Big Data texnologiyalarının təhlükəsizlik problemləri və həlli yolları

Məkrufə Hacırəhimova[1], Aybəniz Əliyeva[2]
[1,2]İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan
[1]makrufa@science.az, [2]aliyeva.a.s@mail.ru

***Xülasə*—Big data texnologiyalarının meydana gəlməsi verilənlərin təhlükəsizliyi və məxfiliyi üçün yeni problemlər yaradır. Ənənəvi texnologiyalar və metodlar bu problemlərin həlli üçün uyğun və səmərəli deyildir. İşdə böyük verilənlərin təhlükəsizlik problemləri tədqiq olunur. Böyük verilənlərin təhlükəsizliyi və məxfiliyi üçün istifadə olunan texoloji həllər təhlil olunur və onların cari vəziyyəti nəzərdən keçirilir.**

Açar sözlər— big data; verilənlərin təhlükəsizliyi; infrastrukturun təhlükəsizliyi; verilənlərin gizliliyi

I. GİRİŞ

Hazırda şəbəkə və informasiya texnologiyalarının inkişafı verilənlərin həcminin sürətlə artmasına səbəb olur. Müxtəlif üsullarla əldə edilmiş bu böyük verilənlər (ing. Big data) biznes, telekommunikasiya, səhiyyə, maliyyə, informasiya texnologiyaları, təhsil və s. daxil olmaqla bir çox sahələrdə mühüm rol oynayır. Böyük verilənlərin (BV) səmərəli emalı müəssisələrin rəqabət üstünlüyünün artmasına, bir çöx sosial və iqtisadi sahələrdə dəyərin yüksəlməsinə imkan verir. Maksimum gəlirlərin əldə edilməsi məqsədilə dövlət orqanları tərəfindən Big data layihələrinə böyük miqdarda investisiyalar qoyulur, özəl sektorlar gəlirlərin artmasına və resursların optimallaşdırılmasına güclü səy göstərirlər. Lakin Big data konteksində informasiyanın təhlükəsizliyi və gizliliyinə risklərin olması bu texnologiyaların tətbiqini məhdudlaşdırır və inkişafına mane olur [1, 2].

Ümumiyyətlə, verilənlərin təhlükəsizliyi hər zaman vacib məsələlərdən hesab olunmuşdur. Lakin Big data-nın meydana gəlməsi ilə informasiya təhlükəsizliyi baxımından yeni problemlər yaranmışdır. BV-nin mənbələrinin müxtəlifliyi, axın şəklində informasiyanın toplanması və böyük həcmdə informasiyanın “buludlara” miqrasiyası kimi xüsusiyyətləri təhlükəsizlik və gizlilik problemlərini daha da kəskinləşdirmişdir [3, 4]. Bu problemlər verilənlərin toplanması, saxlanması və idarə edilməsi, ötürülməsi və təhlili kimi bütün mərhələləri əhatə edir.

Big data texnologiyalarının informasiya təhlükəsizliyi məsələsi beynəlxalq aləmin, elmi dairələrin diqqətindən kənarda qalmamış, əsas tədqiqat mövzusuna çevrilmişdir. Belə ki, BV-nin təhlükəsizliyinin mühafizəsi üzrə fəaliyyətin standartlaşdırması məsələlərinə Standartlaşdırma üzrə beynəlxalq təşkilat (International Organization for Standardization) və Beynəlxalq elektrotexniki komissiya (International Electrotechnical Commission) (ISO/IEC), Beynəlxalq Telekommunikasiya İttifaqı (International TelecommunicationUnion, ITU), Britaniya Standartlar İnstitutu (British Standards Institution, BSI), ABŞ-ın Milli Standartlar və Texnologiyalar İnstitutu (National Institute of Standards and Technology, NIST) kimi bir sıra aparıcı standartlaşdırma institutları cəlb edilmişdir [5].

Dünyanın bir çox ölkələri BV-nin təhlükəsizliyinin təmini üçün bir sıra qanunları tədqiq etmişdir. ABŞ 2012-ci ilin mart ayında Böyük verilənlərin təhlükəsizliyi ilə bağlı inkişaf strategiyalarını dərc edən ilk ölkə olmuşdur. Həmin ilin iyun ayında Britaniya hökuməti gizliliyin qorunması normaları üçün özəl olaraq “Open Data White Paper” icmali nəşr etdirmişdir. Fransa, Yaponiya, Avstraliya və digər ölkələr böyük verilənlərin təhlükəsizliyinin inkişaf strategiyasını qəbul etdilər [1].

Böyük verilənlər əsas tədqiqat mövzusu olmasına baxmayaraq, mövcud təhlükəsizlik sxemləri tətbiqi proqramların tələblərinə tam cavab verə bilməz. Son illər tədqiqatçılar tərəfindən mövcud təhlükəsizlik sxemləri təkmilləşdirilmiş və ya yeni təhlükəsizlik texnologiyaları irəli sürülmüşdür.

İşin məqsədi BV-nin təhlükəsizliyi və məxfiliyi problemlərinin araşdırılması və bu problemlərin həlli üçün təklif edilmiş texnologiyaların və metodların cari vəziyyətini analiz etməkdən ibarətdir.

II. BÖYÜK VERİLƏNLƏRİN TƏHLÜKƏSİZLİK PROBLEMLƏRİ

İnternet mühitində rəqəmsal informasiyanın həcmi artıqca onun istifadəçilərinin sayı da çoxalır. Nəticədə informasiyanın məxfiliyinə və təhlükəsizliyinə olan təhdidlər artır, yeni təhlükəsizlik problemləri yaranır [6]. Bulud Təhlükəsizlik Alyansı (CSA – Cloud Security Alliance) tərəfindən “Big data” sistemlərinin on təhlükəsizlik problemi [7] verilmiş və onlar infrastrukturun təhlükəsizliyi, verilənlərin gizliliyi, verilənlərin idarə edilməsi, tamlıq və reaktiv təhlükəsizlik kimi dörd qrup üzrə təsnifatlandırılmışdır [8].

A. İnfrastrukturun Təhlükəsizliyi Problemləri

DOI: 10.25045/NCInfoSec.2018.52

1) Paylanmış hesablamaların və digər proseslərin təhlükəsizliyinin təmini

Hadoop platformasının MapReduce kimi paylanmış strukturunun hesablama və emal elementlərinin təhlükəsizliyi əsas etibarilə mühafizə olunmur. Bunun üçün kartoqrafların təhlükəsizliyi, “kartoqraflar” icazəsiz olduqda isə verilənlərin mühafizəsi tədbirləri həyata keçirilməlidir.

2) Qeyri-relyasion verilənlər anbarlarının məxfiliyinin qorunması

NoSQL kimi verilənlər bazaları gizlilik təhdidlərinə səbəb olan bir çox təhlükəsizlik qüsurlarına malikdir. Təhlükəsizliyin bu qüsurları verilənlərin teqləşdirilməsi və ya qeyd edilməsi, həmçinin ötürülməsi və ya toplandıqdan sonra müxtəlif qruplara paylanması zamanı verilənlərin şifrələnməməsi ilə bağlıdır.

B. Verilənlərin Gizliliyi Problemləri

1) Gizliliyi saxlayan miqyaslanan və kompozit data mining və analitika

Kiçik bir məlumat sızması və ya platformada boşluq aralığının olması verilənlərin böyük itkisinə səbəb ola bilər. Nəticədə BV-nin analitikasında problemlər yarana bilər.

2) Girişə nəzarətin əlaqə və şifrələmə metodlarının təhlükəsizliyinin təmini

Verilənlərin mühafizəsinin sadə üsulu bu verilənlərin saxlanması qurğularının, platformalarının təhlükəsizliyinin təminidir. Çünki verilənlərin saxlanması qurğuları təhlükələrə daha həssasdır. Ona görə də şifrələmə metodlarının təhlükəsizliyinin təmin olunması vacibdir.

3) Qranular giriş nəzarəti

NoSQL verilənlər bazası və ya Hadoop paylanmış fayl sisteminin böyük verilənlər anbarlarının qranular giriş nəzarəti üçün güclü autentifikasiya prosesi və məcburi giriş nəzarətinin olması əsas tələblərdəndir.

C. Verilənlərin Idarə Edilməsi Problemləri

1) Verilənlərin saxlanması və ötürülməsi jurnallarının mühafizəsi

İnformasiya daşıyıcılarında saxlanılan tranzaksiya jurnalları və digər məxfi verilənlər müxtəlif səviyyələrə malik ola bilərlər, lakin bu kifayət deyildir. Nümunə üçün, verilənlərin bu səviyyələr arasında trasferi zamanı İT meneceri yerini dəyişən verilənlər barəsində məlumatlı olur. Verilənlərin həcmi fasiləsiz artıqca, miqyaslılıq və əlyetərlik böyük verilənlərin saxlanmasının idarə edilməsi üşün avtomatik çoxsəviyyəliyi (auto-tiering) zəruri edir. Bununla belə, “auto-tiering” üsulu verilənlərin saxlanma yerinə nəzarət etmədiyindən, BV-nin saxlanması üçün yeni problemlər yaranır. Ona görə də bu saxlanclar icazəsiz girişdən qorunmalı və onların əlyetərliyi daim təmin etdilməlidirl [9].

2) Qranular audit

Müxtəlif növ jurnalların düzgün analizi faydalı ola bilər və bu informasiya bütün kiberhücumların və ya zərərli fəaliyyətin (casusluğun) aşkarlanmasında istifadə edilə bilər. Buna görə verilənlərin fasiləsiz yoxlanması ilə yanaşı onların müntəzəm surətdə auditi də çox vacibdir.

3) Verilənlərin mənşəyinin təyini problemləri

Verileniərin təsnifləndirilməsi zamanı onların mənbəyinin bilinməsi zəruridir. Verilənlərin mənşəyinin dəqiq müəyyən edilməsi onun identifikasiyası, yoxlanması və girişə nəzarəti ilə əldə edilə bilər.

D. Tamlıq və reaktiv təhlükəsizlik problemləri

1) Girişlərin yoxlanması və filtrasiyası

Giriş qurğuları BV-nin qorunmasında əsas yer tutur. Saxlanma, emal və digər zəruri işlər giriş qurğularından daxil olan ilkin verilənlərin köməyi ilə həyata keçirilir. Buna görə də, bir təşkilat etibarlı və qanuni giriş qurğularının istifadə olunmasını təmin etməlidir. Hər bir şəbəkə zərərli *girişlərdən* azad olmalıdır [9].

2) Real-vaxt rejimində verilənlərin təhlükəsizliyinin və monitorinqinin təmini

Təhlükəsizlik yoxlamaları və monitorinqin real vaxt rejimində və ya real vaxta olduqca yaxın digər bir rejimdə yerinə yetirilməsi daha uyğun hesab edilir. Qeyd edək ki, ənənəvi platformaların çox hissəsi generasiya olunan verilənlərin böyük həcmi səbəbindən bunu edə bilmirlər.

III. BIG DATA TEXNOLOGIYALARININ TƏHLÜKƏSIZLİYİNİN TƏMİN EDİLMƏSİ METODLARI

Böyük verilənlərin təhlükəsizliyi və məxfiliyi ilə bağlı yuxarıda qeyd olunan problemlərin həlli üçün tədqiqatçılar tərəfindən müxtəlif metodlar təklif olunmuşdur. Aşağıda bu metodlarının bəziləri nəzərdən keçirilmişdir.

A. Verilənlərin məxfiliyinin qorunması metodları

Böyük verilənlərin gizliliyini qorumaq üçün verilənlərin şifrələmə və anonimliyi texnologiyasından istifadə edilir.

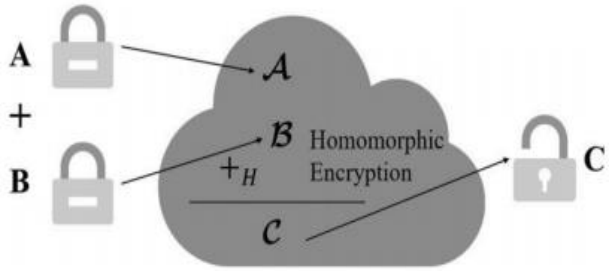
1*) Verilənləri şifrləmə texnologiyaları*

a) Homomorfik şifrləmə: Verilənləri şifrləmə mütodu verilənlərin gizliliyini qorumaq üçün əhəmiyyətli bir vasitədir. Böyük verilənlər sisteminin məlumatı sürətli və effektiv emaletmə qabiliyyəti şifrələmə üçün tələb olunan aparat və proqram təminatının tələblərinə uyğundur və verilənlərin gizliliyinin qorunmasında homomorfik şifrləmənin istifadəsini mümkün edir [10].

Homomorfik şifrləmənin konsepsiyası ilk dəfə R.L.Rivest və həmkarları tərəfindən daxil edilmişdir. Bu metodun üstünlüyü ondan ibarətdir ki, şifrlənmiş mətinlər deşifrə edilmədən, hesablamanın nəticələrinə təsir etmədən birbaşa əməliyyat apara bilirlər. Şəkil 1-də bulud mühitində homomorfik şifrələmə üçün konkret nümunə göstərilmişdir. Şəkildən göründüyü kimi A və B mətnlərinə homoqrafik şifrləmə tətbiq edilir və onlar bulud mühitinə göndərilir.

Homoqrafik şifrləmədən sonra şifrlənmiş mətn deşifrə olunur və düzgün nəticə C qeyd olunur [1].

Ümumiyyətlə, homomorfik şifrləməni şərti olaraq yarımhomomorfik (semi-homomorphic encryption, SHE) və tam homomorfik (fully homomorphic encryption, FHE) kimi iki növə bölürlər. SHE sxemləri yalnız toplama və vurma kimi sadə əməliyyatları dəstəkləyir. İlk dəfə C.Gentry tərəfindən təklif edilmiş FHE sxemi isə ixtiyari çoxhədlinin hesablanmasını dəstəkləyir. FHE sxeminin BGV, Bral2, TSZ, və GSW1321 kimi tanınmış bir çox sxemləri də Big data-nın təhlükəzliyi və məxfiliyinin təmini üçün istifadə olunmuşdur [1, 11].



Şəkil 1 Bulud mühitində homomorfik şifrləmə

Q. Zhang və həmkarları Big data-nın klasterləşməsinin təhlükəsizliyinin təmini üçün PCM (possibilistic cmeans calculation) və CFS (clustering by fast search) klasterləşmə alqoritmlərinə BGV sxemini tətbiq etmişdir. Bundan əlavə, onlar bulud mühitində fərdi verilənlərin şifrələnməsi üçün BGV-nin istifadəsilə məxfiliyi təmin edən iki dərin hesablama modeli hazırlamışdılar [1, 12, 13].

A. Rəhmani və həmkarları təkamül kriptoqrafiyası və TSZ sxemini birləşdirən yeni çoxlaylı metod təklif etmişlər. Bu metodlar şifrələnmiş BV-nin təhlükəsizliyini yüksək səviyyədə təmin etməklə bir-birini tamamlayırlar. Eyni zamanda, onlar orijinal metodun şifrlənmiş mətninin uzunluğunu artırır və emal və analiz zamanı şifrlənmiş mətnlərin istifadə edilməsinə imkan verir [14].

Lakin homomorfik şifrləmə sxemlərində küyə nəzarət etmək çətindir və bulud mühitində aşağı səmərəliliyə malik olurlar. Belə ki, homomorfik şifrləmə bəzi tətbiqi sistemlərin tələblərinə cavab vermir.

b) Təhlükəsiz çoxpartiyalı hesablama (secure multiparty computation, SMC): SMC sxemi ilk dəfə 1982-ci ildə A.C.Yao-nun tədqiqatlarında təklif olunmuşdur. Bu sxem tərəflərə paylanmış bulud mühitində öz fərdi (gizli) girişləri ilə funksiyanı hesablamağa imkan verir. Son illər təhlükəsizlik məsələləri üçün SMC metoduna əsaslanan yeni sxemlər təklif etdilmişdir.

A. Patra və həmkarlarının [15] təklif etdiyi təhlükəsiz asinxron çoxpartiyalı hesablama (asynchronous multiparty computation) sxemi statistik olaraq təhlükəsiz və asinxronlaşdırıla bilən gizli paylaşmanı təmin edir.

A.Y. Sarhan və S.Carr [16] tərəfindən təklif edilmiş təhlükəsizlik sxemi (self-protection data scheme) bulud mühitində verilənlərin sahiblərinə kənar istifadəçilərin ötürdükləri verilənlər üzərində mütləq nəzarətə malik olmasına imkan verir. Müəlliflər girişə nəzarətin təhlükəsizliyi məqsədilə atributlar əsasında şifrələmə, açarları idarəetmə üçün RSA, verilənlərin özünü qoruması üçün aktiv verilənlər paketləri (active data bundles) və trafikin idarə edilməsi üçün müstəqil mobil agentləri kimi dörd inkişaf etmiş texnologiyanı birləşdirmişdir.

[17]-də istifadəçi verilənlər bazasının təhlükəsizliyinin təmini üçün gizli açarın paylaşılmasında SMC sxemini istifadə edən çoxpartiyalı paylaşma sistemi təklif edilmişdir.

[18]-da çoxmənbəli verilənlərin təhlükəsizliyinin analizini dəstəkləmək üçün istifadə edilən gizliliyi qoruyan statistik analiz mühiti, yəni Rmind təqdim edilmişdir.

c) Atributlar əsasında şifrləmə (Attribute-based encryption, ABE): Atributlar əsasında şifrələmə ilk dəfə, A. Sahai və, B. Waters [19] tərəfindən təklif edilmişdir. Bu üsul verilənlərə təhlükəsiz giriş üçün ən uyğun və populyar şifrləmə üsulu olmuşdur [20] İdentifikasiyaya əsaslanan şifrələmədən fərqli olaraq, ABE-də verilənlərin sahibləri həm açarları, həm də şifrələnmiş mətnləri atributlar dəsti ilə qeyd edirlər (nişanlayırlar). Sonra isə istifadəçilər bu atributlara görə hüquqi istifadəçilərin avtorizasiyasını təsdiq etməlidirlər. Bundan sonra, verilənlər açarın və şifrlənmiş mətnin atributlarına uyğun olaraq yüklənir və şifrələnir.

Ənənəvi ABE şifrəli mətn mexanizminə (ciphertext-policy CP-ABE) və açar mexanizminə (key-policy, KP-ABE) əsaslanan iki tipə malikdir. Bu iki tip əsas etibarilə, giriş protokollarına tətbiq edilən üsula görə fərqlənir. CP-ABE sxemində giriş protokolu homoloq şifrəli mətnlə yerləşdirilir. Əksinə, KP-ABE sxemində giriş protokolları istifadəçilərin gizli açarları ilə əlaqəlidir [1].

2) Verilənlərin Anonimlik Metodu

Verilənlərin anonimliyi gizliliyin qorunması üçün digər bir əhəmiyyətli texnologiyadır. Bu halda bədniyyətli gizliliyi özündə saxlayan verilənləri əldə etsə də, açar sahənin qiyməti gizli olduğundan ilkin dəqiq verilənləri əldə edə bilmir. Lakin, BV mühitində bədniyyətli birdən çox mənbələrdən verilənlər əldə edə bilər, daha sonra bir mənbənin verilənlərini digər bir mənbənin verilənləri ilə əlaqələndirə və gizli verilənlərin orijinal mənasını tapa bilər.

B. Girişə Nəzarət Metodu (Access Control Technology)

Girişə nəzarət texnologiyası verilənlərin nəzarət altında birgə istifadəsi üçün səmərəli vasitədir. Lakin BV mühitində istifadəçilərin sayı çoxdur, avtorizasiyalar mürəkkəbdir. Ona görə də verilənlərin nəzarətli paylaşımını reallaşdırmaq üçün yeni texnologiyanın istifadəsinə ehtiyac vardır.

1)Rolların intellektual analizi üsulları (Role Mining technology)

Rollar əsasında girişə nəzarət metodu (Role-based access control, RBAC) geniş istifadəyə malikdir. Istifadəçilərin

razılığını əldə etmək, hüquqların idarə olunmasını sadələşdirmək, gizliliyin qorunmasına nail olmaq üçün icazələrin alınması ilə bağlı istifadəçilərə görə rolu təyin edir. İlkin mərhələdə RBAC hüquqlarının idarə edilməsi "yuxarıdan aşağı" rejimini tətbiq etmişdir. Müəssisənin rolları təyin etmək mövqeyinə uyğun olaraq BV sahəsinə tətbiq edildikdə, tədqiqatçılar "istifadəçilər - obyekt" səlahiyyətlərinə əsaslanan "aşağıdan yuxarı" rejiminə keçməyə başladılar. BV sahəsində rolların intellektual analiz üsullarından istifadə edərək, rollar istifadəçinin giriş qeydlərinə əsasən avtomatik olaraq yaradıla, kütləvi istifadəçilər üçün effektiv şəkildə fərdi məlumat xidmətləri təqdim edə bilir. Lakin Role Mining texnologiyası dəqiq, qapalı verilənlərə əsaslanır və BV-yə tətbiq edildikdə dinamik dəyişikliklərin və verilənlər yığımının keyfiyyətinin yüksək olmaması kimi problemlərin həllinə ehtiyac yaranır [10].

2) Risk adaptiv olan girişə nəzarət metodu (Risk Adaptive Access Control)

BV sahəsində təhlükəsizlik üzrə inzibatçı kifayət qədər təcrübəyə malik olmaya, istifadəçilərin əldə edə biləcəyi verilənləri düzgün müəyyən etməyə bilər. Bu problemin həlli üçün risklərə adaptiv olan girişə nəzarət metodundan istifadə olunur. Risklərə adaptiv girişə nəzarətə nail olmaq üçün statistik metodlarlardan və informasiya nəzəriyyəsindən istifadə etməklə *kvant alqoritmi* müəyyən edilə bilər. Eyni zamanda, BV mühitində riski təyin etmək və kəmiyyətcə qiymətləndirmək daha çətindir [10].

C. Verilənlərin Mənşəyi Texnologiyaları (Data Provenance)

Hələ BV konsepsiyasının meydana gəlməsindən əvvəl verilənlərin mənşəyi texnologiyaları (Data Provenance) verilənlər bazasında geniş şəkildə araşdırılmışdır. Verilənlərin mənşəyi metodu nişana (label) əsaslanır. Nişan vasitəsilə hansı verilənin mənbə olduğunu müəyyən etmək və nəticənin düzgünlüyünü asanlıqla yoxlamaq və ya məlumatları minimum dəyərlə yeniləmək mümkündür. Lakin BV-nin təhlükəsizliyi və gizliliyinin qorunması üçün bu texnologiya 1) gizliliyin qorunması və verilənlərin mənşəyi arasında balans; 2) verilənlərin mənşəyi texnologiyasının özünün təhlükəsizliyinin qorunması kimi iki məsələni həll etməlidir.

NƏTİCƏ

Böyük verilənlər iqtisadi və sosial inkişaf üçün yeni bir vasitədir, iqtisadi əməliyyatlara, həyat tərzinə və milli idarəetmə bacarıqlarına getdikcə daha çox təsir edir. BV-nin həcmi fasiləsiz olaraq artdıqca onun təhlükəsizlik və gizlilik problemləri də artır. Big data texnologiyaların informasiya təhlükəsizliyi və gizliliyinə risklərin olması bu texnologiyaların tətbiqinə və inkişafına mane olur, imkanlarını məhdudlaşdır. Tədqiqatlardan da göründüyü kimi böyük həcmli verilənləri idarə etmək üçün bu günə qədər hazırlanmış alətlər və texnologiyalar verilənlərin təhlükəsizliyini və gizliliyini təmin etmək üçün kifayət qədər effektiv deyildir. Ona görə də BV-nin təhlükəsizliyinə böyük əhəmiyyət verilməli, təhlükəsizlik texnologiyaları tədqiq olunmalı və inkişaf etdirilməlidir. Böyük verilənlərin paylanmış saxlanma, pərakəndə idarəetmə, platformalar arasında paylanması və digər xüsusiyyətləri üçün təhlükəsiz şifrələmə və etibarlı hesablama texnologiyaları işlənilməlidir.

Eyni zamanda, BV-nin açıq olması verilənlərin məxfiliyi riskini artırdığından böyük verilənlərin təhlükəsizliyinin inkişaf strategiyaları və etibarlı təhlükəsizlik sistemlərinin işlənilməsinə ehtiyac vardır.

Beləliklə, Big data inkişaf etməkdə olan bir texnologiyadır və onun təhlükəsizlik risklərinə zəmanət vermək qeyri-mümkündür. Odur ki, təhlükəsizlik texnologiyalarının səmərəliliyinin yaxşılaşdırılması gələcək illərdəki tendensiyalardan biri olaraq qalacaqdır.

ƏDƏBİYYAT

[1] R, Bao, Z, Chen and M. S. Obaidat, “Challenges and techniques in Big data security and privacy: A review,” Security and Privacy, vol. 1, no. 4, 2018, pp. 1-8.

[2] Fatima-Z. Benjelloun and A. L. Ayoub, “Big Data Security: Challenges, Recommendations and Solutionsç” Handbook of Research on Security Considerations in Cloud Computing, 2015, pp. 301-313.

[3] R Alguliyev, Imamverdiyev Y. Big Data: Big promises for information security / Proceedings of the IEEE 8th International Conference on Application of Information and Communication Technologies (AICT), Astana, Kazakhstan 15-17 October, 2014, pp.216–219.

[4] R. M. Əliquliyev, M. Ş. Hacırəhimova and A.S. Əliyeva, “Big data-nin aktual elmi-nəzəri problemləri,” İnformasiya cəmiyyəti problemləri, №2, 2016, səh. 37–49.

[5] Д. Пудов, “Проблемы безопасности больших данных,” Открытые системы. СУБД, № 04, 2017, https://www.osp.ru/os/2017/04/

[6] R.Əliquliyev, F.Abdullayeva Bulud texnologiyalarının təhlükəsizlik problemlərinin tədqiqi və analizi // İnformasiya Texnologiyaları Problemləri, №1, 2013, səh. 3–14.

[7] Cloud security alliance. Expanded Top Ten Big Data Security and Privacy Challenges,2013, https://downloads.cloudsecurityalliance.org/

[8] M. Ş. Hacırəhimova. “Big Data” texnologiyalari və informasiya təhlükəsizliyi problemləri, İnformasiya texnologiyaları problemləri, №1, 2016, səh. 49–56.

[9] P. S. Yosepu, “A Study on Security and Privacy in Big Data Processing,” International Journal of Innovative Research in Computer, vol. 3, no. 12, 2015, pp. 12292-12296.

[10] Z. Gang, “Big Data and Information Security,” International Journal of Computational Engineering Research, vol. 5, no. 6, 2015, pp. 17-21.

[11] C. A. Gentry, Fully Homomorphic Encryption Scheme, Palo Alto, USA: Stanford University; 2009, 199 p.

[12] Q. Zhang, L.T. Yang, Z. Chen et al., “PPHOPCM: privacy-preserving high-order possibilistic c-means algorithm for Big data clustering with cloud computing”, IEEE Trans Big Data, 2017, pp.1-11.

[13] Q. Zhang, L.T. Yang, Z. Chen, “Privacy preserving deep computation model on cloud for Big data feature learning,” IEEE Trans Computing, vol. 65, no.5, 2016, pp.1351-1362.

[14] A. Rahmani, A. Amine and R.H. Mohamed, “A multilayer evolutionary homomorphic encryption approach for privacy preserving over Big data,” International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, Shanghai, China: IEEE, 2014, pp.19-26

[15] A. Patra, “Choudhury A, Rangan CP. Efficient asynchronous verifiable secret sharing and multiparty computation,” J. Cryptol., vol. 28, no. 1, 2015, pp. 49-109

[16] A.Y. Sarhan and S. Carr, “A highly-secure self-protection data scheme in clouds using active data bundles and agent-based secure multi-party computation,” International Conference on Cyber Security and Cloud Computing. Columbia University, New York, IEEE, 2014, pp. 228-236.

[17] Dyadic Security White Paper, 2017, 5 p., https://www.dyadicsec.com/.

[18] D. Bogdanov, L. Kamm, S. Laur, et al., “Rmind: a tool for cryptographically secure statistical analysis”, IEEE Trans. Depend Secure Computing, 2016, pp.1-14.

[19] A. Sahai and B. Waters, “Fuzzy identity-based encryption,” International Conference on Theory and Applications of Cryptographic Techniques, French Riviera, France: Springer-Verlag, 2005, pp.457-473.

[20] J.R. Bertini, M. Carmo Nicoletti and L. Zhao, “Attribute-based decision graphs: a framework for multiclass data classification,” Neural Networks, vol. 85, 2017, pp. 69-84.

SECURİTY PROBLEMS OF BİG DATA TECHNOLOGY AND THEİR SOLUTİON WAYS

Makrufa Hajirahimova[1], Aybeniz Aliyeva[2]
[1,2]Institute of Information Technology of ANAS, Baku, Azerbaijan
[1]makrufa@science.az, [2]aliyeva.a.s@mail.ru

Abstract -- The formation of Big data technologies create new challenges for data security and privacy. Traditional technologies and methods are not appropriate and efficient for solving these problems. In this article has been researched security problems of big data.
The technical solutions that used to ensure the security and privacy of big data have been analyzed and their current status has been reviewed.

Keywords -- Big data; data security; infrastructure security; data privacy