AZƏRBAYCAN RESPUBLİKASININ PREZİDENTİ YANINDA
ELMİN İNKİŞAFI FONDU

QRANT LAYİHƏLƏRİ MÜSABİQƏSİ

Kompleks elmi-tədqiqat proqramları layihələri müsabiqəsi EİF- KETPL-2015-1(25)	LAYİHƏNİN MƏZMUNU VƏ ƏSASLANDIRILMASI

LAYİHƏNİN ADI:

Böyük verilənlər (“big data”) mühitində informasiya təhlükəsizliyinin təmin olunması metodları və alqoritmlərinin işlənilməsi və onların bəzi tətbiqləri

1. Layihənin məqsədi, qarşıya qoyulan məsələləri, aktuallığı, kompleks proqram və multidissiplinar xarakterinin əsaslandırılması

Layihənin məqsədi böyük verilənlər (“big data”) mühitində (erasında) milli informasiya təhlükəsizliyi sisteminin təkmilləşdirilməsi üçün metodların, yanaşmaların və alqoritmlərin işlənilməsi və bir sıra tətbiqlərinin reallaşdırılmasıdır. Bu məqsədə çatmaq üçün aşağıdakı məsələlərin həlli nəzərdə tutulur:

- Müxtəlif Informasiya təhlükəsizliyi obyektlərində toplanmış böyük həcmli verilənlərdə anomaliyaların aşkarlanması üçün metod və alqoritmlərin işlənilməsi;
- Milli informasiya infrastrukturuna olan xidmətdən kütləvi imtina (DDoS – Distributed Denial of Service) hücumların aşkarlanması üçün metod və alqoritmlərin işlənilməsi;
- Sosial medianın analizi əsasında milli təhlükəsizliyə təhdidlərin və cəmiyyətdə anomal proseslərin aşkarlanması üçün text mining (mətnlərin intellektual analizi) yanaşmalarının təklif edilməsi;
- Hədəfyönümlü hücumların aşkarlanması və analizi üçün modellərin işlənilməsi;
- Məxfiliyi təmin etməklə fərdi məlumatların intellektual analizi üçün metod və alqoritmlərin işlənilməsi;
- Təklif olunmuş metod və alqoritmlərin korporativ şəbəkələrdə tətbiqi.

Hazırda insan fəaliyyətinin müxtəlif sferalarında böyük həcmli verilənlər (Big Data) istehsal olunur. Məlumdur ki, onlar böyük informasiya və bilik mənbəyidirlər. Lakin bu cür böyük həcmli verilənlərin analizi və onlardan yeni informasiya və biliklərin əldə olunması müasir elmin qarşısında aktual və həlli çətin olan multidissiplinar problemlər qoymuşdur.

Strateji əhəmiyyətinə görə Big Data-nı yeni iqtisadiyyatın, başqa sözlə biliklər iqtisadiyyatının nefti adlandırırlar. Dünya İqtisadi Forumunun 2011-ci ildə Davosda keçirilən toplantısında "Big Data yeni iqtisadiyyatın neftidir" tezisi qəbul edilmişdir. Qəbul edilmiş sənəddə Big Data xam neft kimi dəyərləndirilir. Big Data qiymətlidir, lakin o da xam neft kimi emal edilməzsə, istifadə oluna bilməz. Onun dəyər əldə etməsi üçün analiz olunmalıdır. Dünya İqtisadi Forumunun 2012-ci ildə keçirilən iclasında isə Big Data valyuta və ya qızıl dəyərində yeni iqtisadi aktiv kimi dəyərləndirilmiş, beynəlxalq inkişaf üçün Big Data-nın imkanlarını müzakirə edən sənəd hazırlanmışdır.

Dünyanın bir çox inkişaf etmiş ölkələri – ABŞ, Böyük Britaniya, Fransa, Yaponiya, Çin və Koreya gələcəkdə davamlı və dayanıqlı inkişafı təmin etmək üçün özlərinin Big Data strategiyalarını müəyyənləşdirmiş və müvafiq sənədlər qəbul etmişlər. Məsələn, 2012-ci ilin martında ABŞ Prezident Administrasiyası hökumətin müdafiə, kibertəhlükəsizlik, səhiyyə, energetika, ekologiya, elm və yüksək texnologiyalar sahəsində üzləşdiyi problemlərin həllində Big Data-dan istifadənin rolunu tədqiq etmək məqsədilə "Big data tədqiqat və inkişaf təşəbbüsü" adlı sənəd qəbul etmiş və onun icrası üçün ilkin olaraq 200 milyon dollar vəsait ayırmışdır.

Bir sıra ölkələrin informasiya təhlükəsizliyi strategiyalarında kiber hücumlar, mütəşəkkil kiber cinayətkarlıq və kiber terrorizm milli təhlükəsizliyə əsas təhdidlər kimi xarakterizə olunur. Aydındır ki, ənənəvi informasiya təhlükəsizliyi sistemlərinin bu təhdidlərin vaxtında aşkarlanması və qarşısının alınması sahəsində problemləri mövcuddur. Burada əsas problem informasiya təhlükəsizliyi siyasətinin pozulması hallarının aşkarlanması zamanı mövcud texnologiyaların bütün əlyetən informasiya mənbələrindən müxtəlif tipli məlumatları əldə etməsi, saxlaması və emal etməsi imkanlarının məhdud olmasıdır.

Layihənin kompleks xarakteri bir neçə cəhətdə özünü göstərir:

- Qarşıya qoyulan məsələlərin həlli bir-birilə əlaqəli kompleks elmi-nəzəri və tətbiqi xarakterli tədqiqatların aparılmasını nəzərdə tutur
- Big Data mühitində informasiya təhlükəsizliyinin təmin edilməsi kompleks yanaşma tələb edir.
- Böyük həcmli və müxtəlif təbiətli verilənlərin kompleks emalı müxtəlif profilli mütəxəssislərin sıx əməkdaşlığını tələb edir.

Verilənlərin həcminin eksponensial artımı elmdə yeni bir istiqamətin – verilənlər haqqında elmin (Data Science) və yeni mütəxəssislərin – verilənlərin emalı üzrə alimlərin (Data Scientist) meydana gəlməsinə səbəb olmuşdur. Data Science verilənlərin analizi, emalı və təqdimatı (vizuallaşdırılması) problemləri ilə məşğul olur. Verilənlər haqqında elmə bəzən Datalogiya da deyirlər. Data Science bir neçə elmin – Kompüter Elmlərinin, Tətbiqi Riyaziyyat və Riyazi Statistikanın və tədqiq olunan elmi istiqamətin qovşağında olan elm sahəsidir. Verilənlərin emalı üzrə alimlər yuxarıda sadalanan elm sahələri üzrə biliklərə malik universal mütəxəssislərdir.

Analiz göstərir ki, layihə çərçivəsində qarşıya qoyulan informasiya təhlükəsizliyi problemlərinin həlli multidissirplinar yanaşma tələb edir və bunun üçün big data, data mining, paralel və paylanmış hesablamalar, cloud computing, sosial şəbəkə analizi, text mining, web mining, optimallaşma, süni intellekt, riyazi statistika və ehtimal nəzəriyyəsi və s. sahələrin metod və texnologiyalarından geniş istifadə olunmalıdır.

Aparılacaq tədqiqatların kompleks xarakterini və multidissiplinarlığını nəzərə alaraq layihənin işlənilməsinə bir neçə tədqiqat istiqamətini təmsil edən aparıcı tədqiqatçılar cəlb edilmişdir.

2. Layihənin annotasiyası

Müasir dövrdə verilənlərin mənbələrinin, həcminin, sürətinin, etibarlılığının və dəyərinin artması (big data) şəraitində informasiya təhlükəsizliyinin təmin edilməsi ən aktual problemlərdən biridir. Bu problemin həlli üçün layihədə big data analitikası, maşın təlimi, paylanmış və paralel hesablamalar, sosial media analitikası və data mining metodları və alqoritmlərinin təklif edilməsi nəzərdə tutulur. Layihə çərçivəsində aşağıdakı nəticələrin əldə olunması nəzərdə tutulur:

- Sosial mediada milli informasiya təhlükəsizliyinə təhdidləri aşkarlamaq üçün effektiv data mining metodlarının təklif edilməsi;
- Big Data analitikası texnologiyalarının imkanlarından istifadə etməklə, normal və anomal davranış profillərinin aşkarlanması üçün metod və alqoritmlərin işlənilməsi;
- Milli informasiya infrastrukturuna olan DDoS hücumların və hədəfyönümlü hücumların erkən mərhələdə aşkarlanması üçün big data analitikası yanaşmalarının təklif edilməsi;
- Böyük həcmli fərdi məlumatların intellektual analizi zamanı məxfiliyi təmin etmək üçün metodlar və alqoritmlər.

Alınacaq nəticələr milli informasiya təhlükəsizliyi sisteminin təkmilləşdirilməsi, milli informasiya fəzasının monitorinqi, analizi və effektiv qərarların qəbul edilməsi zamanı istifadə edilə bilər. Bu nəticələr eyni zamanda böyük həcmli verilənlərin analizini tələb edən sahələrdə, məsələn, ekologiyada, tibbdə, maliyyə-bank sistemində, genetikada, molekulyar biologiyada, əczaçılıqda, geologiya və geofizikada və s. istifadə oluna bilər. Əldə edilmiş elmi nəticələr həmçinin tədrisdə Data Science (Verilənlər haqqında elm) sahəsində kadrların (Data Scientist – verilənlərin emalı üzrə mütəxəssis) hazırlanmasında istifadə oluna bilər.

3. Layihənin məzmununu tam əks etdirən açar sözlər və ya söz birləşmələri

Big data; informasiya təhlükəsizliyi; anomaliyaların aşkarlanması; data mining; DDoS (Distributed Denial of Service – xidmətdən kütləvi imtina); Big Data analitika; PPDM (Privacy-preserving data mining – məxfiliyi təmin etməklə verilənlərin intellektual analizi); maşın təlimi; şəbəkə monitorinqi; kiberhücum

4. Layihənin elmi istiqaməti və qarşıya qoyulan problem üzrə qısa icmal

Anomaliyaların aşkarlanması. Anomaliyaların aşkarlanmasının məqsədi normal davranışların əvvəlcədən müəyyən edilmiş çoxluğundan kənara düşən istənilən hadisəni hədəf edir. Anomaliyaların aşkarlanması proqramları fərz edir ki, istənilən bir müdaxilə hadisəsi anomal fəaliyyətin altçoxluğudur. Bu baxımdan, o, sui-istifadənin aşkarlanmasından açıq-aşkar fərqlənir, sonuncuda hücumları göstərmək üçün əvvəlcə anormal davranışın siqnaturası müəyyən edilir. Anomaliyaların aşkarlanmasında əvvəlcə kiber infrastrukturun sağlamlıq və həssaslığını əks etdirən normal davranış profili müəyyən edilir. Müvafiq olaraq, verilənlərdə gözlənilən normal davranışa uyğun gəlməyən istənilən şablon anomaliya kimi müəyyən edilir [4, 27].

Yeni hücumlar meydana çıxdıqda və normal davranışlar olduğu kimi qaldıqda anomaliyaların aşkarlanması yeni və ya qeyri-adi hücumları aşkarlaya və erkən xəbərdarlıq edə bilər. Sui-istifadənin aşkarlanması kimi, anomaliyaların aşkarlanması da normal və anomal davranış profilləri arasında aydın sərhədə əsaslanır, bu sərhəd normal davranış profilini anomal hadisələrdən fərqləndirməyə xidmət edir.

Maşın təlimi üsulları anomaliyaların aşkarlanması sistemlərində normal profilin qurulması və müdaxilələrin aşkarlanmasında əsas rol oynayır. Anomaliyaların aşkarlanmasında normal davranışa uyğun nişanlanmış verilənlər adətən əlyetər olur, anomal davranışa uyğun verilənlər isə mövcud olmur. Öyrədilən maşın təlimi metodlarına hücum olmayan təlim verilənləri lazımdır. Lakin real şəbəkə mühitində bu cür təlim verilənlərini əldə etmək çətindir. Belə təlim verilənlərinin olmaması maşın öyrənməsində tanınmış verilənlərin balanslı olmayan paylanmasına gətirib çıxarır. Bundan başqa, dəyişən şəbəkə və ya xidmətlər mühitində normal profil nümunələri də dəyişəcək. Təlim və test verilənləri arasında belə fərqlər müdaxilələrin öyrədilən aşkarlanması sistemlərində yüksək yalnış-pozitiv faizlərə gətirib çıxarır. Öyrədilməyən anomaliya aşkarlanması sistemləri öyrədilən analoji sistemlərin nöqsanlarını aradan qaldıra bilər. Buna görə, yarım öyrədilən və öyrədilməyən maşın təlimi üsulları tez-tez istifadə olunur [10].

DDoS hücumlar. Adından da göründüyü kimi, DDoS (Distributed Denial of Service – xidmətdən kütləvi imtina) hücumlarında, seçilmiş hədəfə yönəlmiş trafik

generasiya edən çox böyük sayda paylanmış hostlar (kompüterlər) iştirak edir [7, 9]. Hücum vektorlarını iki qrupa bölmək olar: gücə əsaslanan hücumlar və semantik hücumlar. "Sel hücumları" kimi də tanınan gücə əsaslanan hücumlar şəbəkənin buraxma zolağının zəbt olunmasına yönəlmişdir, çox böyük sayda saxta sorğular generasiya etməklə çox saylı paylanmış hostlardan trafiki birləşdirərək hədəfi çökdürür. Bu hədəflər tətbiqi proqramlar, serverlər (hostlar) və ya infrastruktur ola bilər. Məsələn, tətbiqi proqramlara gücə əsaslanan hücuma nümunə kimi provayderin şəbəkəsində bütün elementlərə çox böyük sayda SSH (Secure Shell – təhlükəsiz örtük) giriş cəhdləri ola bilər. Gücə əsaslanan hücum öz paylanmış təbiətinə görə kütləvi ola bilər.

DDoS hücumların aşkarlanması alqoritmlərini iki əsas qrupa təsnif etmək olar: siqnaturaya əsaslanan və davranışa əsaslanan. Siqnaturaya əsaslanan aşkarlama üsulunda əldə edilən trafik əvvəlcədən müəyyən edilmiş hücum nümunələri ilə müqayisə edilir. Bu üsul hücum edənlərlə onların "zombi" kompüterləri arasında kommunikasiyanı aşkarlamaq üçün faydalı ola bilər. Kommunikasiya şifrləndikdə bu üsul səmərəsizdir. Davranışa əsaslanan yanaşmanın əsas ideyası trafik şablonları əsasında trafik üçün nəyin normal davranış olmasını müəyyən etməkdir. Normal davranışlardan hər hansı bir sapma bədniyyətli hesab edilə bilər. Ümumiyyətlə, sapmaları daha asan konfiqurasiya və müxtəlif şərtlərə adaptasiya etmək üçün sərhədlər qiymətləri müəyyən edilir [14, 19, 20, 21].

Sosial medianın analizi. Sosial media istifadəçilərinin spektri olduqca müxtəlifdir. Adi istifadəçilər sosial mediadan ünsiyyət, tanışlıq, gündəlik həyata aid məlumatların, şəkillərin paylaşımı vasitəsi kimi yararlanırlar. Sosial medianın səmərəli əks əlaqə imkanları onu əlverişli kommunikasiya və təsir kanalına çevirir. Son dövrlər dövlət hakimiyyəti orqanları, siyasi partiyalar, vətəndaş cəmiyyəti institutları, özəl sektor sosial medianın bu potensialından geniş istifadə etməyə çalışırlar. Sosial media özü ilə bir sıra təhlükələr də gətirir. Bu təhlükələr fərdlərə, sosial qruplara, bütövlükdə dövlətə və cəmiyyətə yönələ bilər. Sosial şəbəkələrin fərdlərə yönəlik təhlükələri barədə elmi ədəbiyyatda müfəssəl məlumat verilir və konkret tövsiyələr təklif edilir. Son illərdə sosial medianın milli təhlükəsizliyə təhdidlər yarada biləcəyi narahatlıqları bütün dünya ölkələrində dövlət hakimiyyəti orqanlarının nümayəndələri tərəfindən dəfələrlə bəyan edilir. Burada müxtəlif risk ssenariləri mümkündür – terrorçular tərəfindən sosial medianın geniş istifadə edilməsi, xarici qüvvələr tərəfindən ölkənin daxili siyasətinə təsir aləti kimi istifadə edilməsi və s. Bu narahatlıqların təcrübi əsası

da var və “Ərəb baharı” sübut etdi ki, sosial media kütlələri yönləndirmək, hadisələri dramatik şəkilləndirmək, sosial dəyişikliklər, inqilablar etmək üçün güclü silahdır [3, 8, 11, 15, 16, 17, 18, 22, 26, 28].

Dövlət hakimiyyəti orqanlarının informasiya siyasətinin əsas məqsədləri vətəndaşları öz fəaliyyətləri haqqında məlumatlandırmaq və kütləvi kommunikasiya vasitələrinin köməyi ilə vətəndaşlarla əks əlaqəni təşkil etməkdir. Eyni zamanda, dövlət orqanları münaqişə, sosial gərginlik yarada bilən, yanlış ictimai rəy formalaşdıran, hakimiyyət orqanlarının nüfuzuna ziyan vura bilən informasiya təhdidlərinə operativ reaksiya verməyə borcludurlar.

Hədəfyönümlü hücumlar. Ənənəvi kiber-hücumlardan fərqli olaraq hədəfyönümlü hücumların əsas məqsədi kiber infrastruktura ziyan vurmaq deyil, qiymətli verilənlərə çıxış əldə etmək və onu mümkün olduqca uzun müddət – aylar, hətta illər ərzində saxlamaqdır. Qiymətli verilənlər dedikdə təşkilatın intellektual mülkiyyəti (proqram məhsulunun ilkin kodları, alqoritmlər, müştərilər bazası, istənilən digər korporativ sirlər) nəzərdə tutulur. Hədəfyönümlü hücumların böyük sinfini APT (Advanced Persistent Threat – Təkmil Davamlı Hücum) hücumları təşkil edir. Termin xarici kəşfiyyat/dövlət tərəfindən dəstəklənən hücumları təsvir etmək üçün 2006-cı ildə ilk dəfə ABŞ hərbçiləri tərəfindən işlədilmişdi [2, 6, 12, 13].

APT təhdidlərin xüsusi növüdür, konkret dövlət strukturlarına, şirkətlərə, təşkilatlara və hətta şəxslərə yönəlir. Bu növ təhdidlərin əsas fərqli cəhəti yalnız onların hədəflərinin konkretliyi deyil, həm də qabaqcıl informasiya texnologiyaları ilə yanaşı, psixologiya, sosial mühəndislik və s. metodlarına əsaslanan ən müasir yanaşmalardan istifadə etməsidir.

Fərdi məlumatların təhlükəsizliyi. Big data şəxsi həyatın gizliliyinə, vətəndaş azadlıqlarının pozulmasına potensial təhdidlər yaradır, dövlət və korporativ nəzarət imkanlarını artırır. Şirkətlərin marketinq məqsədləri üçün Big Data analitikasından istifadə edərək şəxs barəsində gizli məlumatlar əldə edə bilərlər. Eynilə, analitika üçün verilənlərin anonimləşdirilməsi istifadəçi məxfiliyini qorumaq üçün kifayət deyil. Buna görə də, fərdi məlumatları intellektual analizi zamanı məxfiliyinin pozulması hallarının qarşısını almaq üçün müvafiq yanaşmalar, metodlar və texnologiyaların işlənilməsi vacibdir [1, 5].

PPDM (Privacy-preserving data mining – məxfiliyi təmin etməklə verilənlərin intellektual analizi) yanaşmasının məqsədi data mining və ya maşın təlimi metodlarının tətbiqi ilə alınmış məxfi informasiyaya icazə olmayan istifadəçilərin

girişinin qarşısını almaqdır. Tədqiqatçılar məxfiliyi qorumaq üçün PPDM-də data mining və maşın təlimi alqoritmlərində bir çox üsullardan istifadə edirlər [1, 23].

Mövcud PPDM üsulları altı prosedurdan ibarətdir: məxfiliyi saxlamaq üçün ilkin verilənlərin modifikasiyası, verilənlərin toplanması, məxfiliyi saxlamaq üçün aqreqasiya verilənlərinin modifikasiyası, PPDM alqoritmləri, konkret fərdi verilənlər üçün mining nəticələrinin rekonstruksiyası və PPDM nəticələrinin qiymətləndirilməsi. İlkin verilənlərin modifikasiya edilməsi fərdi verilənlərdə həssas məlumatların açıqlanmasının və ya fərdlərin məxfiliyinin pozulmasının qarşısını almağa xidmət edir. Geniş istifadə edilən data mining və ya maşın təlimi üsullarından fərqli olaraq, PPDM giriş verilənlərinin modifikasiya edilməsini tələb edir [24].

PPDM alqoritmlərinin əksəriyyəti nəzəri olaraq təklif edilib və onların yalnız kiçik bir qismi real praktiki situasiyalar üçün realizə edilmişdir və ya real verilənlərdən istifadə edilərək test edilmişdir. Bu isə onların istifadəçilərə təmin edəcəyi təhlükəsizlik səviyyəsini birqiymətli müəyyən etməyi çətinləşdirir. Lakin analiz edilmiş PPDM üsulları əsasında bir neçə tədqiqat istiqaməti təklif etmək olar. Birincisi, intellektual analizin dəqiqliyi ilə PPDM alqoritmlərin məxfiliyi pozma səviyyəsini balanslaşdırmaq üçün optimal həllər axtarmaq lazımdır. İkincisi, çoxtərəfli kriptoqrafik hesablamalarda etibarlı üçüncü tərəf axtarışı həlli vacib problemlərdindir. İştirakçılar tez-tez narahat olurlar ki, üçüncü tərəf özünə fayda saxlamaq üçün icazəsiz tərəflərlə əlaqələrə şirniklənə bilər. Üçüncü tərəf kompüterlərdə saxlanan gizlilik məlumatlarını gizli paylaşa bilər. Üçüncüsü, məxfi məlumatların pulla qiymətləndirilməsi PPDM alqoritmlərinin dəqiqliyi və xərcləri üçün istifadə edilə bilər. Belə xərclərə məxfilik sızmaları, hesablama xərcləri və miqyaslama daxildir. Məlumatların pulla dəyəri PPDM istifadəçilərini və məxfi məlumatların sahiblərinin mənfəət və məxflilik arasında başqa balanslaşdırmaya gətirir. Dördüncüsü, şəbəkə monitorinqi və profilləşdirmə üsullarında PPDM səhiyyə və tibbi məlumatlar kimi fərdi məlumatların böyük onlayn axını kimi ortaya çıxır. Məlumatların sayının böyük olması, əlamətlər çoxluğunun böyük ölçüləri, şəbəkə trafik axınlarının dinamik təbiəti şəbəkə monitorinqi gizlilik qorunmasını bir çox digər tətbiqi proqramlarla müqayisədə daha çətin edir [1, 25].

ƏDƏBİYYAT

[1] Aggarwal, C.C. and P.S. Yu. **Privacy-preserving data mining: models and algorithms**. New York: Springer, 2008.

[2] Alguliyev R., Imamverdiyev Y. Big data: big promises for information security // **IEEE 8th International Conference on Application of Information and**

Communication Technologies, 2014.

[3] Batrinca B., Treleaven P. C. Social media analytics: a survey of techniques, tools and platforms // **AI & Society**, 2015, vol. 30, no. 1, pp. 89–116.

[4] Bhuyan, M. H., Bhattacharyya, D. K., & Kalita, J. K. Network anomaly detection: methods, systems and tools // **IEEE Communications Surveys and Tutorials**, 2014, 16(1), 303–336.

[5] **Big Data and Privacy: a technological perspective** // Report to the President (Executive Office of the President President's Council of Advisors on Science and Technology), 2014, 76 pp.

[6] Cardenas A., Manadhata P. K., and Rajan S. P. Big data analytics for security // **IEEE Security & Privacy**, 2013, vol.11, No.6, pp.74-76.

[7] Chen Y., Hwang K., Ku W.S. Collaborative detection of ddos attacks over multiple network domains // **IEEE Transactions on Parallel Distributed Systems**, 2007, vol.18, no.12, pp.1649-1662.

[8] Chen Y. "Research on social media network and national security // **Lecture Notes in Electrical Engineering**, 2013, vol. 205, pp 593-599.

[9] Douligeris C., Mitrokotsa A. DDoS attacks and defense mechanisms: classification and state-of-the-art // **Computer Networks**, 2004, vol.44, no.5, pp.643-666.

[10] Dua, S., Du, X. **Data Mining and Machine Learning in Cybersecurity,** CRC Press**,** Taylor & Francis, 2011, 248 pp.

[11] Əliquliyev R. M., İmamverdiyev Y. N., Abdullayeva F. C. **Sosial şəbəkələr**. Bakı, İnformasiya Texnologiyaları, 2010.

[12] Əliquliyev R. M., Hacırəhimova M.Ş. Big data fenomeni: problemlər və imkanlar // **İnformasiya texnologiyaları problemləri**, 2014, №2, s.3-16.

[13] Əliquliyev R.M., İmamverdiyev Y.N., Yusifov F.F. Cəmiyyətin informasiya təhlükəsizliyinə dair bəzi konseptual baxışlar // **İnformasiya cəmiyyəti problemləri**, 2011, №2(4), s.3-9.

[14] Gulisano V., Callau-Zori M., Fua Z., Jiménez-Peris R., Papatriantafilou M., Patiño-Martínez M. STONE: a streaming DDoS defense framework // **Expert Systems With Applications**, 2015, vol.42, pp.9620-9633.

[15] Hogben G. (ed.), **Security issues and recommendations for online social networks**. ENISA Position Paper No.1, October 2007.

[16] Khondker H. H., Role of the new media in the arab spring // **Globalizations**,

2011, vol. 8, no. 5, pp.675-679.

[17] Montagnese A. **Impact of social media on national security**. Centro Militare di Studi Strategici: Research Paper 2011 STEPI - AE-U-3. 2012.

[18] Pang B., Lee L. Opinion mining and sentiment analysis // **Foundations and Trends in Information Retrieval**, 2008, vol. 2, no. 1-2, pp.1-135.

[19] Peng T, Leckie C, Ramamohanarao K. Survey of network-based defense mechanisms countering the DoS and DDoS problems // **ACM Computing Survey**, 2007, 39(1).

[20] Shameli-Sendi A., Pourzandi M., Fekih-Ahmed M., Cheriet M. Taxonomy of distributed denial of service mitigation approaches for cloud computing // **Journal of Network and Computer Applications**, 2015, vol.58, pp.165-179.

[21] Shiaeles S.N., Katos V., Karakos A.S., Papadopoulos B K. Real time DDoS detection using fuzzy estimators // **Computer Security**, 2012, vol.31, no.6, pp.782-790.

[22] **Social media in strategic communication** (SMISC). http://www.darpa.mil/opencatalog/SMISC.html

[23] Vaidya, J. and C. Clifton. Privacy-preserving data mining: why, how, and when // **IEEE Security and Privacy**, 2004, no.2, pp.19-27.

[24] Verykios, V.S., A. Elmagamid, E. Bertino, Y. Saygin, and E. Dasseni. Association rule hiding // **IEEE Transactions on Knowledge and Data Engineering**, 2004, vol.16, no.4, pp.434-447.

[25] Verykios, V.S., E. Bertino, I.N Fovino, L.P. Provenza, Y, Saygin, and Y. Theodoridis. State of-the-art in privacy preserving data mining // **ACM SIGMOD Record,** 2004, vol.33, no.1, pp.50-57.

[26] Wright D., Hinson M., Examining how public relations practitioners actually are using social media // **Public Relations Journal**, 2009, vol.3, no.3, pp.1-33.

[27] Yen T.-F. et al. Beehive: Large-scale log analysis for detecting suspicious activity in enterprise networks // **Proc. Annual Computer Security Applications Conference**, 2013, pp.199-208.

[28] Zafarani R., Abbasi M.A., Liu H. **Social media mining: an introduction**. Cambridge University, 2014, 382 pp.

5. Layihənin elmi ideyası

Layihənin elmi ideyası informasiya infrastrukturuna olan kütləvi, hədəfyönümlü kiber hücumları və təhdidləri erkən mərhələdə identifikasiya etmək üçün böyük zaman intervalında müxtəlif mənbələrdən toplanmış böyük həcmli məlumatların

analizi üçün big data analitikası texnologiyalarından istifadə etməyi nəzərdə tutur.

Big data analitikası böyük həcmdə verilənlərin emalı nəticəsində informasiya təhlükəsizliyi hadisələri barəsində yeni biliklər, səbəb-nəticə əlaqələri, qanunauyğunluqlar aşkarlamağa, məkan və zamana görə paylanmış, ilk baxışda əlaqəsiz görünən faktları əlaqələndirməyə, uyğunsuzluqlar tapmağa imkan verir. Bu informasiya təhlükəsizliyi obyektlərinin davranışında anomaliyaları aşkarlamağa, hədəfyönümlü hücumlar barəsində erkən xəbərdarlıq etməyə, informasiya təhlükəsizliyi sistemlərində generasiya edilən səhv xəbərdarlıqların sayını əhəmiyyətli dərəcədə azaltmağa şərait yaradardı. Big Data texnologiyaları əsasında dövlət miqyasında informasiya təhlükəsizliyinin vəziyyətini real zamanda mərkəzləşdirilmiş şəkildə monitorinq etmək, meydana çıxan situasiyalar üzrə əvvəlki təcrübə əsasında effektiv qərarlar qəbul etmək mümkündür. Big data texnologiyaları sürətli emal və məlumatların müxtəlif növlərinin analizi ilə anomaliyaların real zamanda aşkarlanmasına imkan verə bilər.

Hazırda informasiya təhlükəsizliyi sistemlərinin çoxu siqnaturaların və məlum təhdidlərin davranış modellərinin axtarışına əsaslanır. Belə sistemlər siqnaturaları hələlik məlum olmayan yeni hücumlara, o cümlədən APT-hücumlarına qarşı gücsüzdürlər. Big Data texnologiyaları əsasında bu hücumların vaxtında aşkarlanmasına ümid edilir.

Dövlətin, cəmiyyətin və fərdlərin informasiya təhlükəsizliyinə ən böyük təhdid mənbələrindən biri də sosial mediadır. Bir çox halda informasiya təhlükəsizliyinin pozulması barəsində ilkin anonsları sosial mediada aparılan müzakirələrdə, postlarda və s. sezmək mümkündür. Sosial medianı müntəzəm analiz etməklə, informasiya təhlükəsizliyinə təhdidləri vaxtında aşkarlamaq və qərar qəbul edən şəxsləri əvvəlcədən xəbərdar etmək olar. Məlumdur ki, sosial mediada toplanan informasiyanın həcmi həddindən artıq böyükdür və onu ənənəvi texnologiyaların köməyilə analiz etmək mümkün deyildir. Ona görə də sosial medianı informasiya təhlükəsizliyi baxımından analiz etmək üçün Big Data Analitikası texnologiyalarından istifadə olunması təklif edilir. Bu məqsədlə lazım gəldikdə mövcud yanaşmaların imkanlarından istifadə etmək, onları təkmilləşdirmək, həmçinin yeni yanaşmalar və texnologiyaların təklif edilməsi nəzərdə tutulur. Sosial media mənbələrinin qabaqcıl text mining üsulları ilə analizi bu sahədə xüsusi potensiala malikdir.

Məlumdur ki, fərd barəsində toplanan məlumatların həcmi və çeşidi artıqca fərdin özəl həyatı "şəffaflaşır". Fərdi məlumatların "şəffaflaşmasına" qarşı dayanıqlı intellektual analiz metodları və alqoritmlərinin işlənilməsi nəzərdə tutulur.

6. Layihə üzrə tədqiqatın strukturu

Layihə **12** mərhələdən ibarətdir:

1-ci mərhələ:

- Big Data sahəsində aparılan elmi tədqiqatların müasir vəziyyətinin analizi və tətbiq sahələrinin tədqiqi

2-ci mərhələ:
- İnformasiya təhlükəsizliyi sahəsində Big Data texnologiyalarının multidisiplinar problemlərinin tədqiqi
- Müxtəlif Informasiya təhlükəsizliyi obyektlərində toplanmış böyük həcmli verilənlərdə anomaliyaların aşkarlanması üçün metod və alqoritmlərin işlənilməsi

3-cü mərhələ:
- Müxtəlif Informasiya təhlükəsizliyi obyektlərində toplanmış böyük həcmli verilənlərdə anomaliyaların aşkarlanması üçün metod və alqoritmlərin işlənilməsi

4-cü mərhələ:
- Müxtəlif Informasiya təhlükəsizliyi obyektlərində toplanmış böyük həcmli verilənlərdə anomaliyaların aşkarlanması üçün metod və alqoritmlərin işlənilməsi
- Milli informasiya infrastrukturuna olan DDoS hücumların aşkarlanması üçün metod və alqoritmlərin işlənilməsi

5-ci mərhələ:
- Müxtəlif Informasiya təhlükəsizliyi obyektlərində toplanmış böyük həcmli verilənlərdə anomaliyaların aşkarlanması üçün metod və alqoritmlərin işlənilməsi
- Milli informasiya infrastrukturuna olan DDoS hücumların aşkarlanması üçün metod və alqoritmlərin işlənilməsi
- Sosial medianın analizi əsasında milli təhlükəsizliyə təhdidlərin və cəmiyyətdə anomal proseslərin aşkarlanması üçün text mining yanaşmalarının təklif edilməsi

6-cı mərhələ:
- Milli informasiya infrastrukturuna olan DDoS hücumların aşkarlanması üçün metod və alqoritmlərin işlənilməsi
- Sosial medianın analizi əsasında milli təhlükəsizliyə təhdidlərin və cəmiyyətdə anomal proseslərin aşkarlanması üçün text mining yanaşmalarının təklif edilməsi

7-ci mərhələ:
- Milli informasiya infrastrukturuna olan DDoS hücumların aşkarlanması üçün metod və alqoritmlərin işlənilməsi
- Sosial medianın analizi əsasında milli təhlükəsizliyə təhdidlərin və cəmiyyətdə anomal proseslərin aşkarlanması üçün text mining yanaşmalarının təklif edilməsi;

8-ci mərhələ:
- Sosial medianın analizi əsasında milli təhlükəsizliyə təhdidlərin və cəmiyyətdə anomal proseslərin aşkarlanması üçün text mining yanaşmalarının təklif edilməsi
- Hədəfyönümlü hücumların aşkarlanması və analizi üçün modellərin işlənilməsi;

9-cu mərhələ:
- Hədəfyönümlü hücumların aşkarlanması və analizi üçün modellərin işlənilməsi;
- Məxfiliyi təmin etməklə fərdi məlumatların intellektual analizi üçün metod və alqoritmlərin işlənilməsi
- Təklif olunmuş metod və alqoritmlərin korporativ şəbəkələrdə tətbiqii

10-cu mərhələ:

- Hədəfyönümlü hücumların aşkarlanması və analizi üçün modellərin işlənilməsi;
- Məxfiliyi təmin etməklə fərdi məlumatların intellektual analizi üçün metod və alqoritmlərin işlənilməsi
- Təklif olunmuş metod və alqoritmlərin korporativ şəbəkələrdə tətbiqi

11-ci mərhələ:

- Hədəfyönümlü hücumların aşkarlanması və analizi üçün modellərin işlənilməsi;
- Məxfiliyi təmin etməklə fərdi məlumatların intellektual analizi üçün metod və alqoritmlərin işlənilməsi
- Təklif olunmuş metod və alqoritmlərin korporativ şəbəkələrdə tətbiq edilməsi

12-ci mərhələ:

- Təklif olunmuş metod və alqoritmlərin korporativ şəbəkələrdə tətbiqi.

7. Layihədən gözlənilən nəticələr, onların elmi və təcrübi əhəmiyyəti

Layihə çərçivəsində aşağıdakı elmi nəticələrin əldə edilməsi gözlənilir:

Böyük həcmli verilənlərdə normal davranış profillərini və onlar əsasında anomaliyaların aşkarlanması üçün metod və alqoritmlər;

Big Data Analitikası metodlarının köməyilə informasiya infrastrukturuna olan kütləvi kiber hücumların erkən mərhələdə aşkarlanması üçün yanaşmalar;

İnformasiya təhlükəsizliyinə olan təhdidlərin aşkarlanması məqsədilə sosial medianın monitorinqi, məlumatların toplanması və intellektual analizi üçün metodlar;

Müxtəlif mənbələrdən kritik informasiya obyektlərinə hücum artefaktlarının aşkarlanması, onların aqreqasiyası və analizi üçün Big Data Analitikası metodları;

Böyük həcmli fərdi məlumatların intellektual analizi zamanı məxfiliyi təmin etmək üçün metodlar və alqoritmlər.

8. Layihə üzrə tədqiqatın nəticələrinin istifadəsi və tətbiqi mümkün olan sahələr

Layihə çərçivəsində əldə ediləcək nəticələr təkcə informasiya təhlükəsizliyi sistemlərində deyil, böyük həcmli verilənlərin toplandığı digər sahələrdə: tibbdə, maliyyə-bank sistemində, molekulyar biologiyada, əczaçılıqda, neft-qaz sənayesində və s. istifadə oluna bilər. Layihə çərçivəsində əldə edilmiş elmi nəticələr və təcrübə big data-nın yaratdığı yeni elmi istiqamətdə – Data Science (Verilənlər haqqında elm) sahəsində kadrların hazırlanmasında istifadə edilə bilər. Təklif olunmuş metod və alqoritmlər əsasında yeni proqram məhsulları hazırlana bilər.

9. Elmi kollektivin xarakterizə edilməsi

Layihə rəhbəri və icraçıların ixtisas və elmi-tədqiqat fəaliyyətini əks etdirən məlumat aşağıdakı cədvəldə verilmişdir:

№	İştirakçılar	İxtisası, elmi dərəcəsi və elmi adı	Uyğunluq dərəcəsi	Qeyd
1	Əliquliyev Rasim	Mühəndis-sistemotexnik; texnika üzrə elmlər	İnformasiya təhlükəsizliyi, sosial şəbəkələr, data	555 elmi əsərin

AZƏRBAYCAN RESPUBLİKASININ PREZİDENTİ YANINDA
ELMİN İNKİŞAFI FONDU

QRANT LAYİHƏLƏRİ MÜSABİQƏSİ

	Məhəmməd oğlu	doktoru, professor, AMEA-nın həqiqi üzvü	mining, text mining, web mining, elektron dövlət, big data sahəsində mütəxəssis	müəllifi
2	Alıquliyev Ramiz Məhəmməd oğlu	Riyaziyyatçı; texnika üzrə elmlər doktoru	Text mining, web mining, data mining, sosial şəbəkələr, big data sahəsində mütəxəssis	128 elmi əsərin müəllifi
3	Rüstəmov Raif Məmməd oğlu	Riyaziyyatçı və kompüter mühəndisi; Riyaziyyat üzrə fəlsəfə doktoru	Big data, data mining, deep learning, qraflar nəzəriyyəsi, həndəsi modelləşmə	31 elmi əsərin müəllifi
4	İmamverdiyev Yadigar Nəsib oğlu	Riyaziyyatçı-mühəndis; texnika üzrə fəlsəfə doktoru	İnformasiya təhlükəsizliyi, biometrik texnologiyalar, audio mining və big data sahəsində mütəxəssis	120 elmi əsərin müəllifi
5	Şıxəliyev Ramiz Hüseyn oğlu	Mühəndis-elektrik; texnika üzrə fəlsəfə doktoru	Kompüter şəbəkələrinin monitorinqi və informasiya təhlükəsizliyi sahəsində mütəxəssis	50 elmi əsərin müəllifi
6	Yusifov Fərhad Firudin oğlu	Mühəndis-elektrik; texnika üzrə fəlsəfə doktoru	Web mining və big data sahəsində mütəxəssis	44 elmi əsərin müəllifi
7	İsayev Orxan Rasim oğlu	Tibb üzrə fəlsəfə doktoru	Molekulyar biologiya; Histologiya, sitologiya və hüceyrə biologiyası; onkologiya sahəsində mütəxəssis	5 elmi əsərin müəllifi
8	Hacırəhimova Məkrufə Şərif qızı	Mühəndis-elektrik; texnika üzrə fəlsəfə doktoru	Text mining; e-dövlət və big data sahəsində mütəxəssis	46 elmi əsərin müəllifi
9	Suxostat Lüdmila Valentinovna	Riyazi modelləşdirmə; texnika üzrə fəlsəfə doktoru	Biometrika və audio mining texnologiyaları sahəsində mütəxəssis	16 elmi əsərin müəllifi
10	Nəbiyev Babək Rasim oğlu	Mühəndis-elektrik	Kompüter şəbəkələrinin monitorinqi və informasiya təhlükəsizliyi sahəsində mütəxəssis	10 elmi əsərin müəllifi
11	Niftəliyeva	Kompüter elmləri	Elektron dövlət, sosial	3 elmi

AZƏRBAYCAN RESPUBLİKASININ PREZİDENTİ YANINDA
ELMİN İNKİŞAFI FONDU

QRANT LAYİHƏLƏRİ MÜSABİQƏSİ

	Günay Yavər qızı		şəbəkələr və text mining sahəsində mütəxəssis	əsərin müəllifi
12	İsazadə Nicat Ramiz	Kompüter elmləri	Text mining sahəsində mütəxəssis	7 elmi əsərin müəllifi

10. Layihə üzrə elmi-tədqiqat işinin yerinə yetirilməsi üçün lazım olan avadanlıq, cihaz və qurğulardan mövcud olanlar haqqında məlumat, əlavə lazım olanların əsaslandırılması

11. Layihə rəhbərinin və icraçıların Elmin İnkişafı Fondu tərəfindən maliyyələşdirilmiş və/və ya hal-hazırda maliyyələşdirilən layihələrdə və digər ölkədaxili, regional və beynəlxalq qrant müsabiqələrində iştirakı barədə məlumat

№	S. A. A.	Hal-hazırda maliyyələşdirilən layihələr	Maliyyələşdirilmiş layihələr
1	Əliquliyev Rasim Məhəmməd oğlu	EİF-2014-9(24)-KETP	EİF-RİTN-MQM-2/İKT-2-2013-7(13)-29/20/1 EİF-2011-1(3)-82/07/1
2	Alıquliyev Ramiz Məhəmməd oğlu	EİF-2014-9(24)-KETP	EİF-RİTN-MQM-2/İKT-2-2013-7(13)-29/20/1 EİF-2011-1(3)-82/07/1
3	İmamverdiyev Yadigar Nəsib oğlu		EİF-RİTN-MQM-2/İKT-2-2013-7(13)-29/18/1
4	Şıxəliyev Ramiz Hüseyn oğlu		EİF-RİTN-MQM-2/İKT-2-2013-7(13)-29/27/1
5	Yusifov Fərhad Firudin oğlu		EİF/GAM-2013-2(8)
6	Suxostat Lüdmila Valentinovna		EİF-RİTN-MQM-2/İKT-2-2013-7(13)-29/18/1

Təqdim olunan layihə həmin layihələrlə uzlaşır və bir-birini tamamlayır

Layihə rəhbəri	*(soyadı, adı, atasının adı)* Alıquliyev Ramiz Məhəmməd oğlu	*İmza, tarix* 21.01.2016