzing module 220 repeatedly performs the above analysis of harmfulness with respect to the specified modified user message. According to one of the embodiments of this technique, obtaining access to one of the corresponding message bases 116, 126 of messaging systems 110, 120, and then extracting the necessary user message from it (as described earlier in this document) can be performed by analyzing module 220 only for those user messages for which the analyzing module 220 has determined that the users who sent these user messages have each reputation indicator, exceeding the minimum reputation threshold,

while analyzing module 220 receives information about the minimum threshold reputation from the local data storage 230, and information about the current value of the user reputation indicator, the analyzing module 220 receives from the user database located in the local data storage 230 based on at least a unique user identifier.

5 In order to identify, in messaging systems, at least one user associated with a suspicious user message (i.e., associated with an analyzed user message, in which the analyzing module 220 has detected at least one suspicious sign as described earlier in this document), in the messaging systems included in the system 100 in order to identify suspicious users and interacting with the device 200 in order to identify suspicious
10 users, for example, in messaging systems 110, 120, the analyzing module 220 (i) receives from each of the messaging systems 110, 120 all user messages stored respectively in message bases 116, 126, each received user message being associated with a unique identifier of a particular messaging system and unique identifier of a specific user; (ii) compares a suspicious user message with each of the specified user
15 messages received using at least one known method for comparing text data, for example, by character-by-character comparison, comparison by keywords that can be extracted from each user message using, for example , a special pre-programmed classifier embedded in the analyzing module 220, and / or by comparing the hash sum calculated by the analyzer module 220 for a suspicious user message, with the hash sum
20 calculated by the analyzing module 220 for each of the user messages received, to detect messages similar to the specified suspicious user message in these messaging systems 110, 120, ensuring that users who send such similar messages are grouped into a group of related users; (iii) in the case when the reputation indicator of at least one of the users in the specified group exceeds the specified reputation threshold limit (as
25 described in more detail earlier in this document), all users from the specified group are classified as suspicious users. It shall be noted that, depending on the embodiment option, in order to obtain the necessary user messages, the analyzing module 220 can either send a corresponding request to each of the messaging systems 110, 120 to receive the necessary user messages in response to the sent request, or directly access
30 each of the message databases 116, 126 or establish a connection with them to extract the necessary user messages from there.

As a supplement or alternative for identifying in messaging systems at least one user associated with a suspicious user message (i.e. associated with the analyzed user

message in which the analyzing module 220 has detected at least one suspicious sign as described earlier in this document), in the messaging systems that make up the system 100 in order to identify suspicious users and to interact with the device 200 to identify suspicious users, for example, in messaging systems 110, 120, the analyzing module

5 220 (i) receives from each of the messaging systems 110, 120 all user information stored respectively in user bases 118, 128 and set each in accordance with the unique identifier of a specific messaging system and unique identifier of a specific user; (ii) compares the identification data of the suspicious user, in particular its unique identifier, with the corresponding identification data of each user (in particular, with his unique

10 identifier having the same type of possible types of user identifiers described above as that of the suspicious user), which is comprised in the specified obtained information about users, in order to identify in these messaging systems 110, 120 the users with the similar identification data; (iii) in the event that users with similar identification data are identified in messaging systems 110, 120, such users are united into a group of the

15 related users; (iv) in the case when the reputation indicator of at least one of the users in the specified group exceeds the specified reputation threshold limit (as described in more detail earlier in this document), all users from the specified group are classified as suspicious users. It shall be noted that, depending on the embodiment option, the analyzing module 220 can either send a corresponding request to each of the messaging

20 systems 110, 120 and obtain all necessary information about users in response to this request, or directly access each of the user databases 118, 128 or connect with them and extract from there all the necessary information about users.

As a supplement or alternative for identifying in messaging systems at least one user associated with a suspicious user message (i.e. associated with the analyzed user

25 message in which the analyzing module 220 has detected at least one suspicious sign as described earlier in this document), in the messaging systems that make up the system 100 in order to identify suspicious users and to interact with the device 200 to identify suspicious for example, in the messaging systems 110, 120, the analyzing module 220

(i) receives from each of the messaging systems 110, 120 all user messages stored

30 respectively in message bases 116, 126, with each received user message corresponding to with a unique identifier for a particular messaging system and a unique identifier for a specific user; (ii) extracts, by means of a corresponding regular expression from the set of regular expressions known to it, from this received user message at least one

reference; (iii) establishes a possible link between each malicious reference comprised in a suspicious user message with each of the extracted references comprised in the received user message; (iv) in case of the establishment of a specified connection between references, unites such users into a group of the related users; (v) in case when the reputation indicator of at least one user in the specified group exceeds the specified reputation threshold limit (as described in more detail earlier in this document), all users from the specified group are classified as suspicious users. It shall be noted that, depending on the embodiment option, the analyzing module 220 can either send a corresponding request to each of the messaging systems 110, 120 and receive all necessary user messages in response to this request, or directly access each of the user databases 118, 128 or to contact them and to extract all necessary user messages from them.

In order to establish the said link between the references in the messaging systems 110, 120, the analyzing module 220 performs at least one of the following operations, wherein it sets at least one of the following for each pair of compared references: (1) whether the domain names have a similar spelling (for example, by comparing them character-by-character, calculating the Levenshtein distance between domain names, comparing their hash sums calculated by analyzing module 220, and / or other well-known methods); (2) whether the domain names are registered to the same person; (3) whether the same personal data are indicated for registered domain names, in particular, phone number, actual address and / or email address; (4) whether the domain names are located at the same IP address; and (5) whether the extracted references have the same single web resource pointer “URL” (for example, by their character-by-character comparison, calculating the Levenshtein distance between these “URLs”, comparing their hash-sums calculated by the analyzing module 220, and / or another well-known methodology), with information about the persons to which domain names are registered, information about personal data (included in the registration data of the domain name) specified for registered domain names, and IP addresses where registered domain names are located, could be automatically received by the analyzing module 220 using, for example, the online service “Whois”, in particular by automatically sending a suitable search query to this online service “Whois” and extracting the necessary information from the response of the online service “Whois” or from the web - pages with the results of the search query by using, for example, a

special parser embedded in the analyzing module 220 and analyzing, for example, the text of the response of the online service "Whois" or the html-code of the named web page.

According to one of the embodiments of this technique, as an addition or alternative, the said link between references in the messaging systems 110, 120 can also be established by the analyzing module 220 by comparing for each pair of compared references the history of changes in IP addresses, running services, and the history of domain names , DNS server history, DNS record change history, SSL keys, SSH fingerprints, executable files, and other parameters of network resources. It shall be noted that the existence of a connection between the compared links can be established or determined by the analyzing module 220 based on the coincidence of at least one of the above parameters of network resources. In particular, in one of the variations of this embodiment of this technique, the relationship between network resources located at the analyzed references can be established by the analyzing module 220 by creating a well-known mathematical model in the form of a graph, wherein the vertices of the created graph correspond to at least the first network resource and at least the second network resource, and the graph edges are connections between at least the first network resource and at least the second network resource under at least one parameter of the said parameters common to at least a first network resource and at least second network resource. In such a variation of the embodiment of this technique described above, the analyzing module 220 may be configured to assign, for example, a well-known machine learning algorithm, to weights at least between the first network resource and the second network resource based on the parameter of the first network resource and the second network resource , wherein the number of connections for a single network resource parameter between one first network resource and the second network resources can be limited by the threshold value. The analyzing module 220 is additionally configured to determine a communication coefficient as a relationship of the number of connections in one parameter between one first network resource and second network resources and the weights of each connection in one parameter between the first network resource and second network resources and configured to delete connections between at least the first network resource and at least second network resource in case the value of a certain communication coefficient is less than a predetermined threshold value.

It shall be noted that in the above cases of forming groups of related users, the analyzing module 220 stores information about such related users in a separate database of related suspicious users stored in the local data storage 230, or in the corresponding subsection of the user database for storing information about related suspicious users.

5 placed in the local data storage 230 (a separate local data storage of associated suspicious users to which the analyzing module 220 can gain an access or with which it can communicate using the communication bus 240, or in a separate remote storage of suspicious users, to which the analyzing module 220 can gain access or with which it can communicate using the communication module 210, connected to the analyzing
10 module 220 via communication bus 240, depending on the embodiment of this technique) as described earlier in this document, analyzing module 220 additionally checks whether this suspicious user belongs to a group of interconnected users by searching for that suspicious user among suspicious users in a separate database of related suspicious users or a subsection of the user database for storage information
15 about related suspicious users in the local data storage 230 based on the unique identifier of this suspicious user.

In one of the embodiments of this technique, the analyzing module 220 may combine a group of interconnected users, formed from suspicious users identified by identifying similar messages (as described in more detail above), a group of
20 interconnected users formed from suspicious users identified by identifying users with similar identification data (as described in more detail above), and a group of related users, formed from suspicious users identified by identifying related references (as described in more detail above) into a general group of related users, followed by its use by the analyzing module 220 in the manner described above.

25 In some embodiments of this technique, users and / or a group of related users may be blocked by one of the above methods for a predetermined period of time, including removing at least part of the messages sent by each one of the separately blocked users and / or each of the blocked users, members of each of the blocked groups of related users in the corresponding messaging system, for example, only user
30 messages with malicious content, or all of them. In one of the versions of this embodiments, if a previously blocked user is unblocked (for example, if a user passes a given unblocking procedure and satisfies all of its conditions), the corresponding messaging system sends to the device 200 to identify suspicious users, information

about the unblocked user, set in accordance with the unique identifier of the unblocked user and the unique identifier of this messaging system, the analyzing module 220 automatically (by default) sets for such an unblocked user an indicator of his reputation at a specified initial value, different from zero and less than the value of the specified minimum reputation threshold.

Fig. 3 shows a flowchart of a method 300 for detecting suspicious users in messaging systems according to this technique. It shall be noted that the method 300 may be performed using a computing processor of any known computing device, in particular using the above-described analysis module 220 of the device 200 to detect suspicious users in the messaging systems 110, 120 shown on Fig. 2.

Method 300 shown on Fig. 3 begins with step 310, wherein a set of user messages are received from at least one messaging system.

In one of the embodiments of this technique, the stage 310 may include an additional operation, wherein the description format of the received user messages is identified, if the identified message description format does not conform to the unified data description format suitable for computing device 200, then the received messages will be converted into the specified unified format .

In another embodiment of this technique, the stage 310 may include an additional operation, wherein voice messages and videos are detected in said received user messages and converted into text.

Subsequently, method 300 proceeds to stage 320, wherein each of the user messages received at stage 310 is analyzed to identify at least one sign of suspiciousness of a message from a given set of signs of suspiciousness of messages.

In some embodiments of this technique, the stage 320 of analyzing each of the received user messages may include at least one of the following sub-steps, wherein: (1) detecting in the analyzed user message at least one malicious reference; (2) detecting at least one malicious payment system account in the analyzed user message; (3) identifying in the analyzed user message at least one malicious email address; (4) revealing in the analyzed user message information about at least one malicious account in a financial institution.

It shall be noted that the above described detection substage in the analyzed user message at least one malicious link that can be executed during step 320 may include extracting at least one link from the user analyzed message to ensure at least one of the

following operations, wherein: (i) analyzing domain name of the extracted reference for harmfulness using at least one analysis methodology; (ii) receiving at least one file located at the extracted reference, followed by its analysis for maliciousness using at least one analysis technique; and (iii) obtaining html-code of the web resource located at the extracted reference, followed by its analysis for harmfulness using at least one analysis method. In addition, the substage of detecting in the analyzed user message at least one malicious link that can be executed during stage 320 may further include performing an operation that determines whether the extracted reference matches one of the known malicious references. In addition, operation (i) of the substage of detecting at least one malicious link in the analyzed user message, which can be performed during stage 320, may additionally include checking whether the analyzed domain name at least partially coincides with one of the known malicious domain names. In addition, operation (ii) of the above substage of detecting at least one malicious link in the analyzed user message, which can be performed at stage 320, may additionally include calculating the hash sum of the resulting file located at the extracted reference, and determining whether the calculated hash sum of the analyzed file matches the hash sum of one of the known malicious files. In addition, operation (iii) of the above described substage of detection in the analyzed user message of at least one malicious link that can be executed during stage 320 may additionally include a search in the resulting html-code of the web resource of the given keywords indicating the malicious nature of the web resource.

In other embodiments of this technique, the above described step of detecting at least one malicious payment system account in the analyzed user message, which can be performed at stage 320, may include extracting at least one payment system account from the analyzed user message, with verification whether the extracted payment system account matches with one of the known malicious payment system accounts.

In some embodiments of this technique, the above described stage of detecting at least one malicious email address in the analyzed user message that can be performed during stage 320 may include extracting at least one email address from the analyzed user message with the verification whether the extracted email address matches one of the known malicious email addresses.

The method 300 may additionally include the step of further analyzing the identification data of the identified users to identify at least one suspicious user sign

from a predetermined set of suspicious user signs, ensuring that an additional assessment of suspiciousness is assigned to each of said identified users, depending on the detected suspiciousness signs of the message, each making its given contribution to the assigned additional evaluation of suspiciousness indicator, ensuring change of the specified user's reputation grade based on the assigned additional suspicion rating.

According to one of the embodiments of this technique, the above described additional step of analyzing the identification data of identified users in method 300 can be performed with respect to only those identified users whose reputation exceeds the specified minimum reputation threshold that is less than the specified reputation threshold limit.

According to another embodiment of this technique, to identify at least one sign of suspicion of each of the identified users when performing the above-described additional stage of the method 300 at least one of the following operations is performed, wherein: (1) establishing whether the avatar, full name and / or nickname of the identified user match the corresponding administrator information in at least one of the messaging systems; (2) detecting if the identified user is a program that simulates user behavior in the corresponding messaging system; (3) revealing whether the persons invited by the identified user to exchange messages within the framework of at least one message exchange channel in the corresponding messaging system are programs simulating a person's messaging activity in the specified messaging system; (4) revealing whether at least one user in the messaging systems has identification data that matches the identification data of the identified user.

According to some embodiments of this technique, when performing the above operation (2), in the framework of performing an additional stage of the method 300, the activity of sending messages by an identified user in the messaging system is analyzed for a specified period of time, and if it is determined that the activity of sending messages by the user corresponds to temporary scheme, the specified user is classified as the program that simulates the user's behavior.

According to other embodiments of this technique, the user messages analyzed in stage 320 are stored in a user message base placed on computing device 200. In one form of this embodiment of this technique, at least one of the user messages stored in the user message base hosted on the computing device 200 is sent to one of the corresponding message exchange systems 110, 120 associated with the specified stored

user message to establish the fact of the change of this user message by the user himself in the specified messaging system, wherein in case of establishing the fact of change of the user message, the specified modified user message is received from the specified messaging system, ensuring that it is reanalyzed for suspicious signs, as described earlier in this document. For the above-described variation of this embodiment of this technique, a request for ascertaining whether the user himself has changed this user message can be sent to the messaging system at least for one saved user message associated with the user whose reputation has exceeded the minimum reputation threshold. In yet another variation of the above described embodiment, at least one of the stored user messages is accessed, via a computing device, to the corresponding messaging system by ensuring that a user message is retrieved from its message database, the attributes of which match the attributes of the specified stored user message, and the hash sums of the specified saved user message and the extracted user message are calculated, wherein in the event of discrepancies between the calculated hash sums, the specified modified user message is reanalyzed for suspicious signs. For the above-described another variation of this embodiment of this technique, obtaining access to the message base of the messaging system with extracting the necessary user message from it can be performed for at least one saved user message associated with a user whose reputation score has exceeded the minimum reputation threshold.

The method 300 then proceeds to step 330, wherein, if at least one specified sign of suspiciousness of a message is detected in the analyzed user message, at least one user associated with the analyzed user message is identified in the specified messaging systems and then stage 340 is executed, whereby each of the identified users is assigned a grade of suspiciousness depending on the identified signs of suspicion message indicators, each making its own specified contribution to the assigned suspiciousness rating, ensuring that the specified user's reputation score is changed based on the assigned suspicious rating, and stage 350, wherein users in at least one of the specified messaging systems are considered suspicious users if the changed value of their reputation score exceeds a predetermined reputation threshold limit.

In one of the embodiments of this technique, when identifying at least one user associated with a suspicious user message, in messaging systems, at stage 330, messages that are similar to the specified suspicious user message are detected in these messaging systems by ensuring that the users who sent such similar messages are united

in a group of the related users, and if the reputation indicator of at least one of the users of the group received at stage 340 exceeds the specified reputation threshold limit , then at stage 350 all users from the specified group are classified as suspicious users.

In another embodiment of this technique, when identifying at least one user associated with a suspicious user message, in messaging systems, in stage 330, in these messaging systems, users with similar identification data are identified while ensuring that these users are combined into a group of related users, in the case when the reputation indicator of at least one of the users in the specified group, obtained at step 340, exceeds the specified reputation threshold limit, then at stage 350 all users from the specified group are referred to as suspicious users.

In another embodiment of this technique, when identifying at least one user associated with a suspicious user message in the messaging systems at stage 330 a possible connection is established between the references extracted from the analyzed messages in the messaging systems with other references extracted from the messages analyzed in these messaging systems, with the provision of the integration of such users into a group of interconnected users when establishing such a connection, wherein if the reputation indicator of one of the users of the said group, obtained at stage 340 exceeds a predetermined reputation threshold limit, then at stage 350 all users of the said group are referred to as suspicious users. In one form of this embodiment of this technique, when establishing a connection between references extracted from analyzed messages in messaging systems, at least one of the following may be established for each pair of references: (1) whether the domain names have a similar spelling; (2) whether the domain names are registered to the same person; (3) whether the same contact information is provided for registered domain names; (4) whether the domain names are located at the same IP address; and (5) whether the extracted links have the same “URL”.

In yet another embodiment of this technique, method 300 may include the additional step of sending information about each of the suspicious users to the appropriate messaging system.

In some embodiments of this technique, method 300 may include the additional step of sending a request to block each of the suspicious users to the appropriate data exchange system.

In other embodiments of this technique, the method 300 may include an additional step, wherein, by means of a computing device, suspicious users are detected in the corresponding data exchange systems.

5 The presented illustrative embodiments, examples and description are merely designed to provide an understanding of the proposed technical solution and are not restrictive. Other possible embodiments will be clear to the specialist from the presented description. The scope of this invention is limited only by the attached claims.

CLAIMS

1. A method for detecting suspicious users in messaging systems performed on a computing device, wherein in accordance with this method:
- 5 receiving multiple user messages from at least one messaging system;
 analyzing each of the received user messages to identify at least one sign of suspiciousness of a message from a given set of signs of suspiciousness of messages;
 if the at least one sign of suspiciousness of the message is identified in the analyzed user message, identifying at least one user associated with the analyzed user
10 message in the specified messaging systems;
 assigning each of the identified users a grade of suspiciousness depending on the identified signs of suspiciousness of the message, each identified sign of suspiciousness contributing to the assigned grade of suspiciousness, ensuring a change in a value of a reputation indicator of the identified user based on the assigned grade of suspiciousness;
15 and
 classifying the users in at least one of the messaging systems as suspicious users in the event if the changed value of their reputation indicator exceeds a predetermined reputation threshold limit.
- 20 2. The method of claim 1, further comprising:
 identifying a description format of the received user messages; and
 if the identified description format does not correspond to a unified data description format suitable for the computing device, converting the received messages into the unified data description format.
- 25 3. The method of claim 1, further comprising:
 detecting voice messages and videos in the received user messages; and
 converting the voice messages and the videos into text.
- 30 4. The method of claim 1, wherein analyzing each of the received user messages comprises performing at least one of following substages:
 - detecting in the analyzed user message of at least one malicious link;

- detecting of at least one malicious payment system account in the analyzed user message;

- detecting in the analyzed user message of at least one malicious email address;
and

5 - detecting in the analyzed user message of at least one malicious account in a financial institution.

5. The method of claim 4, wherein detecting the at least one malicious link in the analyzed user message comprises:

10 extracting at least one reference from the analyzed user message; and
performing at least one of:

- analyzing a domain name of the extracted reference for harmfulness using at least one analysis technique;

15 - obtaining at least one file located at the extracted reference, and analyzing the at least one file for harmfulness using at least one analysis technique; and

- receiving html-code of a web resource located at the extracted reference, analyzing the html-code for harmfulness using at least one analysis technique.

20 6. The method of claim 5, further comprising establishing whether the extracted at least one reference coincides at least partially with one of known malicious links.

25 7. The method of claim 5, wherein analyzing the domain name for harmfulness comprises establishing whether the analyzed domain name coincides at least in part with one of known malicious domain names.

8. The method of claim 5, wherein analyzing the obtained at least one file located at the extracted reference for maliciousness comprises calculating a hash sum of the at least one file, and establishing whether the calculated hash sum of the analyzed file coincides with a hash sum of one of known malicious files.

30 9. The method of claim 5, wherein analyzing the received html-code of the web resource comprises conducting a search in the received html-code for specified keywords indicating malicious nature of the web resource.

10. The method of claim 4, wherein detecting the at least one malicious payment system account in the analyzed user message comprises:

extracting at least one payment system account from the analyzed user message ;

5 and

determining a match of the extracted at least one payment system account with one of known malicious payment system accounts.

11. The method of claim 4, wherein detecting the at least one malicious email address in the analyzed user message comprises:

extracting at least one email address from the analyzed user message; and

determining a match of the extracted email address with one of known malicious email addresses.

12. The method of claim 1, further comprising:

analyzing identification data of the identified users in order to identify at least one suspiciousness sign of the user from a given set of suspiciousness signs of users; and

assigning an additional grade of suspiciousness to each of the identified users, depending on the identified suspiciousness signs of the message, each detected suspiciousness sign contributing to the assigned additional grade of suspiciousness, ensuring the change in the value of the reputation indicator of the said user based on the assigned additional grade of suspiciousness.

13. The method of claim 12, wherein analyzing the identification data comprises analyzing the identification data with respect to only those identified users whose reputation indicator exceeds a specified minimum reputation threshold, which is less than the predetermined reputation threshold limit.

14. Method of any of claims 12-13, wherein identifying the at least one suspiciousness sign of each of the identified users comprises performing at least one of:

- establishing whether an avatar, full name and / or nickname of the identified user matches a corresponding information of an administrator in at least one of the messaging systems;

5 - detecting whether the identified user is a program that simulates a user behavior of the user in the corresponding messaging system;

- detecting whether at least one of the persons invited by the identified user to exchange messages within the framework of at least one message exchange channel in the corresponding messaging system is a program that simulates a user behavior in the specified messaging system; and

10 - detecting whether there is at least one user in the messaging systems with identification data that matches the identification data of the identified user.

15. The method of claim 14, wherein detecting whether the identified user is a program that simulates the user behavior in the corresponding messaging system comprises:

15 analyzing an activity of sending messages by the specified user in the specified messaging system for a specified period of time; and

if it is found out that the activity of sending messages by the user corresponds to a specific timing scheme, classifying the specified user as the program that simulates the user behavior.

20

16. The method of claim 1, wherein identifying the at least one user associated with the analyzed user message in the specified messaging systems comprises:

25 detecting messages that are similar to the analyzed user message in these messaging systems, ensuring the association of users who sent such similar messages to a group of related users; and

if the reputation indicator of at least one of the users in the group of related users exceeds the predetermined reputation threshold limit, classifying all users from the group of related users as suspicious users.

30 17. The method of claim 1, wherein identifying the at least one user associated with the analyzed user message in the specified messaging systems comprises:

detecting users with similar identification data in these messaging systems by ensuring that such users are combined into a group of related users; and

if the reputation indicator of at least one of the users in the group of related users exceeds the predetermined reputation threshold limit, classifying all users from the group of related users as suspicious users.

5 18. The method of claim 1, wherein identifying the at least one user associated with the analyzed user message in the specified messaging systems comprises:

 establishing a possible link between references extracted from the messages being analyzed in the messaging systems with other references extracted from the messages being analyzed in these messaging systems, with a provision of an integration
10 of such users who sent similar messages into a group of interconnected users when establishing such link; and

 if the reputation indicator of one of the users in the group of interconnected users exceeds the predetermined reputation threshold limit, classifying all users from the group of interconnected users as suspicious users.

15

19. The method of claim 18, wherein establishing the link between the references extracted from the analyzed messages in the messaging systems comprises establishing, for each pair of references, at least one of:

- whether domain names have a similar spelling;
- 20 - whether domain names are registered to a same person;
- whether same personal data are specified for registered domain names;
- whether domain names are located at a same IP address; and
- whether extracted references have a same "URL".

25 20. The method of claim 1, further comprising sending information about each of the suspicious users to the appropriate messaging system.

21. The method of claim 1, further comprising sending a request to block each of the suspicious users to the appropriate messaging system.

30

22. The method of claim 1, further comprising saving the analyzed user messages to a user message database located on the computing device.

23. The method of claim 22, further comprising, for at least one of the saved user messages, sending a request to the appropriate messaging system associated with the saved user message to establish whether the user has changed the user message in the specified messaging system, wherein, if it is established that the user message has changed, receiving the changed user message from the specified messaging system with provision of its re-analysis for signs of suspiciousness.

24. The method of claim 23, wherein sending the request comprises sending the request to the messaging system for at least one saved user message associated with a user whose reputation indicator has exceeded a minimum reputation threshold.

25. The method of claim 22, further comprising, for at least one of the saved user messages, accessing, via a computing device, the corresponding messaging system to retrieve a user message from its message base, attributes of which match attributes of the saved user message; calculating hash sums of the saved user message and the retrieved user message; and, in the event of a mismatch of the calculated hash-sums, carrying out a repeated analysis of the specified changed user message for signs of suspiciousness.

26. The method of claim 25, wherein accessing the corresponding messaging system to retrieve the user message from its message base is carried out for at least one saved user message that is associated with a user whose reputation indicator has exceeded a minimum reputation threshold.

27. The method of claim 1, further comprising blocking the identified suspicious users in the respective messaging systems by means of a computing device.

28. A computing device for detecting suspicious users in messaging systems, configured to communicate with messaging systems and comprising memory for storing machine-readable instructions and at least one computing processor configured to execute the machine-readable instructions ensuring the implementation of a method for detecting suspicious users in messaging systems according to any of claims 1-27.