Detecting cyber threats through social network analysis: Short survey

K.Shwetha, Assistant Professor, Department Of Cyber Security, SICET, Hyderabad
N. Madhu Bhavani, Assistant Professor, Department Of Cyber Security, SICET, Hyderabad
V. Swathi, Assistant Professor, Department Of Cyber Security, SICET, Hyderabad

Abstract

Thisarticleconsidersashortsurveyofbasicmethodsofsocialnetworksanalysis,whichareusedfordetecting cyberthreats.Themaintypesofsocialnetworkthreatsarepresented.Basicmethodsofgraphtheoryanddata mining, that deals with social networks analysis are described.Typical securitytasks of social networkanal- ysis, such as community detection in network, detection of leaders in communities, detection experts in net- works, clustering text information and others are considered.

Keywords:socialnetworkanalysis,datamining,threats,socialnetworksecurity.

Introduction

The use of the Internet social networks makes it possible to communicate with old friends, make new ac- quaintances,expresstheirthoughtsonaverywideaudience,joingroupsofinterest.Bycoverageofaudiences some groups in social networks and popular bloggers can compete with many media.

According to efficiency information transmission social networks are often superior to most of media, they areabletodisseminateinformationaroundtheworldinseconds,therebyexpeditingtheprogressofoperation, but this does not mean that television and radio have lost their popularity.

In modernconditionsthereisasymbiosisofmajortelevisiongiantswithsuchnetworksasWikiLeaks,Face- book, Twitter, YouTube, reinforcing the ultimate effect of informational influence.

The rapid development of social networks and ability to collect information from them led to a noticeable increase of interest to social network analysis and the occurrence of its new methods become increasingly popularandtheyareusedinvariousfields[5,6]:expertsearchsystem,gatheringateamofspecialists,social recommendations,searchenginespeopleanddocuments,marketing,communications,advertising,andmany others. Nowadays social network analysis (SNA) is used to study a variety of economic and organizational phenomena and processes [25, 42, 61].

TheGlobalRiskReport2017[87]justpublishedbytheWorldEconomicForumaheadofitsannualmeeting in Geneve, continues to receive attention from the business world and reaches a level of credibility in a crowdedarenaofforecasting. Cyberattacksonbusinesseswillincreasewith adenialofservice,databreach, cloudprovidercompromiseandextortionbeingmajorconcernsforITdepartments.Thesensitivegeopolitical context,theriseofcyberattacksandmajordatabreachesandhacks,aswellastheglobalinsurgencyofviolent extremismandradicalizationhaveledmanycountriestotheadoptionofsecuritymeasuresandcounterterror- ism laws that have increased scrutiny and restrictions on the participation of societal actors [87]. In this case many countries in the world create the National Cyber Security Centres and Cyber Security Strategy Docu- ments [68, 69, 89, 97].

Social network analysis (SNA) is used effectively to counter money laundering, identity theft, online fraud, cyberattacks,andothers.Inparticular,theSNAmethodsareusedintheinvestigationofmanyillegalopera- tions with securities and investments, for the prevention of riots and others [9, 85].

Moreover, social networks are increasingly used in the interests of information and psychological influence. Theyprovideopportunitiesintermsofinfluenceontheformationofpublicopinion,theadoptionofpolitical, economic and military decisions, influence the enemy's information resources and distribution of specially prepared information (disinformation) [19-22, 34].

Thus,thetaskofcollection,monitoringandanalysisofsocialnetworksisimportantandactualforinformation security. A review of main SNA methods used to contend for social networks threats is conducted in this work. The basic directions for studies in the field of social networks security are given; at the moment the most popular computer social networking are described; main methods and algorithms used in various net- workanalysismodelsarebrieflypresented; primaryapproachestonetworkanalysisareshown;sometypical tasks of social network analysis are given.

Socialnetworkingsites

The worldwide accessibility to the Internet is one of the defining phenomena of the present times reshaping the world as we know it [84]. The love child of the World Wide Web is social media, which comes in many forms,includingblogs,forums,businessnetworks,photo-sharingplatforms,socialgaming,microblogs,chat apps,andothersocialnetworks.Thenumberofsocialnetworkusersworldwideis2.34billion,anditwas 1.91 billion in 2014. In 2018, it is estimated that there will be around 2.67 billion social media users around the globe and is expected to reach some 2.95 billion by 2020 [84].

Letspresent top 15social networkingsites in the world [90].Much moresocial networkingwebsites present in the list of social networking websites [57].

FacebookisthebiggestsocialmedianetworkintheInternet,bothintermsoftotalnumberofusersandname recognition. It was founded on February 4, 2004. The number of active Facebook's users worldwide is 1.87 billion. Facebook is leading the pack with a huge margin in front of Youtube. The share of time spent on Facebook via mobile device is 68%.

YouTube is the largest and the most popular video-based social media website. It was founded on February 14, 2005.The number of monthly active Twitter users is 1 billion.Twitter's worldwide revenue is $2.53 bn. Instagramisa visual socialmediaplatform,likePinterest. It wasfoundedonOctober 6,2010,hasmorethan 600 million monthly active users, it is owned by Facebook.

Twitter is the social media platform that limiting your posts to 140 characters. It was founded on March 21, 2006 and has more than 313 million monthly active users.

Reddit is social news and entertainment networking website where registered users can submit content such asdirectlinksandtextposts.RedditwasfoundedonJune23,2005,hasmorethan234millionmonthlyactive users.

Vine Camera is a rapidly growing video sharing social media app that allows users to share 6-second video clips with their followers. Vine was founded in June 2012, and has more than 200 million monthly active users.

Ask.fm is a questions and answers network, launched in June 2010, and has more than 160 million monthly active users.

Pinterest is a relatively newcomer in the social media arena. This platformconsists of digital bulletin boards wherebusinessescanpintheircontent,launchedinMarch2010,hasmorethan150millionmonthlyactiveusers.

Tumblr is the social network platform that has not limited in the type of content that can be shared: quote posts, chat posts, video and photo posts, audio posts. Tumblr was founded in February 2007, has more than 115 million monthly active users.

Flickr isanonlineimageandvideohostingplatform. It wascreatedonFebruary10,2004 and hasmorethan 112 million monthly active users.

Google+ is an interest-based social network, its SEO (Search Engine Optimization) value alone makes it a must-use toolforanysmallbusiness,launchedonDecember 15, 2011,it has more than111 million monthly active users.

LinkedInishands-downthemost popularsocial mediasitefor professional networking, launched onMay5, 2003,has more than 106 million monthly active users.

VKisthelargest Europeanonlinesocial media andsocial networkingservice,launchedinSeptember 2006, has more than 90 million monthly active users.

ClassMatesisasocialnetworkingservicethathelpuserstofindclassmatesandcolleaguesfromkindergarten, primary school, high school, college, workplaces, and the U.S. military, founded on November 17, 1995, it has more than 57 million monthly active users.

Meetup is online social networking portal that facilitates offline group meetings in various localities around the world. It was founded in 2002, has more than 30 million monthly active users.

Threatsinsocial networks

Thesedaysthereisawholerangeofdifferentthreatsinsocialnetworks.PaoloAltoNetworks[72]described threatsforbussinessocialnetwork,socialmediathreatsandsomepreventivemeasures[9,24,62,79].

1. Social engineeringis the most popular tactic for cybercriminals. Social networks allowattackersto find confidential information that can be used for property and moral damages.
2. Friends. The trust to those who entered in the “friends” list is always higher than to random people. On the one hand, this is good, since forming a loyal audience around the company, brand or person. But on the other hand, it is an opportunity for attackers.
3. Possibility of substitution of person or masquerade: for sure it is not clear exactlywho hide their actions behindthenameof friendsorhidingbehindphotosfriendsinsocial networkprofile.It ispossiblebythe IP-address of sender to gather at least some information about himin the correspondence bye-mail, that is not work in social network.
   Thismasqueradeispossibleatthecorporatelevelalso.Theresultofsuchmaliciousscriptcanbephishing, the organization of “black PR” or “Antipiar”. There were many instances where it was not clear who created the site on behalf of any company – it is created a problem for the original brand.
4. Stealingpasswords and phishing. As the identificationof social networks uses passwords,it is sufficient to know the sequence of characters and can be possible to send advertising, some information on behalf of others, or to motivate recipients to any negative action, in particular to pass on the link and run the maliciouscode,anddoother(oftenillegal)cases.Besides,somecompaniesusesocialnetworktopromote their own products, and the theft of administrator group password allows to steal the group itself. To obtainconfidentialinformationtraditionally,phishing,dummysites,socialengineering,andmoreothers areused.ProtectionagainsttheseattackmethodsareconsideredDLP-system(DataLossPrevention)and reputation technologies that are integrated into a variety of anti-virus products.
5. URLshorteningservicesusage.Inrecentyears,URLshorteningservicesallowtomaskunwantedwebsite address under the short link are especially popular. In fact, the domain redirects the visitor. Today there isanactivestruggleagainsttheserisks–URLshorteningservicebegantouseimprovedmechanismsfor the detection of spam and other threats. However, for users of social networking this threat is keeping – alluring messages and offers from familiar contacts that have been hacked, often lead to downloading malicious software or display unwanted web pages.
6. Using the same user names and passwords on the corporate network and external social resources. As a result,hackingprofilesofsocialnetworkuserssignificantlyincreasestheriskofpenetrationtocorporate resources on behalf of one of the company’s employees.
7. Web-attack. As social networks are web-based applications, they can be used by hackers to organize at-tacksonvulnerabilitiesinbrowsers.ThetoolsforsuchattackscanbeTrojanapplications,fakeantiviruses, social worms, which are used to spread own friends lists and other. Their main goal is to get into the informationsystemofsocial networkvisitor and gainafootholdinit.Such traditional toolsasanti-virus software,thatareabletoworkinrealtimeandblockthedownloadofmaliciouscodeareusedforprotection.
8. Information leakage and compromising company employee’s behavior. Social networks can be used to organize leaks of important information for the company, as well as to undermine its reputation. Such attack can conductinternal employeeswhoare dissatisfiedwith theleadership, orspecially embedded

insiders. In social networks persons often behave quite differently from the corporate communication environments, and it is possible that shocking publication and rough replicas can cause some damage to thereputationoftheiremployers.DLP-systemsandproductsfortheanalysisofpublicationsontheInter- net intended to protect against these threats.

9. Thegrowthoftraffic,especiallyviewingvideosources.
10. Inducementofminorsforsexualpurposes(grooming).
11. Contentwithsignsofincitementtoracial,ethnicorreligioushatred,propagandaoftotalitariansects.
12. Propagandaandpublicjustificationofterrorism[85].
13. Cyberhumiliationandcyberbullying.
14. Promotionanddistributionofdrugs.

Toprotectfromthisthreats,theinformationsecurityservicessolve next problems:

- detectionofinformationattacks:definethenodesfromwhichtheattackismade,theoptimalplacement of signal points;
- preventinginformationattacks:estimatedcostoftheattackontheobjectofattackanddefense costs;
- formationanddestructionofdifferentnetworks:socialand/orinformation;
- detectionofintruderscommunities:suchasterrorists,trackingmaliciousactivity.

Thefollowingdirectionstocounterinformationandpsychologicalimpactofvirtualcommunitiescanbe identified [34, 74]:

- forcemethods–serversclosure,trafficshaping;
- legal and regulatory practices – criminal responsibility of organizers and participants of the virtual communities;
- Internet censorship;
- monitoringandanalysisofsocialnetworks.

Lets consider advantages and disadvantages of each method. Thefirst two methods are effective in the short term, but they have some disadvantages: the lack of geographical boundaries and limitations for instant dissemination, collection,processingand useofinformation – beyond thescope oflaws legal regulationof any government; anonymity; easily accessible variability of information in electronic form. Censorship works poorly in democratic states based on freedom of speech.

Methods of monitoring and analysis of social networks are more effective in the long term, but require the involvementofspecialistsinvariousfieldsofscience. Asvirtualsocialgroupshavetheabilitytoreorganize, themaintaskofmonitoringandanalysisofvirtualcommunitiesthatrepresentathreattothenationalsecurity ofinformationisnottheirdestruction,butmanagementandcontroloftheiractivitiesbyavarietyofmethods.

Monitoringandanalysisofsocial networks

A great number of special software to monitor and analyze the Internet environment were developed by this time. Major functions of these systems are:

- monitoring:providesautomatedinformationsearchintheInternetenvironment,todetermineandchange the keywords to information search using information retrieval languages;
- analysis:automaticprocessingofinformationflows,revealingfactsandevents,visualizationofanalyti- cal data in the form of digests, charts, graphs, and other types of reports.

Monitoringreferstothe processofcontinuous informationcollectionfromsocial networksinorderto maintain furtheranalysis.So,thesearchconductinthescientificinvestigationsisconsideredbyaglobalsearchenginefor socialnetworks[33,73,74,75,93],andthedevelopmentofcommercialsearchenginesforspecialapplications [40,57,79,81],whichdonottakeintoaccountthepeculiaritiesoffunctioningdiscussionpages.

Somepossibleapproachestosocialnetworksanalysis

Inpresenttimethefourmainapproachesareallocatedinsocialnetworksanalysis[1,82]:structural,resource, regulatory and dynamic.

Structural approach is focused on the geometric shape of network and intensity of interactions (weight of edges),socharacteristicssuchastherelativepositionofthevertices,center,transitiveinteractionsareinvestigated.Themethodsofstatisticalanalysis,clusteringandclassificationalgorithmsareusedinstructuralanal- ysis and analysis of connections behavior. The behavior of vertices in the process of clustering and typical temporalcharacteristicsofsocialnetworksarestudied.Forexample,thedetectionofphishingattacks(define thenodesfromwhichtheattackismade,theoptimalplacementofsignalpoints),trackingmaliciousactivity, how the structure of the network changes in the process of growth or how behavior and distribution of con- nectedcomponentsofthegrapharechanged.Agreatimportanceisgiventotheallocationofcommunitiesin socialnetworks(forexample,detectionofradicalcommunitiesandtrackingtheiractivity).Thegoal istotry todefinethenetworkregions,withinwhichthereisanactiveinteractionofparticipants.Algorithmically,this problem can be attributed to the problem of graphs division. It is necessary to divide the network into dense regions on the basis of behavior of links between vertices. Computer social networks are dynamic that leads to difficulties in terms of identifying communities. In some cases it is possible to integrate the information content of the network in the process of community determination. Then, the content is additional means to identify group members with similar interests.

Resource approach considers the participantspossibilityof attractingindividual and networkresources orto achieve certain goals and differentiates participants, who are in identical structural positions of the social network, according to their resources. Knowledge, prestige, wealth, race, gender can serve as the individual resources. Status, information, fund are understood as network resources. For example, determining signifi- cance(influence)ofsocialsiteasapoliticalplatform(alsodefinitionofsocialnetworksimportanceanddegree of their overlap, importance of individual communities, individual users); determining importance of topics discussed(e.g.,publicdisorder),events(e.g.,theextensionofRussiansanctions,processes(e.g.,thesituation in Ukraine) and persons (e.g., Putin), as well as the attitude to them [49-53, 60].

The analysis of social networks content is an important task of this direction. The network content serves as a source for a wide range of applications that focus on the extraction and data analysis. The use of network contenthelpstoimprovethequalityofconclusionsinsocialnetworksanalysissignificantly,forexampleinthe problemsofclusteringand classification. Four types of network content analysis can be identified [12, 13].

1. The methods of random walks are used in the analysis of general information with arbitrary data types. Oneofthemostwell-knownalgorithms bysuchmethodsisthereferencerankingalgorithm(PageRank). This algorithm can also be used for search and classification of entities and participants in the social network,toassesstheprobabilityofvisitingaparticularvertex.Itisnaturalthatverticesarebetterlocated withstructuralpointofviewandhaveahigherweight,and,therefore,theyaremoreimportant.Themeth- ods of random walk also could be useful to bring together participants in the group relative to the most influential members.
2. Forsensoryandflowanalysistheuseof dataintegrationtechniquescomingfromsensorsanddataavail- ableonthesocialnetworks.Modernmobilephonessupportusersinteractionwitheachotherdynamically inrealtime,dependingontheirlocationandstatus.Theyareusedtoobtaininformationaboutapersonor a combination of the properties of objects that are monitored.
3. Analysisofmultimedia.Therearemanysites(Flickr,YouTube,andothers)fortheexchangeandsharing ofmedia:photo,video,audio.Inthepresenceoftagsorcommentsmultimediaanalysiscanbereducedto the text information analysis in network.
4. Analysis of textual information. A lot of text information contains in various forms, for example, it is possible to add comments, links to posts, blogs or news articles in the social network. Sometimes, users can tag each other, which is also a form of text information in the form of links. The placement of tags (labels or keywords) describing various objects: images, text, video is of particular interest. Under this approach, properties of tags flow, models tagging, semantic tagging, imaging tags, applications for their placement, etc. are studying.

*Normativeapproach*studiesthelevestoftrustbetweentheparticipants,andtherules,regulationsandsanctions influencingonbehaviorofparticipantsinthesocialnetworkandprocessesoftheirinteractions.Inthiscase,the socialrolesofanalysisareassociatedwiththenetworkedge,forexample,thedefinitionoforganizers,managers and implementers of illegal actions; the relationship manager and a subordinate, friendships or family con- nections. Since social networkingis based on the interaction between the various participants, it is natural to assume that this interaction has influence on the participants in terms of their behavior. The issues for this direction: how to simulate the influence on the basis of information about the participants; how to simulate the spread of influence; who is influenced in the process of distribution.

Social networks contains a lot of personal information about the participants, for example, interests, friends, demographicsandother.Thiscanleadtounauthorizeddisseminationofpersonalinformationinthenetwork. Inmakingdecisionsoftasksofsuchtypeitisusefultoapplythemodelsbasedonconfidentialitymechanisms. Functional roles of social network participants are important for the effectiveness and sustainability of the socialnetwork, thesocial networkcanbea tooltoidentifyexpertsin aparticularfield.Toidentifyexpertsin social networks, for example, approach of ant colony optimization (ACO) is used.

Often,inreality,theexpertsformanetwork,whichcorrespondstoasocialnetworkorcompanyorganizational structure.Inadditiontoexpertstheso-calledbrokery–peoplewhoplayroleofmediatorinthesocialnetwork by linking a group of people, establishing communication between professionals and thereby giving them access to information are of interest.

Dynamicapproach isadirectioninthestudyofsocial networksinwhich researchobjectsarethechanges in the networkstructureovertime: thereare newparticipants,some participants have stoppedinteraction,there arenewcommunications,someconnectionsareoutdated,astheparticipantsarenolongerinteract(definition of the relationship between terrorist organizations and their members, identification of regularities based on which theeventsassociatedwithterroristsoccur and prediction ofsimilar eventsin thefuture).Thisleadsto changes in the structure of the social networks in general and in some communities. Herewith the questions: according to what principles long-term changes between the major communities in social networks happen, are there any fixed configurations of social networks, how community develops in time, what changes can occur, howitscanbetracked and submitted. Animportant taskistoforecast theformationofconnectionsin socialnetworks.Mostapplicationsfortheanalysisofsocialnetworksaredynamicandmaychangeovertime. The process of forecasting connections may be involved as a network structure and information about the features of different vertices.To solve theseproblems, it isproposedtobuilda varietyof structural and relational models.

Basicclassesofmethodsusedforsocialnetworksanalysis(SNA)

ThefollowingbasicclassesofmethodsusedinSNAcanbedistinguish:methodsforanalyzinggraphs,statistical techniques, data mining, methods of optimization theory and theory of algorithms. The allocation of separatemethodsofsemanticandtextanalysis isalsoconvenientforsystemsclassification.Inthiscase,you have topayattentiontothesystemsupport oflanguage, bymeans of whichtheusersofsocial networkcommunicate.TwomaingroupsofbackgroundSNAaregraphstructuresandmodels,includingrandomwalkson graphs and data mining methods.

Graph analysis

GraphmodelsandanalysismethodsplayanimportantroleinSNA,becauseanysocialnetworkcanbemathematicallyrepresentedasagraph $G=(V,E)$,where Vistheset ofvertices,Eistheset ofedgesofthe graph, Nisthenumber ofvertices.Graphmodelsofsocial networksareusedtomodeltheeconomicandcommunication links of people, analyze the processes of information dissemination, find community and related subgroups, on which the entire social network can be divided. The participants and edges indicate the existence ofrelationsbetweentheminthesocialnetworkvertices.Relationshipsinanetworkcanbeclassifiedaseither directional or nondirectional. As a rule, there are two types of relationships: friendship (people are familiar with each other) and interests (people are included in one group with the same interests).

To analyze the graph of social network it is convenient to use the graph density defined as the ratio of the number of edges in the analyzed graph to the number of edges in a complete graphwith the same number of vertices (the complete graph is the graph in which all vertices are connected to each other). In addition, the networkcanbecharacterizedbyquantitiessuchasthenumberofpathsofagivenlength(thepathisasequence ofverticeslinkedtogether),theminimumnumber ofedges,theremoval ofwhichdividesthegraph intoseveral parts and others [14].

The analysis of the centrality and other local properties. To determine the relative importance (weight) of graph vertices (that is how participant within a specific network is an influential), the concept of centrality (measureofclosenesstothecenterofthegraph)isintroduced.Centralitycanbedeterminedindifferentways, so there are different measures of centrality [38].

Degreecentralityofthevertexisthenumberofedgesincidenttoit(incidencemeanslinkbetweenthevertex and edge).Thereareincoming andoutgoing communication.The incomingonecharacterizes thepopularity of a participant, the outgoing one implies their sociability. Obtained value can be normalized by dividing to thetotalnumberofparticipantsinthenetwork.Inotherwords, thedegreecentralitysuggeststhatmoreinflu- ential participant is the one who has the most friends or who is included in a larger number of communities.

Closeness centrality is a measure of information dissemination in the network from one participant to the others.Theshortestpathinthegraphisusedasameasureofthedistancebetweentwoparticipants.Thus,the participant's direct friends are at the distance equal to 1, friends of friends are at the distance equal to 2, and so on. The closeness is defined as the inverse of the normalized sum of all distances. Closeness centrality allows to understand how close a participant is to all other members of the network.

Another characteristic of a participant is his importance in the dissemination of information. Betweenness centralityevaluatestheparticipantsinthiscontext.Itiscalculatedasthenumberofshortestpathsbetweenall pairs of participants passing through the considered participant.

Eigenvector centrality shows the relationship between the centrality to a participant and centralities of his friends. The participant, who has many connections with those who have many connections also has high eigenvector centrality. Thus,themoretheparticipanthasfriendsandthemoretheircentrality,thegreaterhis eigenvectorcentrality. The measure of eigenvector centrality is difficult to calculate and it is computed with the help of specialized software only.

CentralitycanbecalculatedusingthealgorithmPageRank,whichisusedinGoogle.PageRankisanalgorithm used by Google Search to rank websites in their search engine results. PageRank is a way of measuring the importance of website pages. PageRank works by counting the number and quality of links to a page to de- termine arough estimate ofhow important the websiteis.The underlyingassumption is that more important websites are likely to receive more links from other websites. Thus, PageRank is the method of calculating the weight of the page by counting importance of references to it [54]. Apart from the listed methods of determining centrality, there are a large number of introduced non-classical methods of calculation of this characteristic of the network.

Theimportant characteristics of the networkconnections are balance and transitivity. Balance isthe absence of situations such as "positive interaction (friendship, partnership) between 1st and 2nd participants and be- tween 1nd and 3d, but negative interactions (enmity, rivalry) between 2st and 3d". It is argued that balanced networksarepsychologicallymorecomfortableforparticipantsand morestablecomparedtotheunbalanced one [43]. Transitivity is the fulfillment of conditions of the form "if there is an interaction between 1st and 2ndparticipantsandbetween2ndand3d,theinteractiontakesplacebetween1stand3dones".Thesecharac- teristics describe the local connections of participants and often used in the analysis of dyads and triads.

ThelevelоftrustistheusefulfeatureinSNA.Analgorithmforcalculatingthelevelоftrust(TrustRank)was originally created to separate the informative web pages from spam [36]. If we talk about this algorithm in termsof sites,theexpertsmanuallyestimatethetrustlevel forasmall number ofsitesthat canbeconsidered reliable.These sites are taken as the standard. Further,the algorithmis based on the assertionthat good sites are seldomlinked to bad, but the bad one is veryoftenrefered to the good one.TrustRankis the value which givesanestimateofwhetheryoucantrustaparticularsite,assumingthatitcontainsnospam. Themorelinks onthesite,thelesstrust "ispassed"oneachsuchlink.Thedegreeoftrusttosite(TrustRank)decreaseswith increasing distance between it and the original sample of standard sites.

Optimization on graphs. The travelling salesman problem can be distinguished among the most important optimizationproblemsrelatedtographs[92].Thisisoneofthemostfamouscombinatorialoptimizationtasks, which is to find the most profitable path, passing through the given vertexes at least once, then return to the original. Under the conditions of the task the criteria of the route profitability (the shortest, the cheapest, the cumulativecriteria,etc.)andthecorrespondingmatrixofcost,distanceandthelikearespecified.Antcolony optimization algorithmis an efficient polynomial-time one for finding approximate solutions of the problem of traveling salesman and as well as similar tasks of route search on graphs. The approach is to analyze and use the behaviors of ants searching for the path from the colony to the food source, and is a metaheuristic optimization [4].

Data mining

Nowadaysalmosteveryoneusessocialnetworkingsites.Manycompaniesareeagertoanalyzehugeamounts ofsocialnetworkdatatotakeadvantageofthissocialphenomenon.Socialnetworkdataminingisoneof the hottest research topics. The application of efficient data mining techniques has made it possible for users to discovervaluable,accurateandusefulknowledgefromsocialnetworkdata[2,15,59,61].

Data miningisa multidisciplinaryareawhichisoriginatedanddevelopedonthebasisof suchsciencesasapplied statistics, patternrecognition, artificialintelligence, database theory, theoryof algorithms and others. The main feature of data miningis acombination of broad mathematical tools (fromthe classical statistical analysisto new cybernetic methods) and the latest advances in information technology. Data mining methods and algorithms include: artificial neural networks, decision trees, symbolic rules, algorithms of the nearest neighbor and k-nearest neighbors, support vector machines, Bayesian networks, linear regression, correlation and regressionanalysis;hierarchicalclusteranalysismethods, thenon-hierarchicalones,includingk-meansandk-mediansalgorithms;associationruleslearning,includingapriorialgorithm;enumerationmethods;evolution-aryprogrammingandgeneticalgorithms;avarietyofmethodsfordatavisualizationandmanyothermethods [17].Mostoftheanalyticalmethodsusedindataminingtechnologyarewell-knownmathematicalalgorithms and methods. New in their application is the possibility of their use in solving various concrete problems, it will be possible due to modern hardware and software.

The most common tasks of data mining are classification, clustering, association, forecasting and visualization.Theclassificationisthemostsimpleandcommon taskof datamining.Asaresultofaproblemsolution ofclassification,thesignswhichcharacterizegroupsofobjectsofinvestigateddataset(classes)aredetected. A new object canbeattributedtooneoranother class basedonthesesigns.To solve theproblemof classifi- cation the nearest neighbor and k-nearest neighbor methods, Bayesian networks, induction of decision trees and neural networks are used [63].

Clusteringisalogicalextensionoftheideaofclassification.Thistaskismorecomplicated,clusteringfeature is that the object classes are not originally predestined. The result of clustering is a partition of objects into groups. In contrast to the tasks of classification, cluster analysis does not require a priori assumptions about thedataset,doesnotimposerestrictionsontherepresentationoftheobjects,itallowstoanalyzevarioustypes ofdata(interval,frequency,binary).Clusteranalysisallowstoreducethedimensionofthedata,tomakethem visible. The methods of cluster analysis can be divided into two groups: hierarchical and non-hierarchical. Each group includes a variety of approaches and algorithms [64].

The essence of the hierarchical clustering is sequential merging of smaller clusters in the large ones or separationoflargeclustersintothesmallerones.Theadvantageofhierarchicalclusteringmethodsistheirvisibil- ity. Hierarchical algorithms associated with the construction of dendrograms, which are the result of hierarchical cluster analysis. Dendrogramdescribesthe closeness of individual points and clustersto each other,it is a graphic sequence of merging (separation) of clusters.

The methods of hierarchical cluster analysis are not suitable with a large number of observations. In such cases, non-hierarchical methods based on separation, which are iterative methods of fragmentation of initial set are used. During separation new clusters are formed as long until a stopping rule will be executed. This non-hierarchical clustering consists in dividing the data set into a certain number of individual clusters.

Oneofthemostpopularmethodsofdataanalysisistheprincipalcomponentanalysis,whichoriginatesfrom statistical analysis applied [76]. This is one of the main ways to reduce the dimension of observation space, losingtheleastamountofinformation.Itisusedinmanyareas,includingeconometrics,bioinformatics,image processing, data compression, social sciences.

Socialnetworkanalysisandcommunitiesinterest

The interest of researchers relates to the fact that it provides a new set of explanatory models and analytical tools that are outside the ordinary quantitative methods. At the same time, a wealth of mathematical tool, allowing to build very complex models of social interactions describing almost any social system has accumulated in this field [31, 32].

GartnerAnalyticalAgencyin2012publishedareportcalled“HypeCycleforEmergingTechnologies”[45]. Accordingto the report, technology“Social Analytics” and “BigData” are nowon the so-called “Peakof inflated expectations”.Inparticular,studiesareactivelyengagedinsocialdataatCarnegieMellonUniversity,Stanford, Oxford,INRIA,aswellasthecompany’sFacebook,Google,Yahoo!,LinkedInandmanyothers. Companyowners of online social networking services (Facebook, Twitter) are actively invest in the development of improved infrastructure (Cassandra, Presto, FlockDB, Thrift) and algorithmic (new search algorithms and recommendationsofusersofgoodsandservices)solutionsforhandlinglargeamountsofuserdata.Commer- cial companies that provide services to access the repository of social data (GNIP), the collection of social data for a given scenario (80legs), social analytics (DataSift), and also an expansion of existing platforms using social data (FlipTop) are successfully developed.

Thus, experts from research centers and companies around the world use social media data for modeling social,economic,politicalandotherprocessesfrompersonaltogovernmentlevelstodevelopmechanismsof these processes actions, as well as create innovative analytical and business applications and services.

Therearemanyorganizationsinterestedinsocialnetwork:InternationalNetworkforSocialNetworkAnalysis (INSNA) – the professional association for researchers interested in social network analysis [41], NetLab is aninterdisciplinaryscholarlynetworkstudyingtheintersectionofsocialnetworks,communicationnetworks, and computer networks [70], Center for Computational Analysis of Social and Organizational Systems (CASOS)atCarnegieMellon[11],Orgnet–SocialNetworkAnalysissoftware&servicesfororganizations, communities, and their consultants [71], the International Survey Center conducts research on social, economic and political issues using survey data from large, representative national samples from many nations [88].

Social networkanalysis software (SNAsoftware) is the software which facilitates quantitative or qualitative analysisofsocialnetworks,bydescribingfeaturesofthenetworkeitherthroughnumericalorvisualrepresentation [82]. The collection of social networkanalysis tools and libraries is showed in Social NetworkAnalysis Software andSoftwareforSocialAnalysis[82,83].Top30SocialNetworkAnalysisandVisualizationToolsarealsodescribed[91].20+FreeandOpenSourceSocialNetworkAnalysisSoftwareispresented[29].

Sometasksofsocialnetwork analysis

Communitydetectioninnetwork

Communitiesinthenetworkarecharacterizedbythepresenceofalargenumberofconnectionsbetweentheir participantsandsignificantlyfewercontactswithotherparticipants.Thesimplestcaseisacommunity,where eachparticipantisassociatedwitheachother,andtheothermemberscannotbeincludedinthisgroup,asthey havenocommunicationwithmembersofthecommunity(clique).Cliqueisthemostcompletesubgraph ofa givengraph.Thedetectionofcommunitiesisanimportantproblem,includingclassificationbynetworkmembers,andasaresult,theidentificationofhomogeneousgroups,groupsofleadersorgroupsofcriticalconnections[8,16,46,58].Communitymaycorrespondtogroupsofwebpagesthathavesimilartopics[27],groups of related individuals in social networks [30], etc.

The detection of community is actually analogue to clustering, traditional task of data mining in relation to various social networks. The approaches to the allocation of target groups by identifying communities allow theirsimulation,followedbytheusemodelsofinformationinfluenceandmanagement[35].Atthesametime SNAinvestigatesthestructureofrelationshipsbetweenparticipantsofvariousapplicationareasbydetecting the implicit links between them involving graph theory [26]. More detailed overview of the community detection methods can be found in Fortunato S. [28].

Stabilityanalysisofcommunity

Theanalysisofexplicitandimplicitcommunitiesallowstostudythestabilityofsocialstructures.Toanalyze the stability of a group structure over time the following technique is typically used. First three-dimensional matrix is constructed where rows represent the estimates of interactions of participant with all the other participants, submitted by the participants themselves. The columns are participant's own estimates of interaction.Thetimeperiodsarelocatedonthethirdaxis.Furthergraphshowsthestructuralchangesofcommunity overtime.Thereafter,thetechniquesfordimensionalityreductionareapplied(forexample,principalcompo- nent analysis), i.e., the projection of the vertices into Euclidean space of reduced dimension to describe the relationships between the rows and columns of the matrix is considered. As a result, you can visualize the changes of network user status against the backdrop of changes in subgroups status [43].

Obtained projection can be clustered using a standard iterative clustering algorithms (for example, k-means) orthehierarchicalones[47].Theadvantageofhierarchicalmethodsisthepossibilitytorepresentthecluster-ingresultinadendrogram.Inthiscasewecanobtainnotonlyapartitionofthegraphintogroups,butalso the hierarchyof groupsandsubgroupsinthe graph. The basic difficultyof such methods is toselect suitable measure of distance (the shortest path between vertices) or a measure of similarity. Clustering may be performed not only from below upwards but also from top down, i.e., first, the whole network is considered as one group, and the consistent separation of one communication takes place at each iteration.

Detectionofleadersin communities

Thesearchforleadersinthecommunityisanimportanttask,sinceinthestudyandmodelingtheinformation influenceitisimportanttohavedataaboutthenatureoftheinteractionsofcommunitymembers,theconnec- tion between them and the laws of information flows distribution. According to Goyal A. et al. [31], some participantisaleaderif,afterhiscommittingofaspecificaction,asignificantnumberoftheotheronesrepeat the same action at a given time interval. The value of betweenness centrality is one of the keys in finding leaders because the more often the communication pathways pass through a vertex in a network, the higher the degree of its information interaction with other vertices [39].

The problems of leaders detection are widespread in many areas. For example, the hypothesis of influential members are considered in relation to the marketing tasks [96], the choice of many individuals to offer any product or innovations [44], the distributionand maximization of the influence inthe competitive socialnetworks and attraction of followers, viral marketing [10], dissemination of social influence [18, 80] and etc.

Detectionexpertsinnetworks

Socialnetworkcanbeatooltofindexpertsinparticularfield.Expertsdetectionisrelated totheproblemsof trust identification and influence distribution, and is a problem of information dissemination in the network. From this viewpoint, the spread of expert influence is transitive, i.e., the influence is transmitted from one node to another, decreasing with each involved experts node [7, 94].

Ant colony optimization approach (ACO = Ant Colony Optimization) is used, for example, for experts definitioninsocialnetworks.Amoredetailedoverviewofexpertsmethods detectioncanbefound,forexample, in Charu C. [13].

Evolutionindynamicsocialnetworks

Thereare newparticipantsin socialnetworks overtime, some participants stopped interaction, generatenew links,somelinksbecomeoutofdate,astheparticipantsarenolongercommunicate.Thisleadstochangesin social networks structure and in some communities. Thus, two important questions appear: 1) according to whichruleslong-termchangesbetweenmajorcommunitiesinsocialnetworksoccur;2)howtodevelopcom- munity over time; and to find out changes that can occur, capabilities of their tracking and presenting.

Toinvestigatethedynamicsofnetworktheapproaches areused[6,37,23].Thesimulationofnetworkgraph evolution exploring different strategies for network building and showing the location of edges has an important role in networks evolution. For example, Lescovec et al. [55] discovered that the network density increases bypower law over time.

Among the works, representing algorithmic tools for analyzing network evolution, can be identified Tantipathananandh C. et al. [86], the algorithms of estimation affiliation of user community and its change over time is proposed. The focus is on determining the approximate user clusters and evolution clusters.

Sometimes the analysis of network graphs development is convenient to conduct based on the paradigm of extracting association rules and analysis of frequency models. The rules of graph evolution, a new type of frequencymodelsareintroducedandtheproblemofsearchfortypicalmodelsofstructuralchangeindynamic networksareconsidered. Amore detailedoverviewofmodels andmethods ofsocialnetworks evolution can be found in the following sources [1, 13, 78].

Links prediction

The research aimed at identifying and predicting possible links between the vertices in the future are useful forinformationextractionofinterestfromsocialnetwork.Thelinksaredynamicandcanchangegreatlyover timeinmostapplicationsfortheanalysisofsocialnetworks.Thenetworkstructureandinformationaboutthe features of different vertices can be involved in processes of prediction links.

Thetaskoflinkpredictionconsistsindeterminingwhethertwoparticularverticesareconnectedtoeachother throughacertaintimeinterval.Thiscomputationalproblem,whichisbasedontheanalysisofevolutionsocial network in time, called the link prediction problem. To solve it, an automatic modeling of process of social network development with attraction of some network characteristics such as the number of common neighbors,theshortestpath,verticesinfluence,timeoffirstenterintoasocialnetworkisused.Tosolvethesetasks itisproposedtobuildavarietyofstructuralandrelationalmodels.Therearemodelsoflinkprediction,based onmachinelearning,usingpersonalinformationofnetworkuserstoimprovetheaccuracyofprediction[56]. Sometimes hierarchical, stochastic (Markov) and relational models to detect connections between users are used.

In other models [48] it is offered to take users properties as a basis, and, for example, the presence of many linksnumber(intheblogosphere)canbeexplainedbycomparingdemographicgroups,commoninterests orby geographicalproximity.Thereviewoflinkspredictionmodelsandmethodsispresentedinsources[3,67].

Clusteringtextinformationbasedonthefrequencyanalysis

Havingcollectedandclusteredtextdatafromthesocialnetwork, itispossibletoidentifythemaintopicsand events discussed socialnetworkingusersindifferentcitiesandcountries. Currently, there are many methods using which the problem of classification and clustering of texts can be solved. On this basis, many systems are implemented by semantic text processing.

Oneof themainmethodsoffrequencyanalysisiscountingthenumber ofoccurrencesofeachwordindocument.Basedonreceivedinformation,it ispossibletomakeso-called “tagcloud”– avisualrepresentationof the words weight in the document. Therefore, to evaluate a word weight correctly, it is necessary to use measuresthatwillnotonlycountthenumberofoccurrencesofwordinadocument,butalsotakeintoaccount thenumberofoccurrencesofwordsinotherdocuments.TheTF-IDFisanexampleofsuchmeasures[77].A measure of inductive Word2vec is often applied in the work of researchers dealing with cluster analysis of text information in various search engines [95, 98].

Theprincipleofactionmeasureistofindrelationsbetweenthewordcontextaccordingtotheassumptionthat the words in similar contexts have a tendency to mean similar things, i.e., be semantically close. Word2vec analyzes theusageof wordcontextsandconcludesthattheyarecloseinmeaningornot.Thealgorithmsthat underpin Word2vec, detailed in the works [65, 66].

Conclusion

This article presents a consist overview of main SNA methods, which are used for informationsecurity. The main types of social network threats, main trends in the field of social networking safety are considered by researches. The basic methods and algorithms of graph theory and data mining, which are used in the SNA are briefly described. Typical security tasks of SNA, such as community detection in network, the detection of leaders in communities, the detection experts in networks, clustering text information based on the frequency analysis and others are presented.

References

1. Aggarwal, C., Karthik, S. (2014). Evolutionary Network Analysis: A Survey. *ACM Computing Surveys*, *47*(1), Article 10.
2. Aggarwal, C. (2011). *Introduction to social network data analytics*. Springer US. Retrieved from doi: 10.1007/978-1-4419-8462-3
3. Ajay Kumar Singh Kushwah, Amit Kumar Manjhvar (2016). A Review on Link Prediction in Social Network. *International Journal of Grid and Distributed Computing*, *9*(2), 43-50.
4. Antcolonyoptimizationalgorithms.Retrievedfromhttps://en.wikipedia.org/wiki/Ant_colony_optimization_algorithms. Accessed 07 March 2017.
5. Batura, T.V. (2013). Modeli i metody analiza komp’yuternykh sotsial’nykh setey [Models and methods ofanalysisofcomputersocialnetworks].*InternationalJournalProgrammnyeProduktyiSistemy*,*3*,130- 137.
6. Bonchi,F.,Castillo,C.,Gionis,A.,Jaimes,A.(2011).SocialNetworkAnalysisandMiningforBusiness Applications, *ACM TIST*, *2*(3), 22-58.
7. Bonchi, F., Castillo, C., Jaimes, A. (2011). Social network analysis andmining forbusiness applications.*ACM TransIntell.Syst.Technol*,*2*(3),1-37.
8. Buzun,N.,Korshunov,A.(2012).*Vyiavlenieperesekayuschihsyasoobschestvvsocialnyhsetyah[Identifying overlappingcommunitiesinsocialnetworks]*.Moscow,18p.
9. Carley,K.,Lee,J.,Krackhardt,D.(2002).Destabilizingnetworks.*Connections*,*24*(3),79-92.

10. Carnes,T.,Nagarajan,R.,Wild,S.M.,VanZuylen,A.(2007).Maximizinginfluenceinacompetitivesocial network:afollower'sperspective.*Proceedingsoftheninthinternationalconferenceonelectroniccommerce*, Minneapolis, USA, 351-360.
11. Center for Computational Analysis of Social and Organizational Systems. Retrieved from http://www.casos.cs.cmu.edu/.Accessed07March2017.
12. Chakrabarti, S. (2002). *Mining the web. Discovery knowledge from hypertext data*. Morgan Kaufmann, 344 p.
13. Charu,C.(2012).*Socialnetworkdataanalytics*.SpringerScience&BusinessMedia,486p.
14. Churakov,A.N.(2001).Analizsocialnyhsetey[Analysisofsocialnetworks].*SocialStudies*,1,109-121.
15. Cortizo,J.,Carrero,F.,Gomez,J.,Monsalve,B.,Puertas,E.(2009).IntroductiontoMiningSM.*Proceedings ofthe1stInternationalWorkshoponMiningSM*,1-3.
16. Coscia,M.,Giannotti,F.,Pedreschi,D.(2011).Aclassificationforcommunitydiscoverymethodsincomplex networks.*StatisticalAnalysisandDataMining*,512-546.
17. Datamining.Retrievedfromhttps://en.wikipedia.org/wiki/Data_mining.Accessed07March2017
18. Dodds, P.S., Watts,D.J.(2005).Ageneralized modelof social andbiological contagion.*Journalof Theoretical Biology*, 4, 587-604.
19. Dodonov, A.G., Lande, D.V., Prischepa, V.V., Putyatin, V.G. (2013). *Konkurentnaya razvedka v komp'yuternykhsetyakh[Competitiveintelligenceincomputernetworks]*.Kyiv:IPRINASUkraine,248p.
20. Dodonov, A.G., Lande, D.V., Putyatin, V.G. (2009). *Ínformatsíyní potoki v global'nykh komp'yuternykh merezhakh[Informationflowsinglobalcomputernetworks]*.Kyiv: Naukova Dumka,295p.
21. Dodonov,A.G.,Lande,D.V.,Putyatin,V.G.(2014).*Komp'yuternyyesetiianaliticheskiyeissledovaniya [Computernetworksandanalyticalstudies]*.Kyiv:IPRINASofUkraine,486p.
22. Dodonov, O.G., Gorbachik, O.S., Kuznetsova, M.G. (2003). Informatsiynesuspil'stvo: tekhnolohii ta bezpeka [Information security: technology and security]. *Information and open government as a means of democratizationofsociety:Coll.material"roundtable"*,119-124.
23. Dokuka,S.V.,Valeeva,D.R.(2015).Statisticheskiemodelidlyaanalizadinamikisocialnyhseteyvissledovaniyahobrazovaniya[StatisticalModelsfortheAnalysisoftheDynamicsofSocialNetworksinEducation Studies].*EducationIssues*,1,201-213.
24. Dzyundzyuk, V.B. (2011). Virtual'ni spivtovarystva: potentsiynazahroza dlya natsional′noyi bezpeky [Virtual communities:potentialthreattonationalsecurity].*Statebuilding:Electronicpublication,1*.
25. Easley, D., Kleinberg, J.(2010). *Networks, Crowds, andMarkets: Reasoning about a Highly Connected World*.CambridgeUniversityPress,819p.
26. Ehrlich,K.,Carboni,I.(2005).InsideSocialNetworkAnalysisIBMWatsonResearchCenter.NewYork, USA, *Technical Report*, 5-10.
27. Flake,G.W.,Lawrence,S.,GilesC.L.,Coetzee,F.M.(2002).Self-organizationandidentificationofWeb communities. *Computer*, 3, 66-70.
28. Fortunato,S.(2010).Communitydetectioningraphs.*PhysicsReports*,*486*(3-5),75-174.
29. Free and Open Source Social Network Analysis Software. Retrieved from http://www.butleranalytics.com/20-free-and-open-source-social-network-analysis-software/.Accessed07March2017.
30. Girvan,M.,Newman,M.E.(2002).Communitystructureinsocialandbiologicalnetworks.*Proceedingsof theNationalAcademyofSciencesoftheUnitedStatesofAmerica*,12,7821-7826.
31. Goyal,A.,Bonchi,F.,Laks, Lakshmanan,V.S.(2008). Discoveringleadersfromcommunityactions. *Proceedingsofthe17thACMConferenceonInformationandKnowledgeManagement*,NapaValley,California, USA,499-508.
32. Greenhow,C.(2011).Onlinesocialnetworksandlearning.*OntheHorizon*,19(1),4-12.
33. Grigor'yevA.N.,Lande,D.V.,Borodenkov,S.A.,Mazurkevich,R.V.,Pats'ora,V.N.(2007).*InfoStream. Monitoring novostey iz Interneta: tekhnologiya, sistema, servis, [Monitoring of news from the Internet: technology,system,service]*.Kyiv:OOO"Start-98",40p.
34. Grinenko,I.,Prokof'eva-Yanchilenko,D.(2012).Vplyvvirtual'nykhspil'notnainformatsiynubezpeky: suchasnyystantatendentsiirozvytku[Influenceof virtual communitiesfor informationsecurity: current situationandtrends].*Legal,regulatoryandmetrologicalsupportofinformationsecurityinUkraine*,*1*(23), 18-23.
35. Gubanov, D.A., Novikov, D.A., Chartishvili, A.G. (2010). *Socialnye seti: modeli informacionnogo vliyaniya,upravleniyaiprotivobotstva[Socialnetworks:informationalinfluence,managementandcontention models]*. Monograph, Moscow, 228 p.
36. Gyöngyi, Z., Garcia-Molina, H., Pedersen, J. (2004). CombatingWeb Spam with TrustRank. *Proceedings of theInternationalConferenceonVeryLargeDataBases*,*30*,576-587.

37. Hanneman, R. (1988). *Computer-Assisted Theory Building: Modeling Dynamic Social Systems*. Riverside, University of California, 343 p.
38. Hanneman, R., Riddle,M.(2005).*Introductiontosocial networkmethods*.Riverside, CA: UniversityofCalifornia. Retrieved from http://faculty.ucr.edu/~hanneman/nettext.Accessed07March 2017.
39. Hoppe,B.,Reinelt,C.(2010).SocialNetwork.AnalysisandtheEvaluationofLeadershipNetworks.*The Leadership Quarterly*, 4, 600-619.
40. Horbulin,V.P.,Dodonov,O.H.,Lande,D.V.(2009).*Informatsiynioperatsiyitabezpekasuspil'stva:zahrozy, protydiya,modelyuvannya: monohrafiya[Informationoperationsandsafetyof society:thethreat,resistance, modeling:monograph]*.Intertekhnolohiya,164p.
41. InternationalNetworkforSocialNetworkAnalysis.Retrievedfromhttp://www.insna.org/#.Accessed07 March2017.
42. Jackson,MatthewO.(2010).AnOverviewofSocialNetworksandEconomicApplications.*Handbookof Social Economics*. Retrieved from https://web.stanford.edu/~jacksonm/socialnetecon-chapter.pdf. Accessed 07 March 2017.
43. Johnson,J.,Ironsmith,M.(1994).AssessingChildren'sSociometricStatus: IssuesandtheApplicationofSocialNetworkAnalysis.*JournalofGroupPsychotherapy,Psychodrama&Sociometry*,*47*(1),36-49.
44. Kempe,D.,Kleinberg,J.,Tardos,É.(2003).Maximizingthespreadofinfluencethroughasocialnetwork. *ProceedingsoftheninthACMSIGKDDinternationalconferenceonknowledgediscoveryanddatamining*, Washington, USA, 137-146.
45. KeyTrendstoWatchinGartner2012EmergingTechnologiesHypeCycle.Retrievedfrom http://www.forbes.com/sites/gartnergroup/2012/09/18/key-trends-to-watch-in-gartner-2012-emerging-technologies-hype-cycle-2.Accessed07March2017.
46. Kolomeychenko,M.I.,Chepovskiy,A.A.,Chepovskiy,A.M.(2014).Algoritmvydeleniyasoobschestvvsocialnyhsetyah [Analgorithm fordetectingcommunitiesinsocial networks], *Fundamental andAppliedMathematics*, *19*(1), 21-32.
47. Koren,Y.(2003).OnSpectralGraphDrawing.*Proceedingsofthe9thInternationalComputingandCombinatoricsConference*,Springer,496-508.
48. Kumar,R.,Novak,J.,Raghavan,P.,Tomkins,A.(2004).StructureandEvolutionofBlogspaceCommun. *ACM*,*47*(12),35-39.
49. Lande,D.V.(2005).*PoiskznaniyvInternet.Professional'nayarabota[SearchofknowledgeontheInternet. Professionalwork]*.Moscow:Dialectics,272p.
50. Lande,D.V.(2006).*Osnovyintegratsiiinformatsionnykhpotokov:monografiya[Fundamentalsoftheintegrationofinformationflows:monograph]*.Kyiv:Engineering,240p.
51. Lande, D.V. (2014). *Elementy komp'yuternoyi linhvistyky v pravoviyi nformatytsi [Elements of Computational LinguisticsintheLegalInformation]*.Kyiv:NDIIPNAPrNUkraine,168p.
52. Lande,D.V.,Furashev,V.N.,Braichevsky,S.M.,Grigoriev,A.N.(2006).*OsnovymodelirovaniyaIotsenki elektronnykh informatsionnykh potokov: monografiya [Fundamentals ofmodeling and evaluation of electronic informationflows:monograph]*.Kyiv:Engineering,176p.
53. Lande,D.V.,Snarskii,A.A.,Bessudnov,I.V.(2009).*Internetika:Navigatsiyavslozhnykhsetyakh:modelii algoritmy[Internetics:Navigationincomplexnetworks:modelsandalgorithms]*.Moscow:TheLibricom Book House, 264 p.
54. Langville,AmyN.,Meyer,CarlD.(2006).*Google'sPageRankandBeyond:TheScienceofSearchEngine Rankings*.PrincetonUniversityPress,224p.
55. Leskovec,J.,Kleinberg,J.,Faloutsos,C.(2005).Graphsovertime:densificationlaws,shrinkingdiameters andpossibleexplanations.*Proc.11thACMSIGKDDIntern.Conf. onKnowledge DiscoveryinDataMining*, NY,177-187.
56. Liben-Nowell,D.,Kleinberg,J.(2003).TheLinkPredictionProblemforSocialNetworks.*Proceedingsofthe 12thInternationalConferenceonInformationandKnowledgeManagement*,NY:ACMPress,556-559.
57. Listofsocialnetworkingwebsites.Retrievedfromhttps://en.wikipedia.org/wiki/List_of_social_networking_websites. Accessed07 March2017.
58. Mona Jalal, An Hay Doan. A Survey on Community Mining in Social Networks. Retrieved from http://monajalal.github.io/assets/pdf/CS784_report.pdf.Accessed07March2017.
59. MariamAdedoyin-Olowe,MohamedMedhatGaber,FredericStahl(2014).ASurveyofDataMiningTechniquesforSocialNetworkAnalysis.*JournalofDataMining&DigitalHumanities*,1-25.
60. Mason,W.A.,Conrey,F.R., Smith,E.R. (2007).Situatingsocialinfluenceprocesses:dynamic, multidirectionalflowsofinfluencewithinsocialnetworks.*OfficialjournalofSocietyforPersonalityandSocialPsychology*, *11*, 279-300.

61. MatthewA.Russell(2011).*MiningtheSocialWeb:AnalyzingDatafromFacebook,Twitter,LinkedIn,and OtherSocialMediaSites*.O'Reilly,332p.
62. Matvienko, Yu.A. Destruktivnyye setevyyesotsial'nyyestrukturykaksredstvoinformatsionnoyvoynyi ugrozabezopasnostiRossii[Destructivenetworksocialstructuresasameansofinformationwarfareanda threat to Russia's security]. Retrieved from http://old.geopolitica.ru/Articles/1218.Accessed 07 March 2017.
63. Metodyklassifikatsiiprognozirovaniya.Metodopornykhvektorov.Metod"blizhayshegososeda".Bayyesovskayaklassifikatsiya[Methodsofclassificationandforecasting.Supportvectormethod.The"nearest neighbor" method. Bayesian classification]. Retrieved fromhttp://www.intuit.ru/studies/courses/6/6/lecture/176. Accessed 07 March 2017.
64. Metodyklasternogoanaliza.Iyerarkhicheskiyemetody[Methodsofclusteranalysis.Hierarchicalmethods]. Retrievedfromhttp://www.intuit.ru/studies/courses/6/6/lecture/182.Accessed07March2017.
65. Mikolov, T., Chen,K., Corrado, G., Dean, J.(2013). Efficient Estimation ofWord RepresentationsinVector Space.Retrievedfromhttps://arxiv.org/pdf/1301.3781.pdf.Accessed07March2017.
66. Mikolov, T., Sutskever, I., Chen,K., Corrado, G., Dean, J.(2013). DistributedRepresentationsofWords and Phrases and their Compositionality. Retrieved from https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf.Accessed07March2017
67. MohammadAlHasan,MohammedJ.Zaki(2011).ASurveyofLinkPredictioninSocialNetworks.*Social Network Data Analytics*, 243-275.
68. NationalCyberSecurityStrategy2016-2021forUnitedKingdom.Retrievedfromhttps://www.gov.uk/government/uploads/system/uploads/attachment_data/file/564268/national_cyber_security_strategy.pdf.Accessed 07 March 2017.
69. National cyber security strategy for Ukraine. Retrievedfrom http://www.president.gov.ua/documents/962016-19836. Accessed 07 March 2017.
70. NetLab.Retrievedfromhttp://groups.chass.utoronto.ca/netlab/.Accessed07March2017.
71. Orgnet:SocialNetworkAnalysisSoftware&ServicesforOrganizations,Communities,andTheirConsultants.Retrievedfromhttp://www.orgnet.com/index.html.Accessed07March2017.
72. PaloAlto Networks.Top10socialnetworkingthreats. Retrievedfromhttp://www.networkworld.com/article/2213704/collaboration-social/top-10-social-networking-threats.html.Accessed07March2017.
73. Peleshchyshyn,A.M.Syerov,Yu.O.,Berezko,O.L.,Peleshchyshyn,O.P.,Tymovchak-Maksymets',O.Yu., Markovets,O.V.(2012).*Protsesyupravlinnyainteraktyvnymysotsial'nymykomunikatsiyamyvumovakh rozvytkuinformatsiynohosuspil'stva:monohrafiya[Managementprocessesinteractivesocialcommunication inthedevelopment of the informationsociety: monograph].* Lviv, LvivPolytechnic National UniversityPublishing House, 368 p.
74. Peleshchyshyn,O.P.(2013).Analiztaprotydiyazahrozammarketynhoviypozytsiyipidpryyemstvavonlaynspil'notakh [Analysis andresistance threats marketingpositionof the companyin the online community]. *Information Security*, *3*(15), 217-224.
75. Peleshchyshyn,O.P.(2010).Informatsiynitekhnolohiyioblikutaposhukuonlayn-spil'notuzadachisotsial'nohomarketynhu[IndustryAccountingandsearchonlinecommunitiesinthetaskofsocialmarketing]. *ProceedingsoftheNationalUniversity"LvivPolytechnic",Economic*Series,*44*,50-59.
76. Principalcomponentanalysis.Retrievedfromhttps://en.wikipedia.org/wiki/Principal_component_analysis. Accessed 07 March 2017.
77. Ramos,J.(2003).Usingtf-idftodeterminewordrelevanceindocumentqueries.*Proceedingsofthefirstinstructionalconferenceonmachinelearning*.
78. SaoussenAouay,SalmaJamoussi,FaiezGargouri,AjithAbraham(2014).ModelingDynamicsofSocialNetworks:ASurvey.*SixthInternationalConferenceonComputationalAspectsofSocialNetworks(CASoN)*,49- 53.
79. Shantanu Ghosh (2011). Top seven social media threats. Retrieved from http://www.computerweekly.com/tip/Top-seven-social-media-threats. Accessed07March2017.
80. Slabchenko,O.O.,Sidorenko,V.N.,Ponomarchuk,R.A.(2013).Metodyialgoritmyvyiavleniyasoobschestv potencialnyh abiturientov iih liderov v socialnyh setyah [Methods and algorithms foridentifying communities ofpotentialapplicantsandtheirleadersinsocialnetworks].*BulletinofNationalUniversityofKremenchuk*, *1*(78),53-61.
81. Smirnov,A.I.,Grigoriev,V.R.,Kokhtyulin,I.N.,Kuroyedov,B.V.,Sandarov,O.V.(2014).*Global'nayabezopasnost'v tsifrovuyuepokhu:stratagemdlya Rossii[Globalsecurityinthe digitalage:stratagems for Russia]*.Moscow:All-RussianResearchInstituteofGeosystems,394p.
82. Social network analysissoftware.Retrieved from https://en.wikipedia.org/wiki/Social_network_analysis_software.Accessed 07 March 2017.

83. Softwareforsocialnetworkanalysis.Retrievedfromhttps://www.gmw.rug.nl/~huisman/sna/software.html.Accessed 07 March 2017.
84. Statisticsandfactsaboutsocialmediausage.Retrievedfromhttps://www.statista.com/topics/1164/social-networks/. Accessed 07 March 2017.
85. Stohl, C., Stohl, M. (2007). Networksof Terror: Theoretical Assumptions and Pragmatic Consequences.*Communication Theory*, *17*, 93-124.
86. Tantipathananandh,C.,Berger-Wolf,T.,Kempe,D.(2007).Aframeworkforcommunityidentificationin dynamicsocialnetworks.*Proc.13thACMSIGKDDIntern.Conf.onKnowledgeDiscoveryandDataMining,* NY,717–726.
87. TheGlobalRisksReport2017,12thEdition.Retrievedfromhttp://www3.weforum.org/docs/GRR17_Report_web.pdf.Accessed07 March2017.
88. The International SurveyCenter conducts research on social, economic and political issues usingsurveydata fromlarge,representativenationalsamplesfrommanynations.Retrievedfromhttp://internationalsurvey.org/.Accessed 07 March 2017.
89. TheNATOCooperativeCyber Defence Centre. Retrievedfromhttps://ccdcoe.org/cyber-security-strategy-documents.html.Accessed07March2017.
90. Top 15 Most Popular Social Networking Sites (and 10 Apps!). Retrieved from https://www.dreamgrow.com/top-15-most-popular-social-networking-sites/.Accessed07March2017.
91. Top 30 Social Network Analysis and Visualization Tools. Retrieved from http://www.kdnuggets.com/2015/06/top-30-social-network-analysis-visualization-tools.html.Accessed07March2017.
92. Travellingsalesmanproblem.Retrievedfromhttps://en.wikipedia.org/wiki/Travelling_salesman_problem.Accessed 07 March 2017.
93. Tymovchak-Maksymets,O.(2010).Metodyvykorystannyarozshyrenykhmozhlyvosteyhlobal'nykhposhukovykhsystemvzadachiposhukuspozhyvats'kohodosviduvonlaynseredovyshchakh[Methodsofusing theadvancedfeaturesofglobalsearchenginesofsearchconsumerexperienceinonlinemedia].*Proceedings oftheNationalUniversity"LvivPolytechnic":InformationSystemsandNetworks*,*689*,323-331.
94. Ukustov,S.S.,Kravets,A.G.(2012).Podkhodkresheniyuzadachiidentifikatsiivliyatel'nykhrazrabotchikov vsotsial'noysetigitkhab[Approachtosolvingtheproblemofidentifyinginfluentialdevelopersinthesocial networkofgithub].*ProceedingsoftheVolgogradStateTechnicalUniversity*,*15*(102),61-66.
95. WangH.(2014).IntroductiontoWord2vecanditsapplicationtofindpredominantwordsenses.Retrieved fromhttp://compling.hss.ntu.edu.sg/courses/hg7017/pdf/word2vec%20and%20its%20application%20to%20wsd.pdf.Accessed 07 March 2017.
96. Watts,D.J.,Dodds,P.S.(2007). Influentials,Networks, andPublicOpinionFormation. *Journal of Consumer Research*, 4, 441-458.
97. WhitePaperonGermanSecurityPolicyandthe FutureoftheBundeswehr. Retrievedfromhttps://www.bundeswehr.de/resource/resource/MzEzNTM4MmUzMzMyMmUzMTM1MzMyZTM2MzIzMDMwMzAzMDMwMzAzMDY5NzE3MzM1Njc2NDYyMzMyMDIwMjAyMDIw/2016%20White%20Paper.pdf.Accessed 07 March 2017.
98. Yu,M.,Dredze,M.(2014).Improvinglexicalembeddingswithsemanticknowledge. *AssociationforComputationalLinguistics (ACL)*,545-550.